

# Adquisición, gestión y gobernanza de datos

Modelos de Datos

**Clase 3**

MAESTRÍA EN  
SISTEMAS DE INFORMACIÓN  
Mención Data Science

La excelencia no se improvisa



## INTRODUCCIÓN

Los modelos de datos son representaciones abstractas que describen la estructura, organización y relaciones de los datos dentro de un sistema. Sirven como un marco conceptual que guía el diseño de bases de datos y facilita la gestión y manipulación de la información. Existen diferentes tipos de modelos de datos, como el modelo jerárquico, el modelo de red y el modelo relacional, cada uno con sus propias características y usos específicos. A medida que la tecnología y las necesidades de negocio han evolucionado, también lo han hecho los modelos de datos, dando lugar a enfoques más avanzados, como los modelos orientados a objetos y los modelos de datos NoSQL, que se adaptan mejor a las demandas de aplicaciones modernas que requieren flexibilidad y escalabilidad (Date, 2004).

El uso de modelos de datos es esencial para garantizar la integridad y coherencia de la información en cualquier sistema de gestión de bases de datos. Un modelo bien diseñado permite a los desarrolladores y analistas comprender las interrelaciones entre diferentes tipos de datos, facilitando la implementación de consultas y transacciones eficaces. Además, los modelos de datos ayudan a minimizar la redundancia y a optimizar el almacenamiento, lo que resulta en un sistema más eficiente y fácil de mantener. A medida que las organizaciones continúan generando y almacenando grandes volúmenes de datos, la importancia de contar con modelos de datos robustos y bien definidos se vuelve cada vez más crítica para la toma de decisiones basada en datos (Elmasri & Navathe, 2015).

### RDA1

Utilizar cada una de las etapas del proceso de descubrimiento de conocimiento en el contexto de Big Data.

## Clase 3: Modelos de Datos

### Características del Modelo de Datos

Un modelo de datos es una representación abstracta de la estructura de los datos, que permite organizar y gestionar la información de manera efectiva. Las características fundamentales de un modelo de datos incluyen:

**Tabla 1**

*Características de un modelo de datos*

Características	Descripción
<b>Abstracción</b>	Representan datos complejos de manera simplificada, ocultando detalles innecesarios y enfocándose en entidades relevantes y sus relaciones.
<b>Consistencia</b>	Garantizan que la información se almacene y procese de manera coherente, evitando duplicaciones y errores.
<b>Flexibilidad</b>	Deben ser adaptables a cambios futuros, permitiendo la incorporación de nuevos requerimientos sin afectar la integridad de los datos existentes.
<b>Interoperabilidad</b>	Deben ser compatibles con diferentes sistemas y plataformas, facilitando el intercambio de información.
<b>Simplicidad</b>	Deben ser fáciles de entender y utilizar, permitiendo que desarrolladores y usuarios finales interactúen de manera eficiente.
<b>Escalabilidad</b>	Capacidad de manejar incrementos en el volumen de datos sin comprometer el rendimiento, especialmente en organizaciones en crecimiento.

Fuente: Elaboración propia

Estas características son fundamentales para diseñar modelos de datos que no solo sean efectivos en la representación de la información, sino que también se integren bien en el entorno tecnológico de una organización.

### Tipos de modelos de datos

Los modelos de datos son herramientas fundamentales en la estructura y gestión de la información en sistemas de bases de datos. Actúan como un marco conceptual que define cómo se organizan, almacenan y acceden a los datos, permitiendo que tanto los desarrolladores como los analistas comprendan las relaciones y la jerarquía de la información. Existen diversos tipos de modelos de datos, cada uno diseñado para satisfacer diferentes necesidades y requisitos. Entre los más comunes se encuentran el modelo relacional, el modelo jerárquico, el modelo de red y, más recientemente, los modelos NoSQL y orientados a objetos. Cada uno de estos modelos ofrece un enfoque distinto para la organización de los datos, lo que influye en la forma en que se realizan las consultas y se llevan a cabo las transacciones (Elmasri & Navathe, 2015).

