

Bases de Datos

Introducción a las Bases de Datos Relacionales y No Relacionales

Lic. Andy Ledesma García
Lic. Víctor M. Cardentey Fundora
Dra. Lucina García Hernández

Departamento de Computación
Facultad de Matemática y Computación
Universidad de La Habana

Licenciatura en Ciencia de Datos

23 de enero de 2024

¿Qué tan grande es YouTube?

- ▶ 500 horas de video subidas cada minuto

¿Qué tan grande es YouTube?

- ▶ 500 horas de video subidas cada minuto
- ▶ 720000 horas de video subidas cada día

¿Qué tan grande es YouTube?

- ▶ 500 horas de video subidas cada minuto
- ▶ 720000 horas de video subidas cada día
- ▶ Se necesitan **82 años** para ver el contenido de un solo día





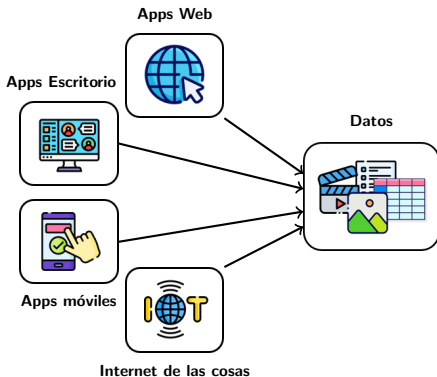
Motivación

Problemas a resolver

Motivación

Problemas a resolver

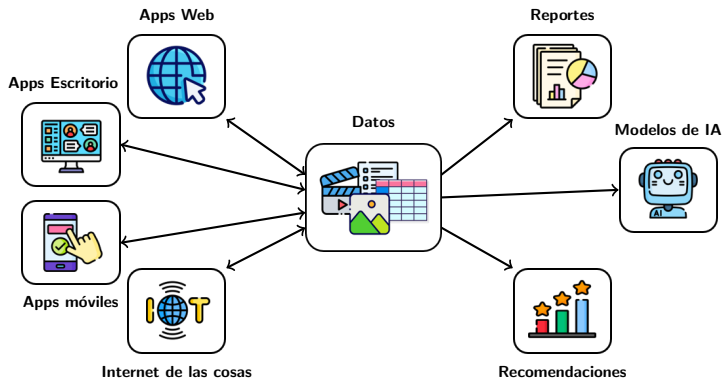
- Garantizar la persistencia de los datos generados por aplicaciones y dispositivos.



Motivación

Problemas a resolver

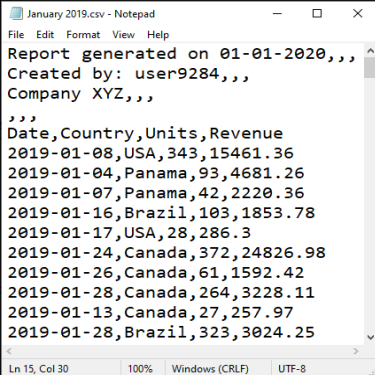
- ▶ Garantizar la persistencia de los datos generados por aplicaciones y dispositivos.
- ▶ Utilizar grandes cantidades de datos de forma eficiente.



¿Cómo?

¿Cómo?

La forma más sencilla de almacenar datos es escribirlos en un fichero



```
January 2019.csv - Notepad
File Edit Format View Help
Report generated on 01-01-2020,,,
Created by: user9284,,,
Company XYZ,,,
,,,
Date,Country,Units,Revenue
2019-01-08,USA,343,15461.36
2019-01-04,Panama,93,4681.26
2019-01-07,Panama,42,2220.36
2019-01-16,Brazil,103,1853.78
2019-01-17,USA,28,286.3
2019-01-24,Canada,372,24826.98
2019-01-26,Canada,61,1592.42
2019-01-28,Canada,264,3228.11
2019-01-13,Canada,27,257.97
2019-01-28,Brazil,323,3024.25
Ln 15, Col 30 100% Windows (CRLF) UTF-8
```

Sistemas orientados a ficheros



Características

- ▶ Se crean nuevos ficheros a medida que se crean nuevos tipos de registros o se eliminan los ficheros.
- ▶ Cada fichero se opera de forma independiente del resto de archivos en almacenamiento.

