

Gobernanza y Modelado de Datos

El modelado de datos es una técnica fundamental que permite representar cómo se organizan, almacenan y relacionan los datos dentro de un sistema. Su objetivo es proporcionar una visión estructurada y lógica de los datos antes de que estos sean implementados físicamente.

Esta representación se realiza a través de diagramas o esquemas que describen entidades, atributos, relaciones y reglas de negocio. En esta presentación, exploraremos los fundamentos del modelado de datos, su gobernanza y las mejores prácticas para implementarlo correctamente.

 **por Kibernetum Capacitación S.A.**



Preguntas de activación



En la sesión anterior hablamos sobre la arquitectura de datos y sus componentes. ¿Cómo crees que el modelado de datos se inserta dentro de esa arquitectura?



¿Qué elementos consideras clave para que un modelo de datos sea comprensible por distintas áreas de una organización (por ejemplo, TI, marketing, operaciones)?



Si tuvieras que diseñar una base de datos desde cero, ¿cuál sería el primer paso que tomarías y por qué?

¿Qué es el Modelado de Datos?

Representación Estructurada

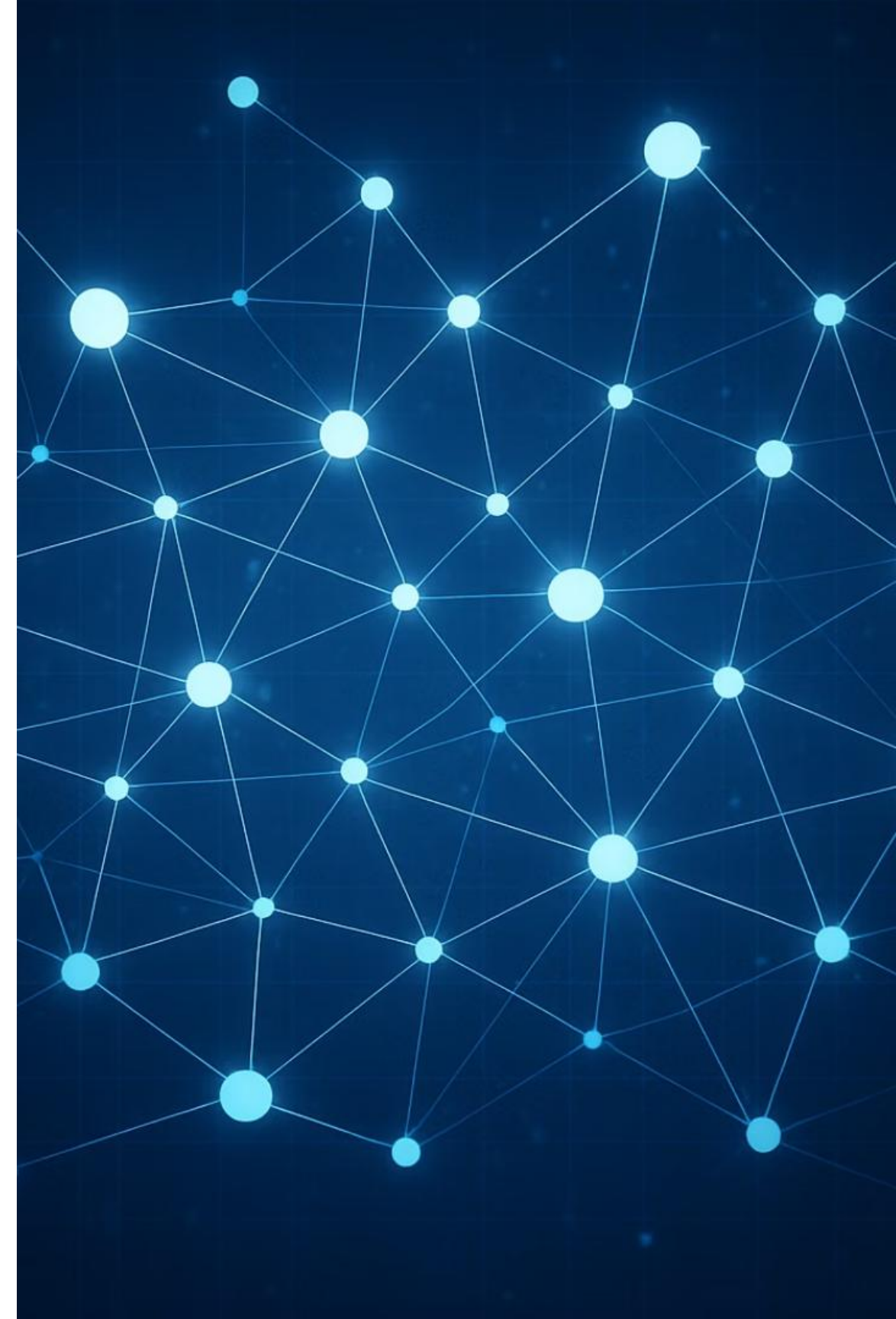
Técnica que permite visualizar cómo se organizan, almacenan y relacionan los datos dentro de un sistema antes de su implementación física.