El modelo relacional, por ejemplo, se basa en tablas y relaciones entre ellas, lo que facilita la manipulación y recuperación de datos mediante el uso de lenguajes de consulta como SQL. Por otro lado, el modelo jerárquico organiza los datos en una estructura de árbol, donde cada registro tiene una única raíz y puede tener múltiples hijos, lo que limita su flexibilidad, pero puede ser eficiente para ciertos tipos de consultas. En contraposición, los modelos NoSQL están diseñados para manejar grandes volúmenes de datos no estructurados y semi-estructurados, permitiendo una escalabilidad horizontal y flexibilidad que los modelos tradicionales no pueden ofrecer. Así, elegir el modelo de datos adecuado es crucial para el diseño de sistemas eficientes y efectivos que satisfagan las necesidades específicas de las organizaciones y sus usuarios (Date, 2004).

**Figura 1**

*Tipos de modelos de datos.*



Fuente: <https://telefonicatech.com/blog/modelado-de-datos-lo-bien-hecho-bien-parece>


### **Modelo lógico, modelo conceptual y modelo físico**

En el diseño de bases de datos, la comprensión de los diferentes tipos de modelos es esencial para crear sistemas de información eficientes y coherentes. Los tres modelos fundamentales en este contexto son el modelo conceptual, el modelo lógico y el modelo físico. Cada uno de estos modelos desempeña un papel crucial en el proceso de desarrollo de bases de datos, desde la planificación inicial hasta la implementación final, y aborda diferentes aspectos de la organización y gestión de datos.

El **modelo conceptual** se centra en representar de manera abstracta y de alto nivel los requisitos de información de una organización. Se enfoca en los objetos y las relaciones relevantes para el negocio, sin entrar en detalles sobre cómo se implementarán físicamente estos datos. Herramientas como los diagramas Entidad-Relación (ER) son comúnmente utilizadas para este propósito, permitiendo a los diseñadores visualizar cómo se conectan diferentes entidades y atributos (Elmasri & Navathe, 2015).

A continuación, el **modelo lógico** toma el modelo conceptual y lo traduce a un diseño más detallado, que considera las estructuras de datos específicas, como tablas y relaciones, sin preocuparse por los aspectos físicos del almacenamiento. Este modelo actúa como un puente entre el diseño abstracto y la implementación real, asegurando que la estructura de datos sea coherente y normalizada (Date, 2004).

Finalmente, el **modelo físico** es la representación más detallada y concreta de cómo se almacenarán realmente los datos en el sistema de gestión de bases de datos. Este modelo toma en cuenta aspectos como el tipo de hardware, el sistema de archivos y las optimizaciones necesarias para garantizar un ac-