Sistemas orientados a ficheros



Características

- ▶ Se crean nuevos ficheros a medida que se crean nuevos tipos de registros o se eliminan los ficheros.
- ▶ Cada fichero se opera de forma independiente del resto de archivos en almacenamiento.

Limitaciones

- ▶ Baja eficiencia

Sistemas orientados a ficheros



Características

- ▶ Se crean nuevos ficheros a medida que se crean nuevos tipos de registros o se eliminan los ficheros.
- ▶ Cada fichero se opera de forma independiente del resto de archivos en almacenamiento.

Limitaciones

- ▶ Baja eficiencia
- ▶ Gran redundancia de los datos

Sistemas orientados a ficheros



Características

- ▶ Se crean nuevos ficheros a medida que se crean nuevos tipos de registros o se eliminan los ficheros.
- ▶ Cada fichero se opera de forma independiente del resto de archivos en almacenamiento.

Limitaciones

- ▶ Baja eficiencia
- ▶ Gran redundancia de los datos
- ▶ Pobre control sobre los datos

Sistemas orientados a ficheros



Características

- ▶ Se crean nuevos ficheros a medida que se crean nuevos tipos de registros o se eliminan los ficheros.
- ▶ Cada fichero se opera de forma independiente del resto de archivos en almacenamiento.

Limitaciones

- ▶ Baja eficiencia
- ▶ Gran redundancia de los datos
- ▶ Pobre control sobre los datos
- ▶ Capacidades inadecuadas de manipulación de datos

“... una base de datos es una colección auto-descriptiva de registros integrados.”
—Allen Taylor

“... una base de datos es una colección auto-descriptiva de registros integrados.”
—Allen Taylor

Registro

Datos específicos sobre una entidad u objeto de interés.

*“... una base de datos es una colección **auto-descriptiva** de registros integrados.”*
—Allen Taylor

Registro

Datos específicos sobre una entidad u objeto de interés.

Auto-descriptiva

Se almacenan metadatos (la descripción de su estructura) dentro.

*“... una base de datos es una colección auto-descriptiva de registros **integrados**.”*
—Allen Taylor

Registro

Datos específicos sobre una entidad u objeto de interés.

Auto-descriptiva

Se almacenan metadatos (la descripción de su estructura) dentro.

Integrados

No solo contiene los datos sino también las interrelaciones.

“... una base de datos es una colección auto-descriptiva de registros integrados.”
—Allen Taylor

Registro

Datos específicos sobre una entidad u objeto de interés.

Auto-descriptiva

Se almacenan metadatos (la descripción de su estructura) dentro.

Integrados

No solo contiene los datos sino también las interrelaciones.

Relacionales

Bases de datos

Relacionales

“... una base de datos es una colección auto-descriptiva de registros integrados.”
—Allen Taylor

Registro: datos específicos sobre una entidad u objeto de interés

23082300205	Edgar F.	Codd	23/08/1923
-------------	----------	------	------------

*“... una base de datos es una colección **auto-descriptiva** de registros integradados.”*
—Allen Taylor

Auto-descriptiva: se almacenan metadatos (la descripción de su estructura) dentro del diccionario de datos de la propia base de datos.

Persona

CI : string	Nombre : string	Apellido : string	F. Nacimiento : date
23082300205	Edgar F.	Codd	23/08/1923

*“... una base de datos es una colección auto-descriptiva de registros **integrados**.”*
—Allen Taylor

Integrados: no solo contiene los datos sino también las interrelaciones que se establecen entre estos.

Persona

CI : string	Nombre : string	Apellido : string	F. Nacimiento : date
23082300205	Edgar F.	Codd	23/08/1923

Cuenta

No. Cuenta : string	Balance : decimal	CI Dueño : string
8976	270.98	23082300205

Llave-valor

Bases de datos

Llave-valor

*“... una base de datos es una colección auto-descriptiva de **registros** integrados.”*
—Allen Taylor

Registro: datos específicos sobre una entidad u objeto de interés

```
‘persona_23082300205’: ‘{  
  "CI": 23082300205,  
  "Nombre": "Edgar F.",  
  "Apellido": "Codd",  
  "F. Nacimiento": 23/08/1923  
}’
```