Elementos Principales

Entidades (objetos del negocio), atributos (propiedades), relaciones (conexiones) y reglas de negocio (restricciones).

Objetivo

Proporcionar una visión estructurada y lógica de los datos para facilitar su implementación, mantenimiento y evolución.



Niveles de Modelado de Datos



Modelo Conceptual

Representación abstracta centrada en requerimientos del negocio. Colaboración con usuarios no técnicos.



Modelo Lógico

Traducción técnica del modelo conceptual. Define entidades, relaciones, claves primarias y foráneas, sin considerar un gestor específico.



Modelo Físico

Implementación directa en un sistema de gestión de base de datos (SGBD), con detalles técnicos como tipos de datos, índices y restricciones.

Modelo Conceptual: El Primer Paso



Orientado al Negocio

Utiliza lenguaje común para usuarios técnicos y no técnicos, facilitando la comunicación entre departamentos.



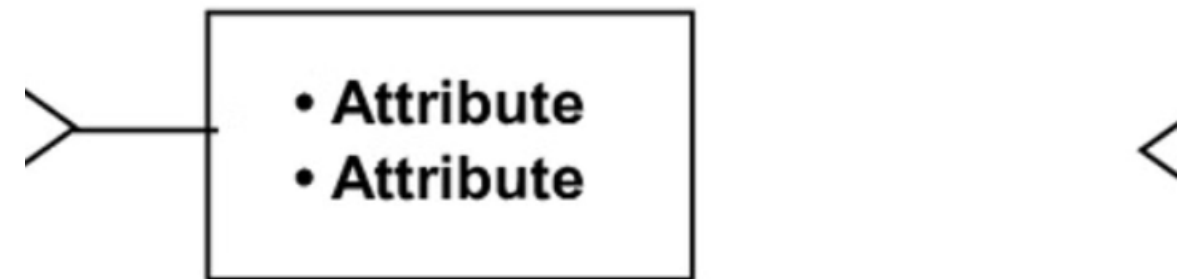
Representa Entidades y Relaciones

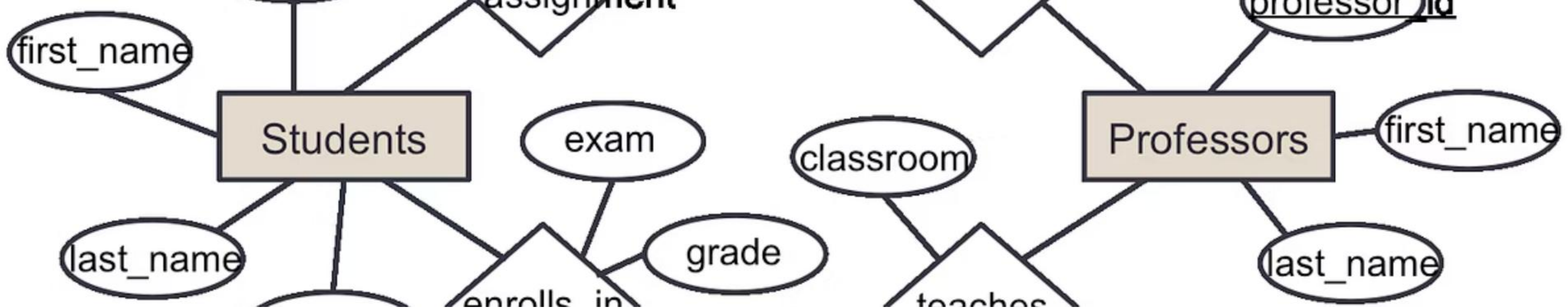
Identifica los objetos principales del negocio y cómo se relacionan entre sí, sin entrar en detalles técnicos.