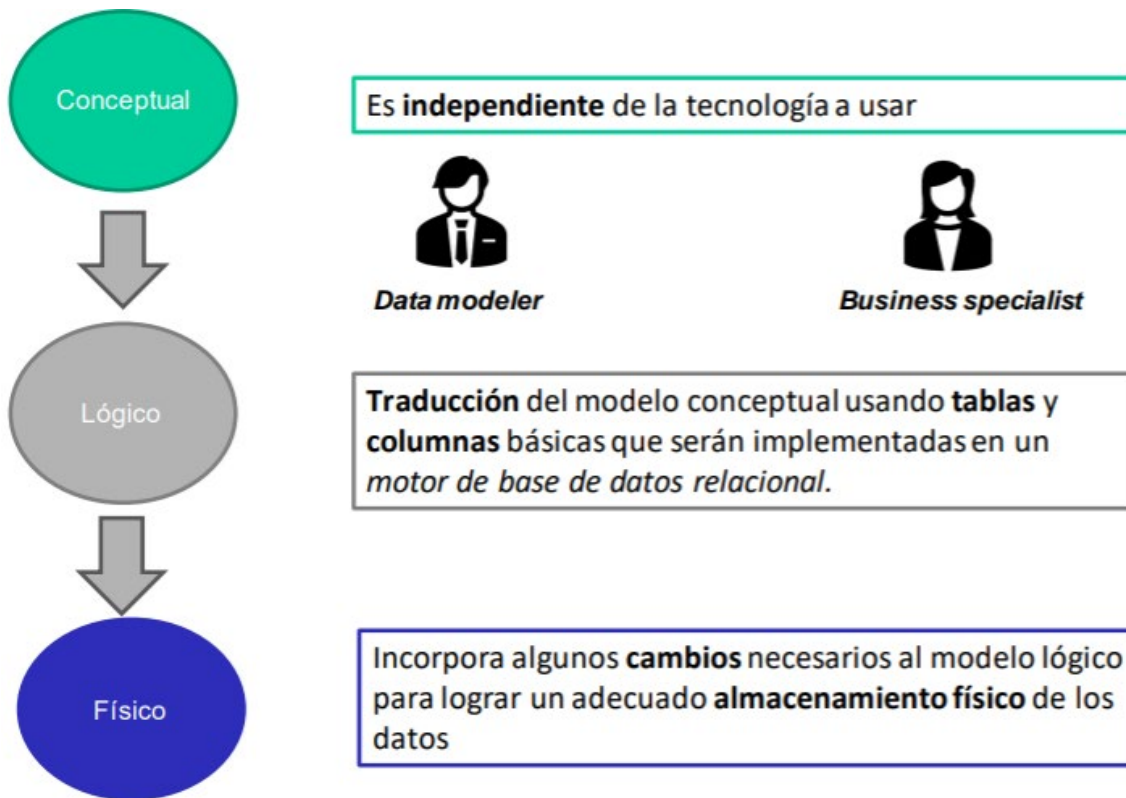


ceso eficiente a los datos. La implementación del modelo físico implica decisiones sobre la indexación, la organización del almacenamiento y otros aspectos técnicos que afectan el rendimiento del sistema (Harrington, 2016).

Así, los modelos conceptual, lógico y físico son interdependientes y esenciales para el diseño eficaz de bases de datos, permitiendo que las organizaciones gestionen su información de manera óptima.

**Figura 2**

*Niveles de abstracción de los modelos de datos*



**Fuente:** Elaboración propia

### Conceptos de Business Intelligence

Business Intelligence (BI) se refiere al conjunto de estrategias, herramientas y tecnologías utilizadas por las empresas para la recopilación, análisis y presentación de datos comerciales. Su objetivo es convertir datos brutos en información significativa que facilite la toma de decisiones informadas. A través de procesos como la minería de datos, el análisis de datos y la visualización de información, las soluciones de BI permiten a las organizaciones identificar tendencias, patrones y oportunidades de mejora en sus operaciones.

El uso de dashboards y reportes interactivos permite a los usuarios acceder a información clave de manera rápida y eficiente, promoviendo una cultura organizacional basada en datos (Turban et al., 2018). Además, Business Intelligence no solo se limita a la recopilación y análisis de datos internos, sino que también puede integrar datos externos, ofreciendo una visión más completa del entorno en el que opera la empresa (Sharda et al., 2014). En este sentido, BI se convierte en un componente crítico para las organizaciones que buscan mantenerse competitivas en un mercado en constante evolución.

### Normalización

La normalización es un proceso clave en el diseño de bases de datos que tiene como objetivo organizar los datos de manera lógica y eficiente, minimizando la redundancia y garantizando la integridad de la información. Al aplicar la normalización, los diseñadores de bases de datos dividen grandes tablas en

estructuras más pequeñas y relacionadas, facilitando así la gestión y el acceso a la información.

Este proceso se lleva a cabo a través de varias etapas conocidas como formas normales, cada una de las cuales establece criterios específicos para asegurar que las dependencias entre los datos sean adecuadamente representadas (Elmasri & Navathe, 2015). Por ejemplo, en la primera forma normal, se eliminan los grupos repetitivos de datos, mientras que en la tercera forma normal se busca eliminar las dependencias transitivas. Aunque la normalización es fundamental para mantener la coherencia y la calidad de los datos, es importante encontrar un equilibrio, ya que un diseño excesivamente normalizado puede complicar las consultas y afectar el rendimiento general del sistema (Date, 2004). En última instancia, la normalización permite a las organizaciones gestionar sus datos de forma más efectiva, asegurando que la información esté organizada, accesible y lista para el análisis.

