

Módulo 1. Las bases de datos

Introducción a las bases de datos

- [Qué es una base de datos](#)
- [Bases de datos relacionales](#)
- [Sistemas Gestores de Bases de Datos](#)
- [Diseño de una base de datos](#)

Qué es una base de datos

Una base de datos es un sistema informático orientado a los datos que pretende recuperar y almacenar la información de manera eficiente y cómoda. Surge en un intento de resolver las dificultades del procesamiento tradicional de los datos, teniendo en cuenta que los datos suelen ser independientes de las aplicaciones.

Las bases de datos nacen como solución al problema que se plantea a los programadores a la hora de diseñar programas eficientes y capaces de manejar grandes cantidades de información.

Estas bases de datos responden a un modelo. Modelar consiste en definir un mundo abstracto y teórico tal que las conclusiones que se puedan sacar de él coinciden con el comportamiento del mundo real.

Un **modelo de datos** es un conjunto de **conceptos**, **reglas** y **conversiones** que permiten describir los datos del problema que se plantea solucionar y su posterior manipulación.

Los modelos de datos que han contado en la historia con un mayor grado de aceptación son el **jerárquico**, **de red** y **relacional**. Aunque la complejidad de las aplicaciones informáticas ha hecho que las **BD NoSQL** estén ganando cada vez más importancia.

Financiado por el Ministerio de Educación y Formación Profesional y por la Unión Europea - NextGenerationEU



Bases de datos relacionales

El modelo relacional nació como consecuencia de los trabajos publicados por Codd en 1970, y su duración en el tiempo es debido en gran medida a que es un modelo bien fundado en bases matemáticas que puede representarse fácilmente usando algoritmos computacionales.

Los **objetivos** fundamentales que buscaba el trabajo de Codd se pueden resumir en:

- **Independencia física:** el modo de almacenamiento de los datos no influye en su manipulación lógica.
- **Independencia lógica:** el añadir, eliminar o modificar objetos de la BD no repercute en el acceso a subconjuntos de los mismos.
- **Flexibilidad:** se puede presentar a cada usuario los datos de una forma distinta.
- **Uniformidad:** las estructuras lógicas de los datos presentan un aspecto uniforme.
- **Sencillez:** es fácil de utilizar y comprender para el usuario final.

En 1985 Codd publicó sus 12 reglas analizando algunos productos comerciales. Éstas son:

1. **Regla de información:** la información está representada explícitamente a nivel lógico y exactamente de un modo, mediante valores en tablas.
2. **Regla de acceso garantizado:** cada uno de los datos se garantiza que sea lógicamente accesible mediante una combinación de nombre de tabla, valor de clave primaria y nombre de columna.
3. **Tratamiento sistemático de valores nulos:** los valores nulos se soportan en los SGBD completamente relacionales para representar la falta de información.
4. **Catálogo en línea** dinámico basado en el modelo relacional: la descripción de la base de datos se representa a nivel lógico del mismo modo que los datos ordinarios.
5. **Regla de sublenguaje completo de datos:** Existen varios tipos de lenguaje. El modelo relacional debe soportar por lo menos un lenguaje relacional que:
 - Tenga una sintaxis lineal.
 - Puede ser utilizado dentro de programas de uso.
 - Soporte operaciones de definición de datos, de manipulación de datos, seguridad e integridad y operaciones de administración de transacciones.
6. **Regla de actualización de vista:** toda vista teóricamente actualizable es también actualizable por el sistema.
7. **Inserción, actualización y supresión de alto nivel:** los datos se pueden recuperar de una base de datos relacional en los sistemas construidos de datos de filas múltiples y/o de tablas múltiples
8. **Independencia física de los datos.**
9. **Independencia lógica de los datos.**
10. **Independencia de integridad:** las limitaciones de la integridad se deben especificar por separado de los programas de la aplicación y se almacenan en la base de datos.

11. **Independencia de distribución:** la distribución de las porciones de la base de datos a las varias localizaciones debe ser invisible a los usuarios de la base de datos.
12. **Regla de no subversión:** si el sistema proporciona una interfaz de bajo nivel de registro, a parte de una interfaz relacional, esa interfaz de bajo nivel no se puede utilizar para subvertir el sistema

Financiado por el Ministerio de Educación y Formación Profesional y por la Unión Europea - NextGenerationEU



Sistemas Gestores de Bases de Datos

Las bases de datos requieren básicamente un software de gestión que facilite las operaciones y las interfaces con los usuarios. Esto es el Sistema Gestor de Bases de Datos, SGBD, o en inglés DBMS, Data Base Management System.

Las **operaciones típicas** que debe realizar un SGBD son las siguientes:

- Aquéllas que afectan a la totalidad de los datos.
 - Creación
 - Reestructuración
 - Consultas
- Las que tienen lugar sobre registros concretos, que suelen llamarse operaciones de actualización:
 - Inserciones
 - Borrados
 - Modificaciones
 - Consultas de selección.

Componentes de los SGBD

Para realizar todas las operaciones nombradas anteriormente es necesario que el SGBD cuente con una serie de componentes:

- **Lenguajes de la base de datos:** lenguaje de definición de datos (LDD), lenguaje de manipulación de datos (LMD) y lenguaje de control de datos (LCD).
- **El diccionario de datos:** es un conjunto de archivos que contienen información acerca de los datos que se almacenan en la base de datos. Se trata de una "metabase de datos", es decir, una base de datos que contienen información sobre la base de datos (datos acerca de los datos)
- **El gestor de la base de datos:** Este componente es el encargado de proporcionar una interfaz entre los datos almacenados y los programas de aplicación que los manejan. Puede verse el gestor de la base de datos como un intérprete entre el usuario y los datos. Toda operación que se quiere realizar "contra" la base de datos debe ser previamente autorizada por el gestor de la misma, el cual, una vez interpretada y validada, o bien realiza la operación devolviendo el resultado de la misma al programa que lo solicitó o bien lo rechaza.