Colaborativo

Se desarrolla en conjunto con los usuarios del negocio para asegurar que refleja correctamente los procesos y necesidades reales.

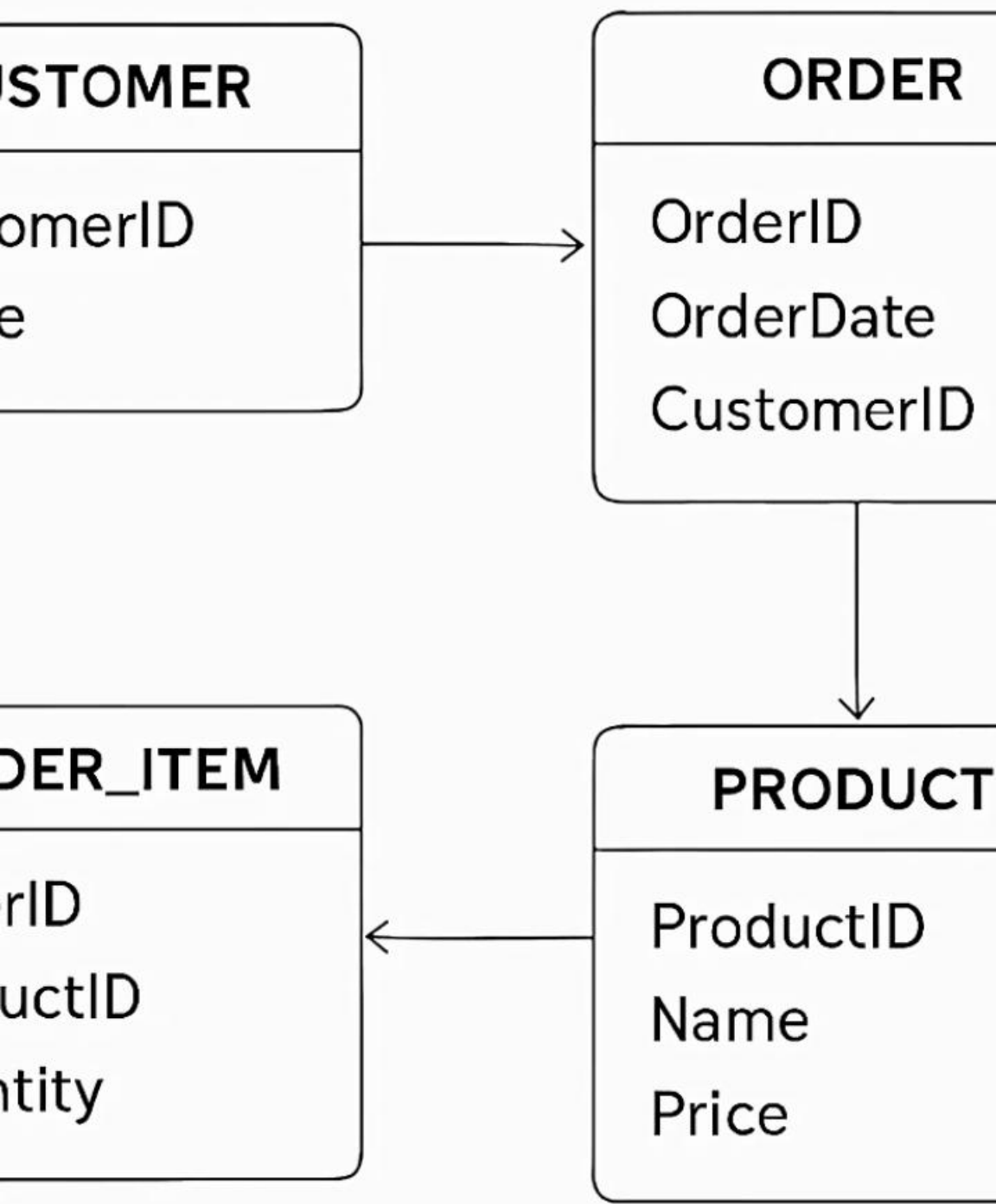




Ejemplo de Modelo Conceptual

Un modelo conceptual representa las entidades principales y sus relaciones de forma abstracta, sin detalles técnicos. Es el primer paso en el proceso de modelado de datos.