Bases de datos

Llave-valor

“... una base de datos es una colección auto-descriptiva de registros integrados.”
—Allen Taylor

Registro: datos específicos sobre una entidad u objeto de interés

```
{ 'persona_23082300205': {  
  "CI": 23082300205,  
  "Nombre": "Edgar F.",  
  "Apellido": "Codd",  
  "F. Nacimiento": "23/08/1923"  
}}
```

Bases de datos

Llave-valor

*“... una base de datos es una colección auto-descriptiva de **registros** integrados.”*
—Allen Taylor

Registro: datos específicos sobre una entidad u objeto de interés

```
‘persona_23082300205’: ‘{  
    "CI": 23082300205,  
    "Nombre": "Edgar F.",  
    "Apellido": "Codd",  
    "F. Nacimiento": 23/08/1923  
}’
```

Bases de datos

Llave-valor

*“... una base de datos es una colección **auto-descriptiva** de registros integrados.”*
—Allen Taylor

Auto-descriptiva: se almacenan metadatos (la descripción de su estructura) dentro del diccionario de datos de la propia base de datos.

```
{ 'persona_23082300205': {  
    "CI": 23082300205,  
    "Nombre": "Edgar F.",  
    "Apellido": "Codd",  
    "F. Nacimiento": 23/08/1923  
  },
```

Bases de datos

Llave-valor

*“... una base de datos es una colección auto-descriptiva de registros **integrados**.”*
—Allen Taylor

Integrados: no solo contiene los datos sino también las interrelaciones que se establecen entre estos.

```
'persona_23082300205': '{  
  "CI": 23082300205,  
  "Nombre": "Edgar F.",  
  "Apellido": "Codd",  
  "F. Nacimiento":  
    23/08/1923  
}',  
  
'cuenta_8976': '{  
  "No. Cuenta": 8976,  
  "Balance": 270.98,  
  "CI Duenyo": 23082300205  
}'
```

Columnares

Bases de datos

Columnares

*“... una base de datos es una colección auto-descriptiva de **registros** integrados.”*
—Allen Taylor

Registro: datos específicos sobre una entidad u objeto de interés

23082300205
Edgar F.
Codd
23/08/1923

Bases de datos

Columnares

“... una base de datos es una colección auto-descriptiva de registros integrados.”
—Allen Taylor

Registro: datos específicos sobre una entidad u objeto de interés

23082300205	43101123919	44011279165
Edgar F.	Michael	Jim
Codd	Stonebreaker	Gray
23/08/1923	11/10/1943	12/01/1944

Bases de datos

Columnares

*“... una base de datos es una colección **auto-descriptiva** de registros integrados.”*
—Allen Taylor

Auto-descriptiva: se almacenan metadatos (la descripción de su estructura) dentro del diccionario de datos de la propia base de datos.

Persona

CI : string	23082300205
Nombre : string	Edgar F.
Apellido : string	Codd
F. Nacimiento : date	23/08/1923

Bases de datos

Columnares

*“... una base de datos es una colección auto-descriptiva de registros **integrados**.”*
—Allen Taylor

Integrados: no solo contiene los datos sino también las interrelaciones que se establecen entre estos.

Persona

CI : string	23082300205
Nombre : string	Edgar F.
Apellido : string	Codd
F. Nacimiento : date	23/08/1923

Cuenta

No. Cuenta : string	8976
Balance : decimal	270.98
CI Dueño : string	23082300205

de Documentos

Bases de datos

de Documentos

*“... una base de datos es una colección auto-descriptiva de **registros** integrados.”*
—Allen Taylor

```
1 {  
2   "CI": "23082300205",  
3   "Nombre": "Edgar F.",  
4   "Apellido": "Codd",  
5   "F. Nacimiento": "23/08/1923",  
6   "Titulo de Doctor": {  
7     "F. Expedido": 1965,  
8     ...  
9   }  
10 }
```

Bases de datos

de Documentos

*“... una base de datos es una colección **auto-descriptiva** de registros integra-
dos.”* —Allen Taylor

```
1 {  
2   "CI": "23082300205",  
3   "Nombre": "Edgar F.",  
4   "Apellido": "Codd",  
5   "F. Nacimiento": "23/08/1923",  
6   "Titulo de Doctor": {  
7     "F. Expedido": 1965,  
8     ...  
9   }  
10 }
```

Bases de datos

de Documentos

*“... una base de datos es una colección auto-descriptiva de registros **integrados**.”*
—Allen Taylor

Integrados: no solo contiene los datos sino también las interrelaciones que se establecen entre estos.