## Modelos Dimensionales

Los modelos dimensionales son un enfoque fundamental en el diseño de bases de datos orientadas a la toma de decisiones y al análisis de datos, especialmente en el contexto de sistemas de Business Intelligence. A diferencia de los modelos relacionales tradicionales, que se centran en la normalización y la minimización de redundancias, los modelos dimensionales priorizan la facilidad de acceso y la comprensión de la información. Este enfoque se basa en la creación de estructuras que consisten en tablas de hechos y tablas de dimensiones.

Las tablas de hechos contienen datos cuantitativos, como ventas o ingresos, mientras que las tablas de dimensiones ofrecen contexto a estos datos, proporcionando información descriptiva sobre los aspectos involucrados, como el tiempo, el producto o el cliente (Kimball & Ross, 2016). Esta organización permite a los analistas realizar consultas complejas de manera más intuitiva y eficiente, facilitando la identificación de tendencias y patrones. Al estructurar los datos de esta manera, los modelos dimensionales no solo mejoran el rendimiento de las consultas, sino que también brindan a las organizaciones una herramienta poderosa para transformar datos en información significativa que apoye la toma de decisiones estratégicas.

En el siguiente enlace denominado [Tabla de hechos y dimensiones](#) se explica con un ejemplo la aplicación de las tablas de hechos y dimensiones desde el minuto 34 hasta el 36:

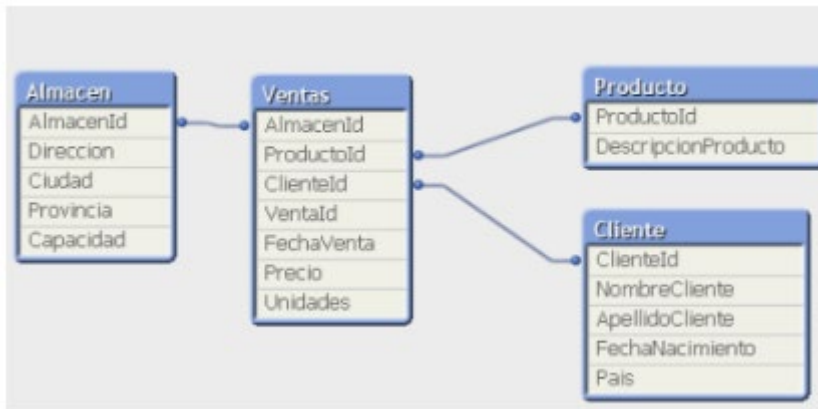
## Tipos de Modelos Dimensionales. Resumen

### Modelo Estrella

- **Estructura:**
  - Una tabla de hechos central conectada a varias tablas de dimensiones.
- **Características:**
  - Simplicidad: Fácil de entender y utilizar.
  - No hay relaciones entre las tablas de dimensiones.
- **Ventajas:**
  - Consultas rápidas.
  - Ideal para análisis de grandes volúmenes de datos.