En este ejemplo de un sistema de ventas, podemos observar entidades como Cliente, Pedido y Producto, conectadas mediante relaciones que indican, por ejemplo, que "un cliente realiza muchos pedidos" y "un pedido contiene varios productos". Este nivel de abstracción facilita la comprensión por parte de usuarios no técnicos.

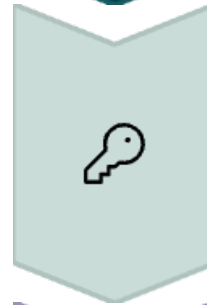


Modelo Lógico: La Traducción Técnica



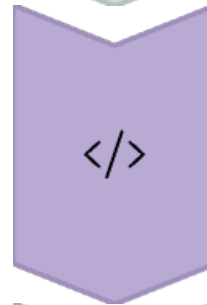
Traducción a Tablas

Convierte las entidades conceptuales en tablas con estructura definida.



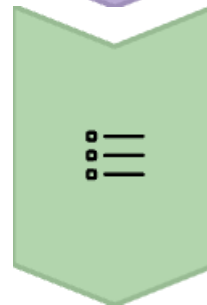
Definición de Claves

Establece claves primarias y foráneas para identificar registros y establecer relaciones.



Independencia Tecnológica

Mantiene independencia del sistema de gestión de base de datos específico.



Atributos Detallados

Define los campos específicos de cada tabla y sus características generales.

Ejemplo de Modelo Lógico

Tabla Cliente

ID_Cliente (PK)

Nombre

Dirección

Tabla Pedido

ID_Pedido (PK)

Fecha

Monto

ID_Cliente (FK → Cliente)

El modelo lógico traduce el conceptual a una estructura de tablas con atributos específicos. Aquí vemos cómo las entidades Cliente y Pedido se convierten en tablas con campos definidos, incluyendo claves primarias (PK) y foráneas (FK) que establecen las relaciones entre ellas.

Modelo Físico: La Implementación Real

Implementación en SGBD

Adaptación a un sistema de gestión de base de datos específico como MySQL, PostgreSQL o Snowflake

Restricciones

Implementación de restricciones como NOT NULL, UNIQUE, CHECK para garantizar la integridad



Tipos de Datos Específicos

Definición de tipos concretos como VARCHAR, INT, DATE según el motor de base de datos