Persona (llave: “23082300205”)

```
1 {  
2   "CI": "23082300205",  
3   "Nombre": "Edgar F.",  
4   "Apellido": "Codd",  
5   "F. Nacimiento": "23/08/1923"  
6 }
```

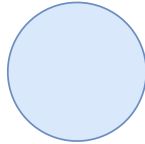
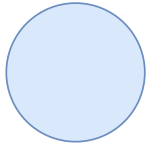
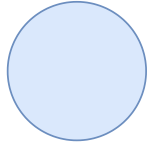
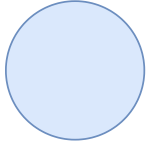
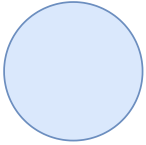
Cuenta (llave: “8976”)

```
1 {  
2   "No. Cuenta": "8976",  
3   "Balance": 270.98,  
4   "CI Duenyo": "23082300205"  
5 }
```

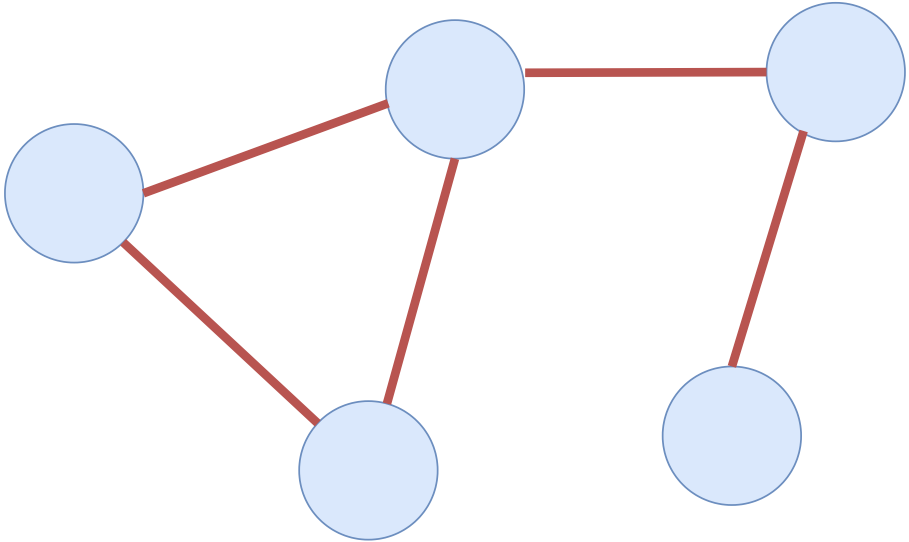
de Grafos

¿Qué es un grafo?

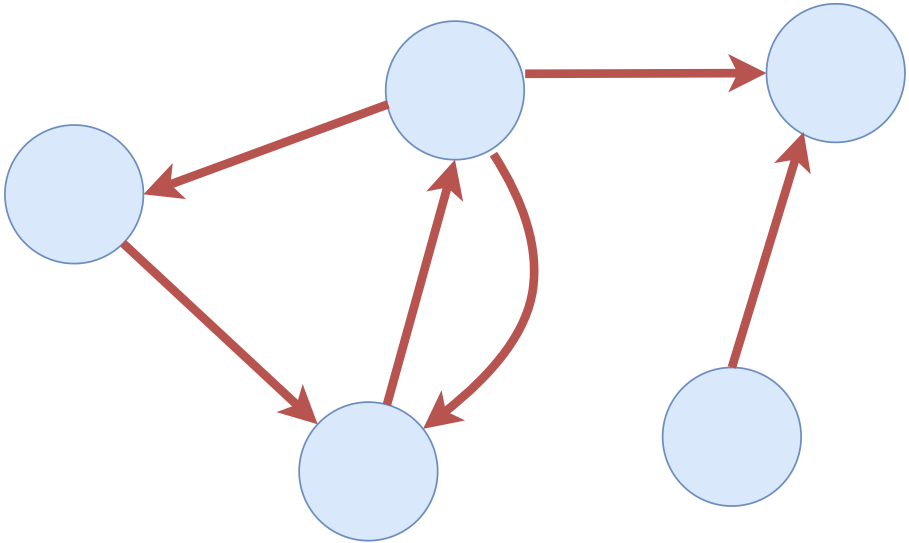
¿Qué es un grafo?