### Figura 3

*Modelo Estrella*



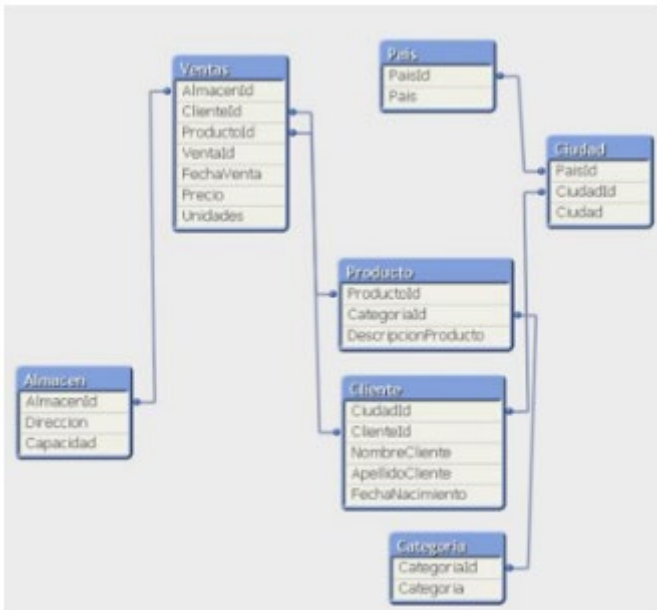
Fuente: Elaboración propia

### Modelo Copo de Nieve

- Estructura:
  - Una tabla de hechos conectada a tablas de dimensiones anidadas.
- Características:
  - Las tablas de dimensiones se descomponen en más tablas relacionadas.
- Ventajas:
  - Menor redundancia de datos.
  - Mejor organización de las dimensiones complejas.

**Figura 4**

*Modelo Copo de Nieve*



Fuente: Elaboración propia

### **Modelo de Constelación de Estrellas**

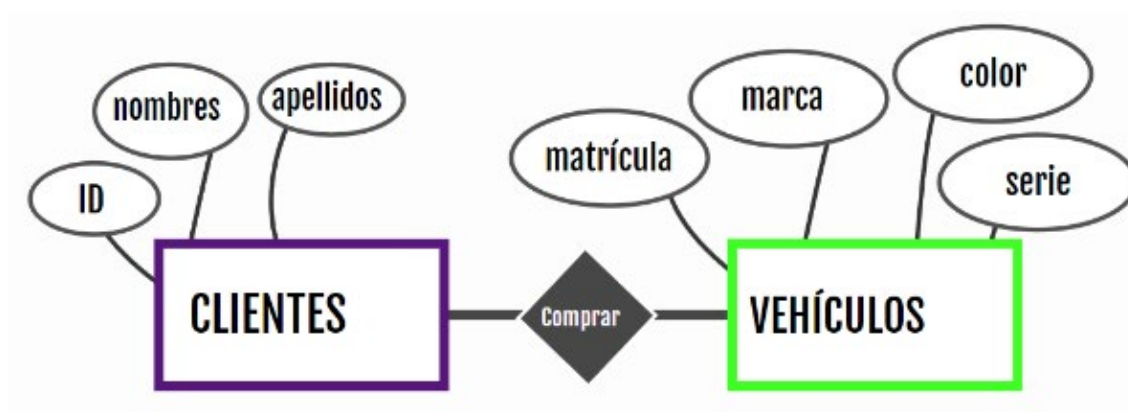
- **Estructura:**
  - Varios modelos estrella que comparten una o más tablas de dimensiones.
- **Características:**
  - Permite múltiples tablas de hechos conectadas a dimensiones comunes.
- **Ventajas:**
  - Flexibilidad para análisis multidimensionales.
  - Ideal para entornos de negocio complejos.

### **Modelo Entidad Relación**

El modelo entidad-relación (ER) es una herramienta fundamental en el diseño de bases de datos que permite representar de manera visual y conceptual la estructura de un sistema de información. Este modelo se basa en la identificación de entidades, que son objetos o conceptos relevantes dentro del dominio de estudio, y las relaciones que se establecen entre ellas. A través de diagramas ER, los diseñadores pueden ilustrar cómo las entidades interactúan y se relacionan, lo que facilita la comprensión de los requisitos del sistema (Chen, 1976).

**Figura 5**

*Modelo Entidad Relación*



Fuente: <https://informaticosinlimites.com/base-de-datos/modelo-entidad-relacion/>

Por ejemplo, en un sistema de gestión de una biblioteca, las entidades podrían incluir “Libros”, “Usuarios” y “Préstamos”, mientras que las relaciones mostrarían cómo los usuarios pueden tomar prestados libros. Este enfoque no solo ayuda a clarificar los requisitos de información antes de la implementación, sino que también actúa como un puente entre los modelos conceptuales y los diseños lógicos de bases de datos. Al adoptar el modelo ER, las organizaciones pueden asegurar una representación coherente de su información, lo que da lugar a una base de datos más organizada y fácil de mantener (Elmasri & Navathe, 2015).