Índices y Optimización

Creación de índices, particiones y otros elementos para mejorar el rendimiento

Ejemplo de Modelo Físico en SQL

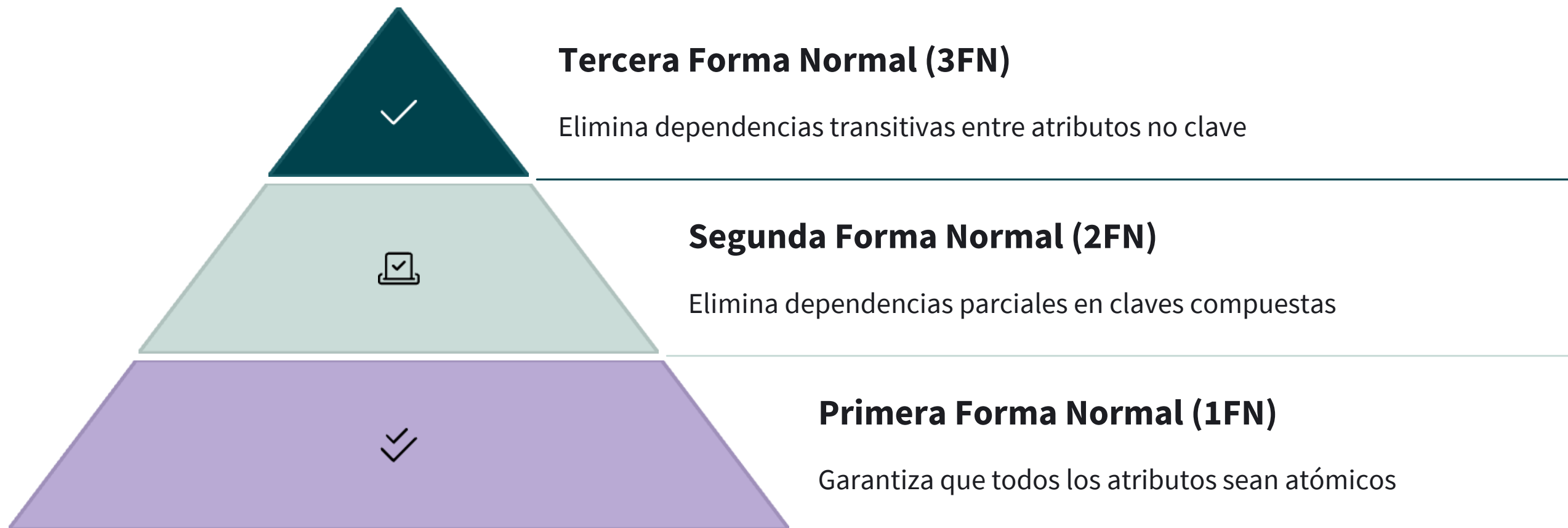
```
3 CREATE TABLE datos;
3     direccion VARCHAR(100) NOT NULL,
4     id_cliente INT NOT NULL,
3     id_producto INT NOT NULL,
3     fecha DATE NOT NULL,
7 }
4 CREATE TABLE pedidos;
4     id_pedido INT NOT NULL,
6     id_cliente INT NOT NULL,
7     id_producto INT NOT NULL,
7     fecha DATE NOT NULL,
7     monto DECIMAL(10,2) NOT NULL,
7     FOREIGN KEY (id_cliente) REFERENCES clientes(id_cliente),
7     FOREIGN KEY (id_producto) REFERENCES productos(id_producto),
7     FOREIGN KEY (fecha) REFERENCES fechas(fecha),
7     PRIMARY KEY (id_pedido)
7 }
```

```
CREATE TABLE Clientes (
  ID_Cliente INT PRIMARY KEY,
  Nombre VARCHAR(100) NOT NULL,
  Direccion VARCHAR(200),
  Ciudad VARCHAR(50),
  Codigo_Postal VARCHAR(10)
);

CREATE TABLE Pedidos (
  ID_Pedido INT PRIMARY KEY,
  Fecha DATE NOT NULL,
  Monto DECIMAL(10,2),
  ID_Cliente INT,
  FOREIGN KEY (ID_Cliente) REFERENCES Clientes(ID_Cliente)
);
```

El modelo físico traduce el diseño lógico a instrucciones específicas para un sistema de gestión de base de datos. En este ejemplo vemos código SQL para crear las tablas Clientes y Pedidos con sus tipos de datos específicos, restricciones y relaciones.

Normalización: Optimizando el Diseño



Ejemplo de Normalización

Tabla No Normalizada

ID_Cliente	Nombre	ID_Pedido	Fecha	Ciudad	Código_Postal
1	Juan	101	45597	Santiago	12345
2	María	102	45598	Valparaíso	54321
1	Juan	103	45599	Santiago	12345

Normalizada (3FN)

Tabla Clientes:

ID_Cliente	Nombre	Ciudad
1	Juan	Santiago
2	María	Valparaíso

Tabla Pedidos:

ID_Pedido	Fecha	ID_Cliente
101	1/11/2024	1
102	2/11/2024	2
103	3/11/2024	1

Tabla Ciudades:

Ciudad	Código_Postal
Santiago	12345
Valparaíso	54321

Metas del Modelado de Datos

Organización Estructurada

Definir cómo se almacenan los datos de forma lógica, clara y sin ambigüedades para facilitar su gestión y acceso.