¿Qué es un grafo?



¿Qué es un grafo?

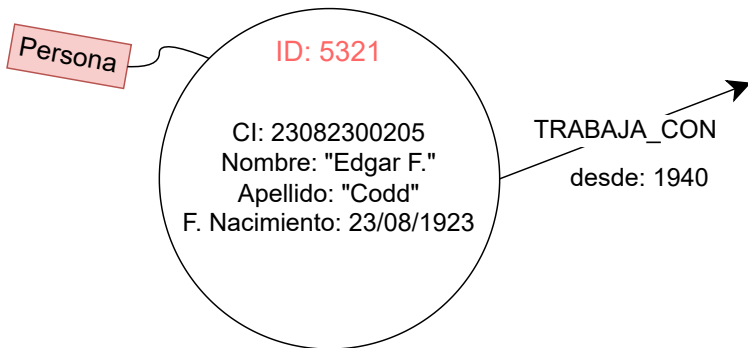


Bases de datos

de Grafos

*“... una base de datos es una colección auto-descriptiva de **registros** integrados.”*
—Allen Taylor

Registro: datos específicos sobre una entidad u objeto de interés

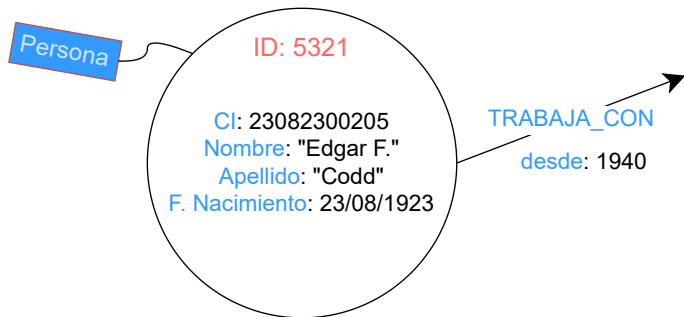


Bases de datos

de Grafos

*“... una base de datos es una colección **auto-descriptiva** de registros integrados.”*
—Allen Taylor

Auto-descriptiva: se almacenan metadatos (la descripción de su estructura) dentro del diccionario de datos de la propia base de datos.

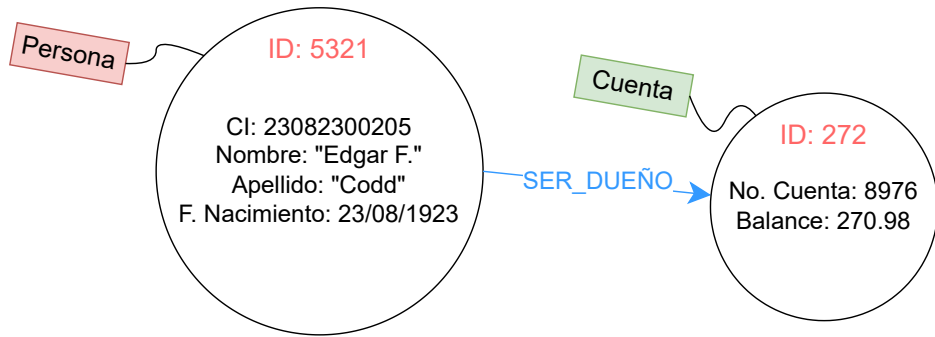


Bases de datos

de Grafos

*"... una base de datos es una colección auto-descriptiva de registros **integrados**."*
—Allen Taylor

Integrados: no solo contiene los datos sino también las interrelaciones que se establecen entre estos.

