

Niveles de abstracción y modelado de datos

Sitio: Agencia de Habilidades para el Futuro
Curso: Administración de Base de Datos 1° D
Libro: Niveles de abstracción y modelado de datos

Imprimido por: RODRIGO PINTO
Día: miércoles, 27 de noviembre de 2024, 07:39

Tabla de contenidos

1. Niveles de abstracción

1.1. Nivel conceptual

2. Uso de los modelos

3. Modelos de datos

3.1. Modelos de datos conceptuales

3.2. Modelos de datos físicos

3.3. Modelos de datos lógicos

4. Ventajas y desventajas de los modelos de datos

Niveles de abstracción

Repasemos el concepto de niveles de abstracción de una base de datos.

Los usuarios no tienen por qué conocer cómo están organizados y almacenados los datos. Por este motivo una base de datos debe presentar los datos de forma que el usuario pueda interpretarlos y modificarlos. Evidentemente esto no lo podemos aplicar a un informático que necesite saber donde se encuentran físicamente los datos para poder tratarlos.

Podemos destacar **tres niveles principales** según la visión y la función que realice el usuario sobre la base de datos.

El **nivel externo** es el nivel de vista del usuario. Cada vista de usuario se conoce como esquema externo, con el que descubrimos alguna parte de la base de datos.

El **nivel conceptual** describe la estructura lógica de toda la base de datos. El esquema conceptual se concentra en describir entidades, tipos de datos, relaciones, operaciones y restricciones.

El **nivel interno** representa el esquema interno o físico. Describe cómo se almacena en disco los datos y los caminos de acceso a la base de datos.

Nivel lógico o conceptual

El **nivel lógico o conceptual** describe qué datos son almacenados realmente en la base de datos y las relaciones que existen entre los mismos. En otras palabras, describe la base de datos completa en términos de su estructura de diseño. Este nivel de abstracción lo usan los/as administradores de bases de datos, quienes deben decidir qué información se va a guardar en la base de datos.

En este nivel de abstracción, los/as administradores de bases de datos deciden que información guardar. Esta decisión consta de diferentes tareas.

1. **Definición de los datos.** Definición de los datos. Se describen el tipo de datos. Longitud de campo todos los elementos direccionables en la base.
2. **Relaciones entre datos.** Se definen las relaciones entre datos para enlazar tipos de registros relacionados para el procesamiento de archivos múltiples.



En el nivel conceptual la base de datos aparece como una colección de registros lógicos, sin descriptores de almacenamiento. En realidad los archivos conceptuales no existen físicamente. La transformación de registros conceptuales a registros físicos para el almacenamiento se lleva a cabo por el sistema y es transparente al usuario.

Uso de los modelos

Los modelos consisten en **sistemas de diagramas o imágenes** que permiten que más personas puedan participar en el diseño del sistema. Si mostramos a un usuario final un centenar de líneas de código, esto será muy difícil para su comprensión; pero si mostramos a ese usuario final un diagrama de actividades, esa misma persona podrá intervenir en el proceso de trabajo.

El **modelado de datos** es el proceso de creación de una representación visual o esquema que define los sistemas de recopilación y administración de información de cualquier organización. Este esquema o modelo de datos ayuda a crear una vista unificada de los datos de una organización. El modelo esboza los datos que recoge la empresa, la **relación** entre los distintos conjuntos de datos y los métodos que se usarán para almacenarlos y analizarlos.

El modelado de datos aporta las siguientes ventajas:

- **Reduce los errores** en el desarrollo de software de bases de datos.
- Facilita la **rapidez y eficacia** en el diseño y creación de bases de datos.
- Facilita la **comunicación** entre los ingenieros de datos y los departamentos de inteligencia empresarial.

Modelos de datos

Los modelos de datos definen cómo se modela la estructura lógica de una base de datos y son entidades fundamentales para introducir la abstracción en una base de datos.

Estos modelos definen cómo los datos se conectan entre sí y cómo se procesan y almacenan dentro del sistema. Su enfoque principal es apoyar y ayudar a los sistemas de información mostrando el formato y la definición de los diferentes datos involucrados, al tiempo que ayudan a evitar la redundancia de datos.