### **Componentes del Modelo Entidad Relación**

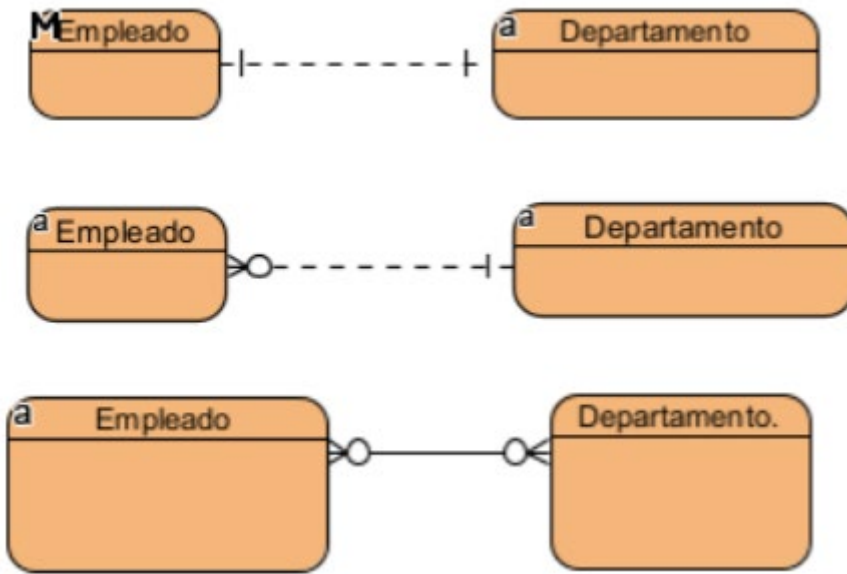
El modelo entidad-relación (ER) se compone de varios elementos clave que facilitan la representación visual y conceptual de una base de datos. En primer lugar, las **entidades** son los objetos o conceptos relevantes dentro del sistema; pueden ser tangibles, como “Clientes” o “Productos”, o intangibles, como “Transacciones” o “Proyectos”. Cada entidad se caracteriza por sus **atributos**, que son las propiedades o características que la describen. Por ejemplo, un “Cliente” puede tener atributos como “Nombre”, “Dirección” y “Teléfono”.

Además, las entidades pueden estar relacionadas entre sí a través de **relaciones**, que describen cómo interactúan los objetos. Estas relaciones pueden ser de uno a uno, uno a muchos o muchos a muchos, dependiendo de la naturaleza de la interacción (Elmasri & Navathe, 2015).

También es importante mencionar las **cardinalidades**, que especifican el número de instancias de una entidad que pueden asociarse con instancias de otra entidad. Finalmente, los **diagramas ER** permiten representar gráficamente estos componentes, ofreciendo una visión clara y accesible de la estructura de la base de datos, lo que resulta esencial para su diseño y comprensión (Chen, 1976). Con estos componentes, el modelo ER proporciona una base sólida para la construcción de sistemas de bases de datos efectivos y bien organizados.

**Figura 5**

*Ejemplos de Cardinalidad*



Fuente: Elaboración propia

### Profundización del modelo Entidad Relación

El diagrama entidad-relación (ER) es una representación visual esencial en el diseño de bases de datos, utilizada para modelar la estructura de datos de un sistema de información. Este tipo de diagrama ayuda a conceptualizar cómo se organizan los datos y cómo las diferentes entidades se relacionan entre sí. En un diagrama ER, las **entidades** se representan como rectángulos y son los objetos clave dentro del sistema, como “Cliente”, “Producto” o “Pedido”. Cada entidad puede tener varios **atributos**, que describen sus características, y estos se representan como óvalos conectados a las entidades. Por ejemplo, un “Cliente” podría tener atributos como “Nombre”, “Dirección” y “Teléfono” (Elmasri & Navathe, 2015).