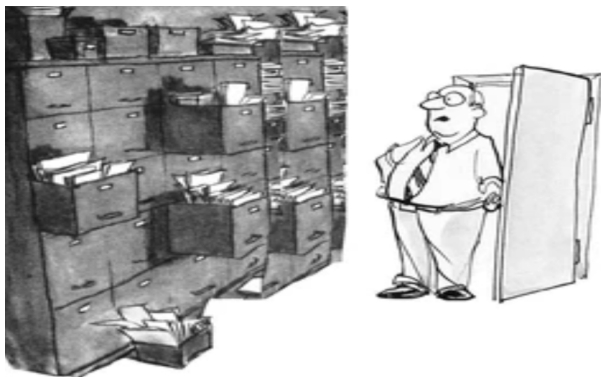


Un cambio de enfoque

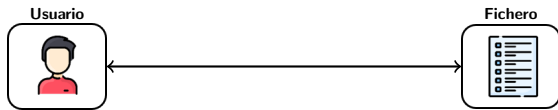
En los sistemas orientados a ficheros los humanos tienen el control sobre los ficheros

Un cambio de enfoque

En los sistemas orientados a ficheros los humanos tienen el control sobre los ficheros



Un cambio de enfoque



El usuario interactúa directamente con los ficheros



El usuario interactúa con la base de datos mediante *software*

¿Qué es este *software*?

Sistemas de gestión de bases de datos (SGBD)

*“Un **sistema de gestión de bases de datos** es un sistema computacional que proporciona **funcionalidades, medios o servicios para manipular** y, en particular, **manejar todos los accesos** a una base de datos o una colección de bases de datos.”*

—C. J. Date

- ▶ **Persistentes:** los datos permanecen en memoria externa

Superando limitaciones

- ▶ **Persistentes:** los datos permanecen en memoria externa
- ▶ **Masivos:** manejan terabytes/petabytes de datos

Superando limitaciones

- ▶ **Persistentes:** los datos permanecen en memoria externa
- ▶ **Masivos:** manejan terabytes/petabytes de datos
- ▶ **Eficientes:** operaciones eficientes gracias al uso de estructuras de datos y algoritmos

Superando limitaciones

- ▶ **Persistentes:** los datos permanecen en memoria externa
- ▶ **Masivos:** manejan terabytes/petabytes de datos
- ▶ **Eficientes:** operaciones eficientes gracias al uso de estructuras de datos y algoritmos
- ▶ **Multi-usuarios:** protocolos para la gestión de accesos concurrentes

Superando limitaciones

- ▶ **Persistentes:** los datos permanecen en memoria externa
- ▶ **Masivos:** manejan terabytes/petabytes de datos
- ▶ **Eficientes:** operaciones eficientes gracias al uso de estructuras de datos y algoritmos
- ▶ **Multi-usuarios:** protocolos para la gestión de accesos concurrentes
- ▶ **Seguros:** consistentes ante accesos por usuarios no autorizados y fallos del sistema

Superando limitaciones

- ▶ **Persistentes:** los datos permanecen en memoria externa
- ▶ **Masivos:** manejan terabytes/petabytes de datos
- ▶ **Eficientes:** operaciones eficientes gracias al uso de estructuras de datos y algoritmos
- ▶ **Multi-usuarios:** protocolos para la gestión de accesos concurrentes
- ▶ **Seguros:** consistentes ante accesos por usuarios no autorizados y fallos del sistema
- ▶ **Disponibles:** al 99.99999 %

Superando expectativas

Conveniencia de los SGBD

Conveniencia de los SGBD

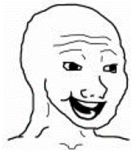
- ▶ **Independencia física de datos:** admiten el cambio de la forma de almacenamiento de los datos, pero la estructura de la base de datos y las operaciones definidas sobre ella no cambian.

Conveniencia de los SGBD

- ▶ **Independencia física de datos:** admiten el cambio de la forma de almacenamiento de los datos, pero la estructura de la base de datos y las operaciones definidas sobre ella no cambian.
- ▶ **Independencia lógica de datos:** proporcionan lenguajes de consulta declarativos, se define lo que se desea pero no cómo alcanzarlo.