- **El administrador de la base de datos:** Es una persona o grupo de personas encargadas de la función de administración de la base de datos
- **Usuarios de la base de datos.**

Sistemas Gestores de Bases de Datos en el mercado

Actualmente las bases de datos relacionales más conocidas son **MySQL** y **Oracle**, ambas pertenecientes a ORACLE y que requieren instalar el sistema gestor de bases de datos en un sistema operativo.



MySQL es **gratis** y de **código abierto**, mientras que Oracle es de **pago** y **privado**. Tanto MySQL como Oracle ofrecen soporte tanto de comunidad como técnico.

Aunque ambos son sistemas basados en modelos relacionales, se pueden encontrar bastantes diferencias entre ellos, siendo ambas opciones muy válidas y potentes para trabajar con bases de datos relacionales.

Para evitar tener que andar con instalaciones en este curso, nos vamos a centrar en herramientas que permiten realizar el diseño de las bases de datos online, sin necesidad de instalar ningún gestor en nuestros sistemas operativos,

Para realizar el diseño de los modelos Entidad-Relación, así como el relacional, se pueden utilizar distintas herramientas que permitan crear diagramas, hay muchas en el mercado. Entre ellas podemos encontrar tanto de escritorio como online, como **ERD Plus** o **Creatly**, que son online o **Dia**, que es una aplicación de escritorio pero también se puede utilizar online a través de la web <https://www.rollapp.com/app/dia> y es la que yo recomiendo utilizar por su semejanza con la notación explicada a lo largo de este curso. Todas las herramientas tienen alguna limitación a la hora de crear los diagramas E/R, como no permitir crear relaciones que no sean binarias, o poner cardinalidades en las relaciones, por ejemplo. Así que elijáis la que elijáis, en algún momento tendréis que idear la manera de representar lo que no permite esa herramienta en concreto.

Durante este curso se va a utilizar la notación de base de datos Chen, y la aplicación Dia la permite.

Hay dos opciones, o bien instalar la aplicación Dia en vuestro ordenador, descargándola desde su web [aquí](#) (fijarse que hay versión para Windows, Mac y Linux, hay que elegir la correcta) o bien registrarse en rollApp (<https://www.rollapp.com>) y luego buscar la aplicación Dia y lanzarla online.

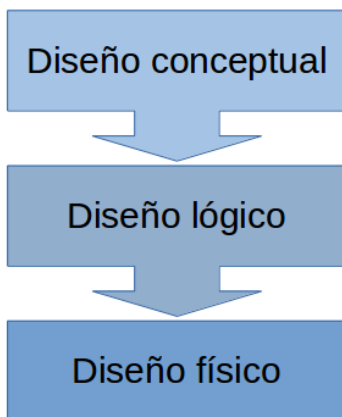
Financiado por el Ministerio de Educación y Formación Profesional y por la Unión Europea - NextGenerationEU



Diseño de una base de datos

El diseño de bases de datos requiere de una metodología lo suficientemente potente para realizar el proceso de modelado de una forma metódica y eficiente, ya que a la hora de realizar este trabajo pueden aparecer problemas difíciles de solucionar más adelante.

La manera óptima de diseñar una base de datos es realizando el proceso en fases. Las fases a seguir para modelar un sistema de bases de datos relacional son:



- **Diseño conceptual:** se obtiene a través del estudio del problema y seleccionando qué elementos se van a modelar. En el caso del modelo relacional, se utiliza el Modelo Entidad-Relación.
- **Diseño lógico:** El objetivo del diseño lógico es convertir el esquema conceptual en un esquema lógico que se ajuste al modelo de SGBD sobre el que se vaya a implementar el sistema. El proceso a seguir para realizar el diseño lógico consiste en tomar el diagrama Entidad-Relación obtenido en el diseño conceptual y obtener las relaciones o tablas propias del modelo relacional, siguiendo unas determinadas reglas de transformación. El diagrama obtenido en este paso puede presentar algunos problemas derivados de fallos a la hora de interpretar el problema real, fallos arrastrados del diseño del diagrama Entidad-Relación o fallos del paso al modelo relacional. Entre otros problemas destacan:
 - Incapacidad para almacenar ciertos hechos.
 - Redundancias, y por tanto, posibilidad de incoherencias.
 - Pérdida de información.
 - Pérdida de dependencias funcionales, es decir, de ciertas restricciones de integridad que dan lugar a interdependencias entre los datos.
 - Aparición, en la base de datos, de estados que no son válidos en el mundo real, es decir, anomalías de inserción, borrado y modificación.

Es por esto que el esquema relacional debe ser analizado para comprobar que no presenta estos problemas. Y esto se hace por medio de la **normalización** de las tablas que, aunque es algo que no abordamos en este curso, es importante nombrarlo para saber que es un paso importante en el diseño de bases de datos relacionales.

- **Diseño físico:** se realiza la implementación propiamente dicha de la base de datos y está estrechamente relacionada al SGBD. Los pasos a seguir para la realización de un buen diseño físico son:
 - Traducir el esquema lógico para el SGBD específico.
 - Diseñar la representación física.
 - Diseñar los mecanismos de seguridad.
 - Monitorizar y afinar el sistema.

Financiado por el Ministerio de Educación y Formación Profesional y por la Unión Europea - NextGenerationEU