Las **relaciones** son otro componente crítico del diagrama ER, representadas por rombos. Estas relaciones indican cómo interactúan las entidades entre sí. Por ejemplo, un “Cliente” puede realizar múltiples “Pedidos”, y cada “Pedido” puede incluir varios “Productos”. Las relaciones pueden tener diferentes cardinalidades, como uno a uno, uno a muchos o muchos a muchos, lo que ayuda a definir la naturaleza de la interacción entre las entidades (Chen, 1976). Esta información es vital para garantizar la integridad de los datos y optimizar el rendimiento de las consultas en la base de datos.

El uso de diagramas ER no solo simplifica la complejidad de la base de datos, sino que también actúa como una herramienta de comunicación eficaz entre desarrolladores, analistas y otros interesados en el proyecto. Facilita la identificación de requisitos y la detección de inconsistencias en la estructura de datos antes de la implementación, lo que puede ahorrar tiempo y recursos en el desarrollo del sistema.

En resumen, el diagrama entidad-relación es una herramienta poderosa que proporciona una base sólida para el diseño y la comprensión de bases de datos, garantizando que los datos se gestionen de manera efectiva y lógica (Elmasri & Navathe, 2015).

En el siguiente enlace se explica de forma resumida el datawarehouse y el datalake, que son estructuras de datos que se nutren de los modelos de datos. Desde el minuto 13 hasta el minuto 17: [Datawarehouse y Datalakes](#).

### Ventajas y Limitaciones del Modelo Entidad Relación

El modelo entidad-relación (ER) es ampliamente utilizado en el diseño de bases de datos debido a sus

numerosas ventajas, aunque también presenta algunas limitaciones que es importante considerar. A continuación, se enumeran las principales ventajas y limitaciones del modelo ER.

### Ventajas

- **Visualización clara:** El diagrama ER proporciona una representación gráfica intuitiva de las entidades y sus relaciones, lo que facilita la comprensión de la estructura de la base de datos tanto para desarrolladores como para usuarios finales (Elmasri & Navathe, 2015).
- **Facilita el diseño:** Al utilizar el modelo ER, los diseñadores pueden identificar fácilmente las dependencias entre las entidades y los atributos, lo que ayuda a crear una base de datos bien estructurada y organizada.
- **Comunicación efectiva:** Los diagramas ER sirven como una herramienta de comunicación eficaz entre los diferentes interesados en un proyecto, permitiendo discutir y ajustar los requisitos del sistema antes de la implementación (Chen, 1976).
- **Adaptabilidad:** El modelo ER es flexible y se puede adaptar fácilmente a diferentes contextos y dominios, lo que permite su uso en una variedad de aplicaciones, desde sistemas de gestión de inventarios hasta bases de datos de clientes.

### Limitaciones

- **Complejidad en sistemas grandes:** A medida que aumenta el número de entidades y relaciones, los diagramas ER pueden volverse complejos y difíciles de interpretar, lo que puede dificultar su uso en grandes sistemas (Elmasri & Navathe, 2015).
- **No especifica detalles de implementación:** Aunque el modelo ER es excelente para representar la estructura lógica de los datos, no aborda los aspectos físicos de la implementación, como la optimización del rendimiento o la gestión de almacenamiento.
- **Dependencia del contexto:** La efectividad del modelo ER puede depender del dominio específico y de los requisitos del negocio. En algunos casos, modelos alternativos pueden ser más adecuados para ciertas aplicaciones (Date, 2004).
- **Requiere conocimientos técnicos:** Para crear y comprender adecuadamente un diagrama ER, se necesita un nivel de conocimiento técnico que no siempre está disponible en todos los usuarios o partes interesadas.