Superando las expectativas

Usuario de SGBD



Lenguaje declarativo

Desarrollador de SGBD



Estructuras de datos

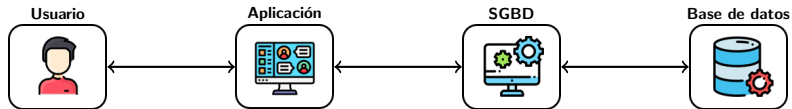
Algoritmos

Optimización

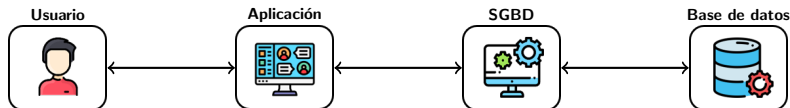
Compilación

Gestión de ficheros

Sistemas de bases de datos (SBD)



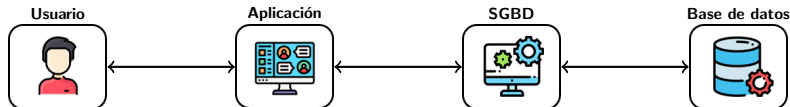
Sistemas de bases de datos (SBD)



*“Un **sistema de base de datos** es un **sistema computacional de mantenimiento de registros**, que se diseña para manejar grandes cantidades de información.”*

—C. J. Date

Sistemas de bases de datos (SBD)



*“Un **sistema de base de datos** es un **sistema computacional de mantenimiento de registros**, que se diseña para manejar grandes cantidades de información.”*
—C. J. Date