La información almacenada en los modelos de datos es de gran importancia para las empresas porque dicta las relaciones entre las tablas de la base de datos, las claves externas y los eventos involucrados.

A continuación, revisaremos ejemplos de cada uno de los tres tipos básicos de modelo de datos:

- **Modelo conceptual**
- **Modelo lógico**
- **Modelo físico**

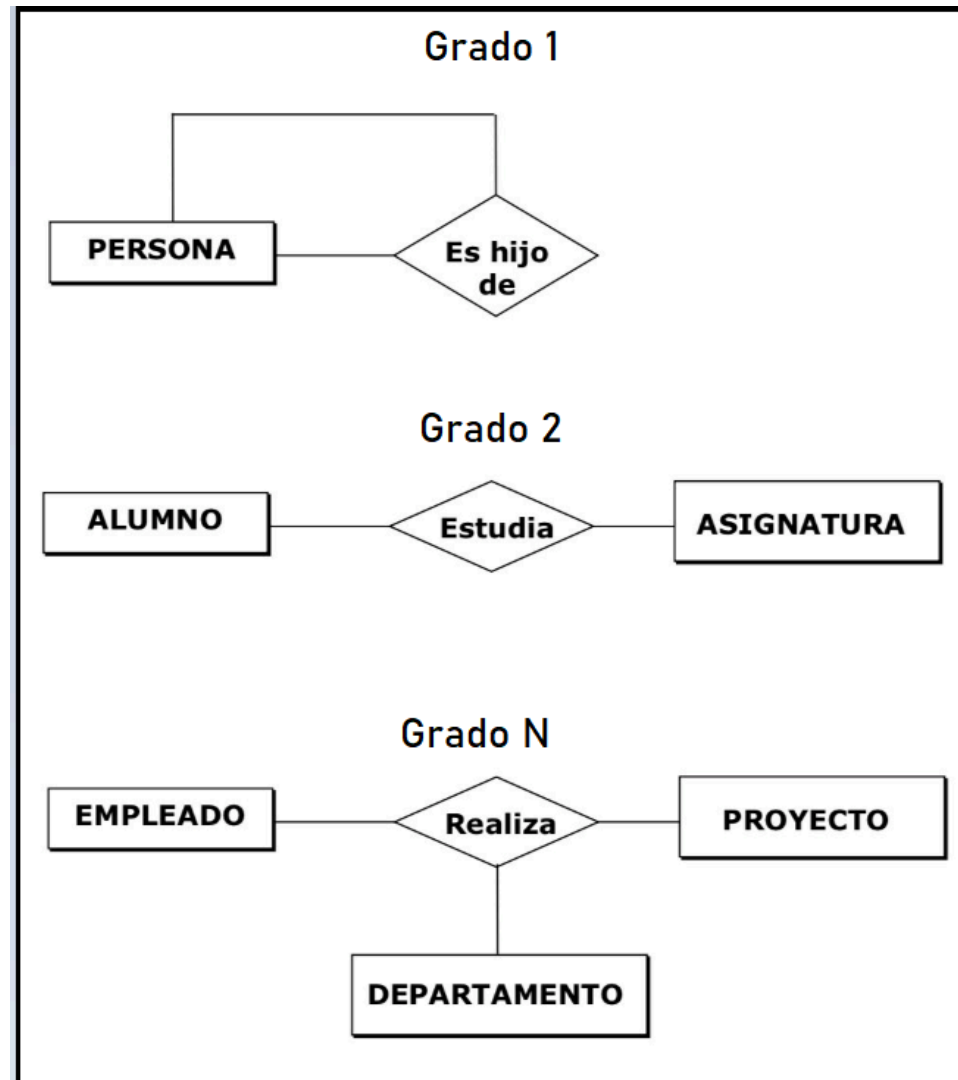
Modelos de datos conceptuales

Un modelo conceptual de datos identifica las relaciones de más alto nivel entre las diferentes entidades. Las características del modelo conceptual de datos incluyen:

- Incluye las entidades importantes y las relaciones entre ellas.
- No se especifica ningún atributo.
- No se especifica ninguna clave principal.