En conclusión, aunque el modelo entidad-relación ofrece ventajas significativas en el diseño y comprensión de bases de datos, también presenta limitaciones que deben ser consideradas. Con un uso adecuado y en el contexto correcto, el modelo ER puede ser una herramienta invaluable para el desarrollo de sistemas de información efectivos.

### Aplicaciones del Modelo Entidad-Relación

El modelo entidad-relación (ER) es una herramienta versátil y ampliamente utilizada en el diseño de bases de datos en diversas áreas. A continuación, se presentan algunas de las principales aplicaciones del modelo ER que destacan su utilidad en el desarrollo de sistemas de información:

- **Desarrollo de bases de datos relacionales:** El modelo ER es fundamental en la creación de bases de datos relacionales, donde se utiliza para definir las entidades, sus atributos y las relaciones entre ellas. Esta aplicación es crucial en sectores como la gestión empresarial, la educación y la salud (Elmasri & Navathe, 2015).
- **Sistemas de gestión empresarial (ERP):** En sistemas ERP, el modelo ER ayuda a estructurar la información relacionada con diferentes departamentos, como finanzas, recursos humanos y logística, asegurando que los datos fluyan de manera eficiente entre las distintas áreas de la

organización (Kumar et al., 2015).

- **Diseño de aplicaciones web:** El modelo ER se utiliza en el diseño de aplicaciones web para definir la estructura de la base de datos que respalda el sitio. Esto es especialmente relevante en plataformas de comercio electrónico, donde se gestionan productos, usuarios y transacciones (Sommerville, 2016).
- **Análisis de sistemas:** Durante la fase de análisis de sistemas, el modelo ER se utiliza para identificar los requisitos de información y cómo los diferentes componentes del sistema se relacionan entre sí. Esto ayuda a los analistas a comprender mejor el dominio del problema y a definir soluciones efectivas (Burch & Phillips, 2014).
- **Educación y formación:** El modelo ER se enseña en cursos de bases de datos y sistemas de información, siendo una herramienta pedagógica efectiva para ayudar a los estudiantes a comprender conceptos fundamentales de la gestión de datos y el diseño de bases de datos (Rob & Coronel, 2016).
- **Migración de datos:** Cuando se realizan migraciones de datos de un sistema a otro, el modelo ER ayuda a mapear las relaciones y dependencias de los datos en el sistema antiguo y a diseñar la estructura del nuevo sistema, garantizando que se conserven la integridad y la coherencia de la información (Elmasri & Navathe, 2015).

## Referencias

- Chen, P. P. (1976). The entity-relationship model—toward a unified view of data. *ACM Transactions on Database Systems*, 1(1), 9-36. <https://doi.org/10.1145/320434.320440>
- Date, C. J. (2004). *An introduction to database systems*. Pearson Education.
- Elmasri, R., & Navathe, S. B. (2015). *Fundamentals of database systems*. Pearson.
- Harrington, J. L. (2016). *Relational database design clearly explained*. Morgan Kaufmann.
- Kumar, A., Singh, A., & Singh, S. (2015). ERP implementation: A case study. *International Journal of Advanced Research in Computer Science and Software Engineering*, 5(6), 1075-1081.
- Rob, P., & Coronel, C. (2016). *Database systems: Design, implementation, & management*. Cengage Learning.
- Sharda, R., Delen, D., & Turban, E. (2014). *Business intelligence and analytics: Systems for decision support*. Pearson.
- Sommerville, I. (2016). *Software engineering*. Addison-Wesley.
- Turban, E., Sharda, R., & Delen, D. (2018). *Decision support and business intelligence systems*. Pearson.

## Definición de los términos citados

### **Tablas de Hechos**

Las tablas de hechos son tablas que almacenan información histórica o transaccional. Suelen contener campos como año, mes, día, fecha, entre otros.

### **Tablas de Dimensiones**

Las tablas de dimensiones, también conocidas como catálogos, son tablas que proporcionan descripciones adicionales a las tablas de hechos. Ejemplos: marca de coche, nombre de productos, dirección de proveedor, etc.



**La excelencia no se improvisa**

síguenos