Funciones

- ▶ Insertar datos
- ▶ Editar datos
- ▶ Eliminar datos
- ▶ Consultar datos

Evolución de las tecnologías de la información

1980

1990

2000

2010



Primeros sistemas
gestores de bases de
datos relacionales
comerciales



Estándar de
lenguaje SQL



Primeros servidores
de bases de datos
relacionales



Procesamiento
analítico de datos

Evolución de las tecnologías de la información

1980

1990

2000

2010



Internet



Negocios online



Modelo orientado a
objetos
XML

Evolución de las tecnologías de la información

1980

1990

2000

2010



Data Warehousing



Redes sociales



Computación móvil



Inicios del Big Data

Evolución de las tecnologías de la información

1980

1990

2000

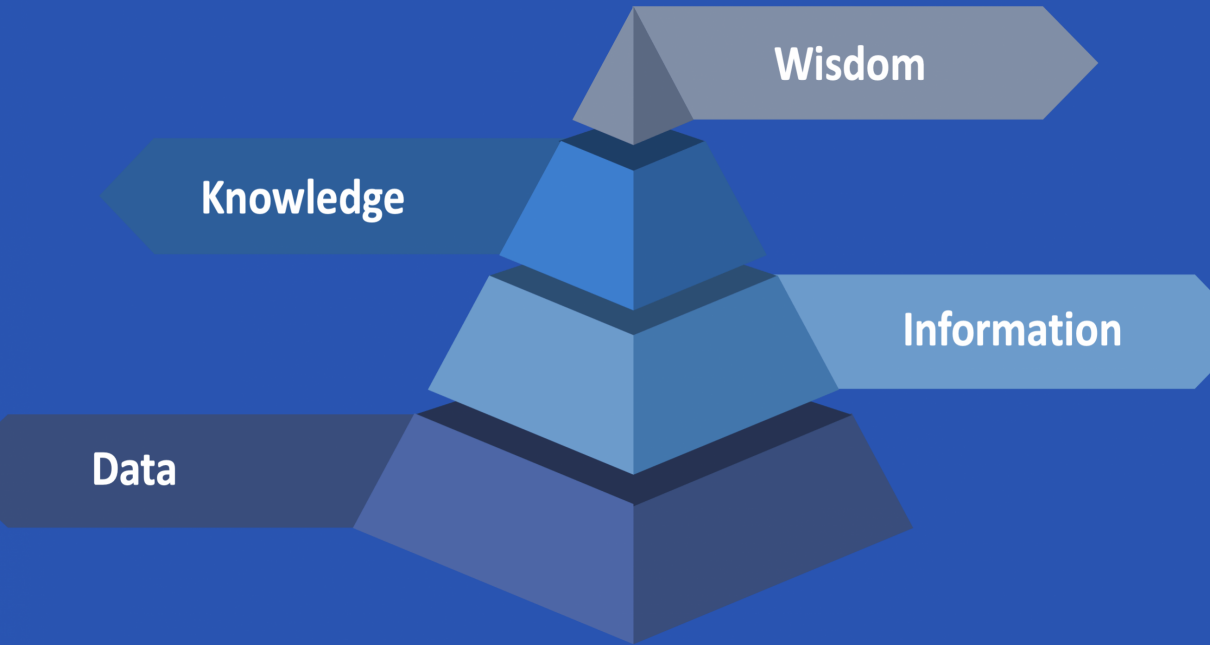
2010



Sistemas gestores
de bases de datos
NoSQL



Computación en la
nube



Objetivos de la asignatura

Proporcionar un conjunto de métodos y herramientas para:

Objetivos de la asignatura

Proporcionar un conjunto de métodos y herramientas para:

- ▶ Diseñar e implementar bases de datos correctas

Objetivos de la asignatura

Proporcionar un conjunto de métodos y herramientas para:

- ▶ Diseñar e implementar bases de datos correctas
- ▶ Evaluar la calidad de bases de datos específicas

Objetivos de la asignatura

Proporcionar un conjunto de métodos y herramientas para:

- ▶ Diseñar e implementar bases de datos correctas
- ▶ Evaluar la calidad de bases de datos específicas
- ▶ Identificar la vigencia del modelo relacional y sus limitaciones

Objetivos de la asignatura

Proporcionar un conjunto de métodos y herramientas para:

- ▶ Diseñar e implementar bases de datos correctas
- ▶ Evaluar la calidad de bases de datos específicas
- ▶ Identificar la vigencia del modelo relacional y sus limitaciones
- ▶ Reconocer casos de uso para bases de datos no relacionales

Estructura de la asignatura

Análisis de Requerimientos

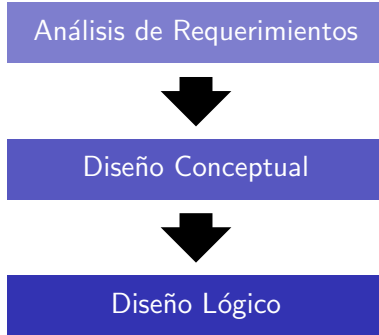
Estructura de la asignatura

Análisis de Requerimientos

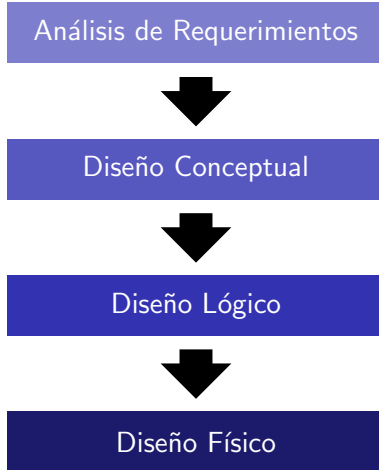


Diseño Conceptual

Estructura de la asignatura



Estructura de la asignatura



Evaluación de la asignatura

Evaluación Sistemática

- ▶ Participación en clases
- ▶ Entrega de tareas orientadas
- ▶ Asistencia



Evaluación de la asignatura

Evaluación Sistemática

- ▶ Participación en clases
- ▶ Entrega de tareas orientadas
- ▶ Asistencia

Nota Final

- ▶ Proyecto integrador
- ▶ Opinión del profesor



Próximo encuentro

Clase Práctica

- ▶ Viernes, 2:55pm, aula 2
- ▶ Instalar *software*
- ▶ Compartir bibliografía
- ▶ Orientación del proyecto



@matcom_database_ds

Próximo encuentro

Clase Práctica

- ▶ Viernes, 2:55pm, aula 2
- ▶ Instalar *software*
- ▶ Compartir bibliografía
- ▶ Orientación del proyecto



@matcom_database_ds

Bases de Datos

Introducción a las Bases de Datos Relacionales y No Relacionales

Lic. Andy Ledesma García
Lic. Víctor M. Cardentey Fundora
Dra. Lucina García Hernández

Departamento de Computación
Facultad de Matemática y Computación
Universidad de La Habana

Licenciatura en Ciencia de Datos

23 de enero de 2024