El requisito para un modelamiento exitoso pasa necesariamente por el “conocimiento del negocio”. Para lograr la meta de representar y organizar los datos y obtener la información que requiere el problema a resolver, se necesita un conocimiento cabal del problema. Modelar significa -en un modo amplio- simplificar la realidad del negocio pero sin perder significancia de sus datos. Se trata de **organizar y clasificar la información en componentes simples que representen la información del negocio**.

La siguiente figura es un ejemplo de un **modelo conceptual de datos**.



Modelos de datos Físicos

Un modelo de base de datos física muestra todas las estructuras de tabla, incluidos el nombre de columna, el tipo de datos de columna, las restricciones de columna, la clave principal, la clave externa y las relaciones entre las tablas.

Las características de un modelo de datos físicos incluyen:

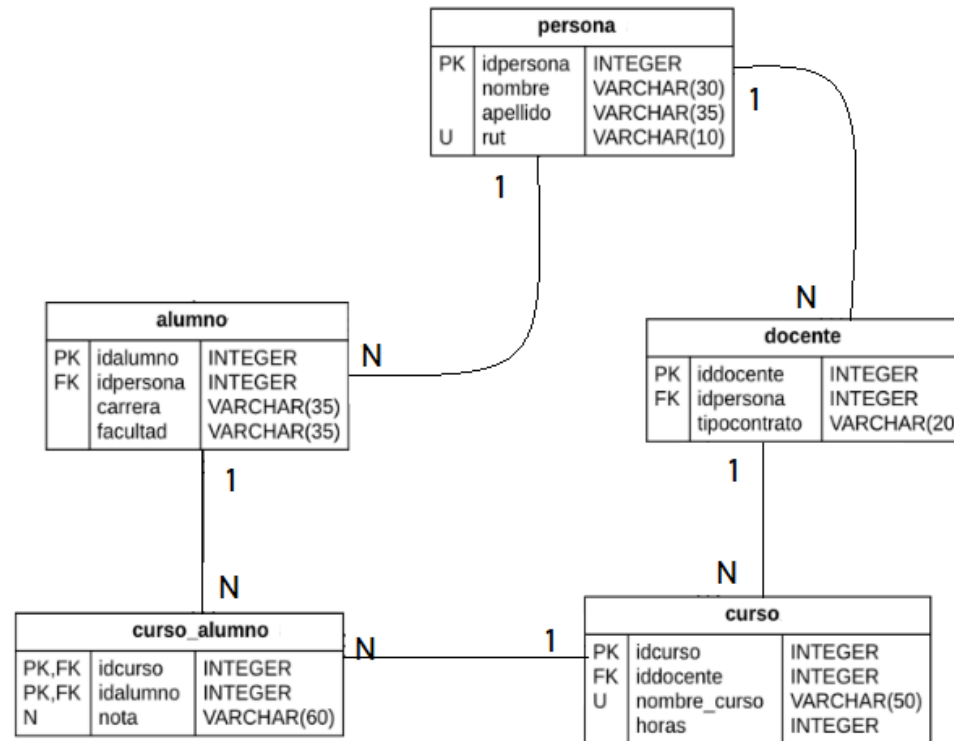
- Especificación de todas las tablas y columnas.
- Las claves externas se usan para identificar relaciones entre tablas.
- La desnormalización puede ocurrir según los requisitos del usuario.

Las consideraciones físicas pueden hacer que el modelo de datos físicos sea bastante diferente del modelo de datos lógicos. Además, el modelo de datos físicos será diferente para distintos Sistemas de Gestión de Base de datos. Por ejemplo, el tipo de datos para una columna puede ser diferente entre MySQL y SQL Server.

Los pasos básicos para el diseño del modelo de datos físicos son los siguientes:

- Convertir **entidades** en tablas.
- Convertir **relaciones** en claves externas.
- Convertir **atributos** en columnas.
- Modificar el modelo de datos físicos en función de las **restricciones/requisitos físicos**.

En el siguiente ejemplo, una persona puede ser alumno y/o docente, un alumno puede cursar varias carreras y un docente puede enseñar en varios cursos.



Modelos de datos lógicos

Un modelo de datos lógicos describe los datos con el mayor detalle posible, independientemente de cómo se implementarán físicamente en la base de datos.

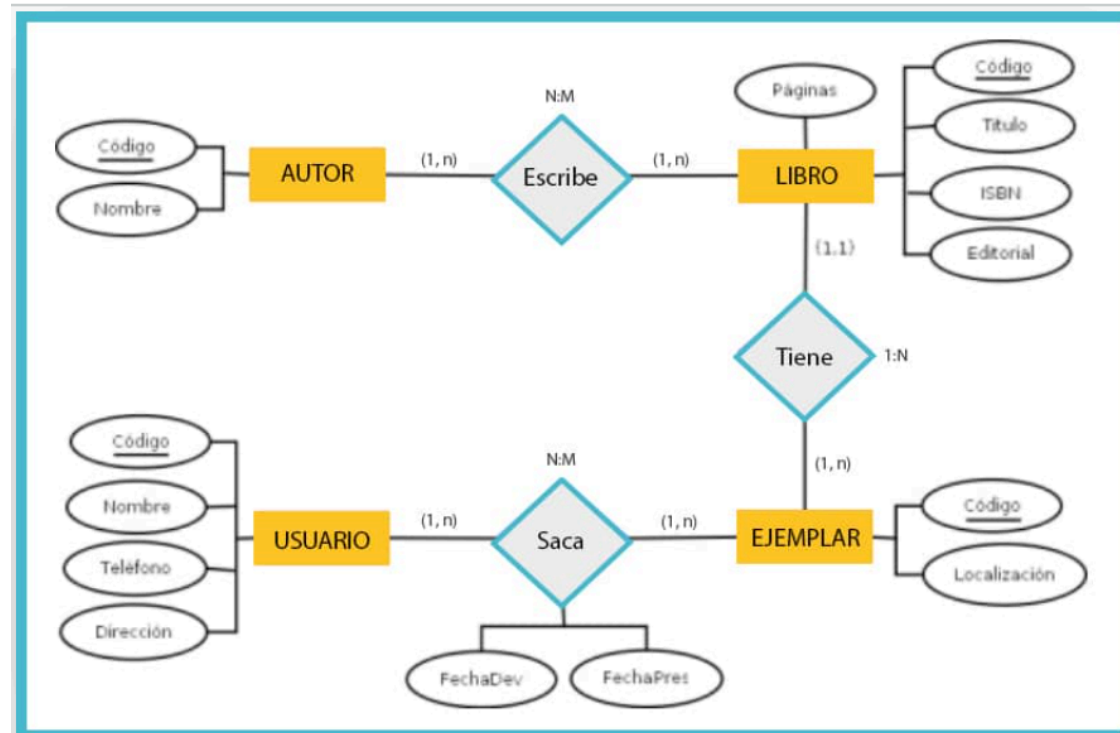
Las características de un modelo de datos lógicos incluyen:

- Se representan las entidades y sus relaciones.
- Se especifican los atributos para cada entidad.
- Se sitúa la clave principal para cada entidad y las claves externas (claves que identifican la **relación** entre diferentes entidades)
- La **normalización** ocurre en este nivel.

A partir de estas características, podemos definir y ordenar los pasos para diseñar un modelo de datos lógicos.

- Especificar **claves primarias** para todas las entidades.
- Encontrar las **relaciones** entre diferentes entidades.
- Describir todos los **atributos** para cada entidad.
- Resolver las **relaciones** de muchos a muchos.
- **Normalización.**

La siguiente figura es un ejemplo de un modelo de datos lógicos.



Ventajas y desventajas de los modelos de datos

Ventajas

- Permite asegurar que los objetos de datos se representen con precisión.
- Es lo suficientemente detallado para ser utilizado para construir la base de datos física.
- La información en el modelo de datos se puede utilizar para definir la **relación** entre tablas, claves primarias y externas y procedimientos almacenados.
- Ayuda a las empresas a comunicarse dentro y entre las organizaciones.
- Permite documentar las asignaciones de datos en el proceso ETL Ayuda a reconocer las fuentes de datos correctas para poblar el modelo.

Desventajas

- Para desarrollar el modelo de datos se deben conocer las características físicas de los datos almacenados.
- Incluso los cambios más pequeños realizados en la estructura requieren modificaciones en toda la aplicación.
- No hay un lenguaje de manipulación de modelos establecido en DBMS.