Consistencia Semántica

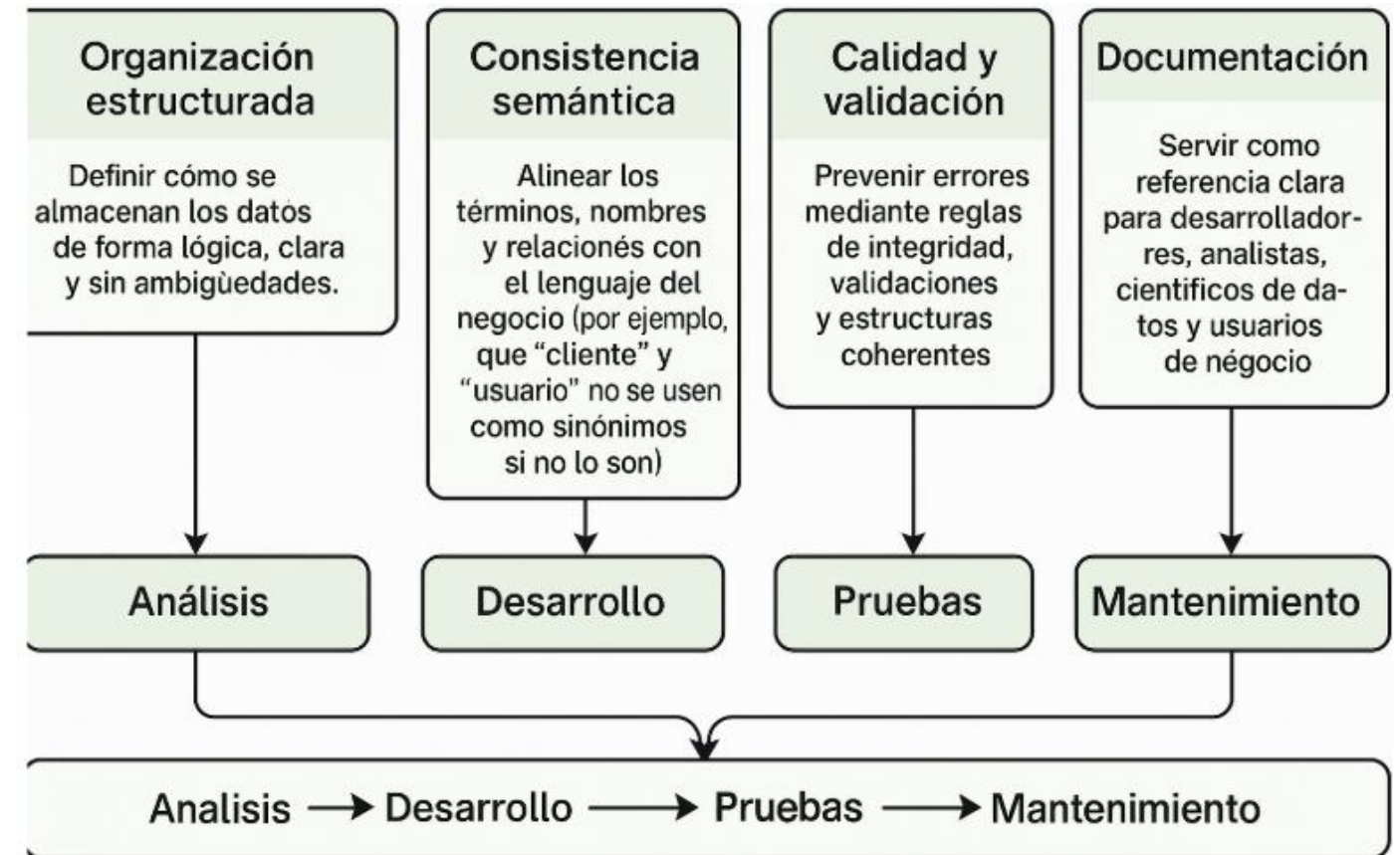
Alinear términos, nombres y relaciones con el lenguaje del negocio, evitando ambigüedades como usar "cliente" y "usuario" como sinónimos si no lo son.

Calidad y Validación

Prevenir errores mediante reglas de integridad, validaciones y estructuras coherentes que garanticen la fiabilidad de los datos.

Facilitar el Mantenimiento

Permitir que los modelos puedan ser modificados o extendidos sin grandes impactos en los sistemas existentes.





Caso Práctico: Problemas por Falta de Modelado

Fragmentación Terminológica

Una empresa de logística utiliza tres sistemas diferentes, cada uno con su propia forma de referirse al mismo concepto: "Cliente" (ventas), "Receptor" (envíos) y "Persona jurídica" (contabilidad).

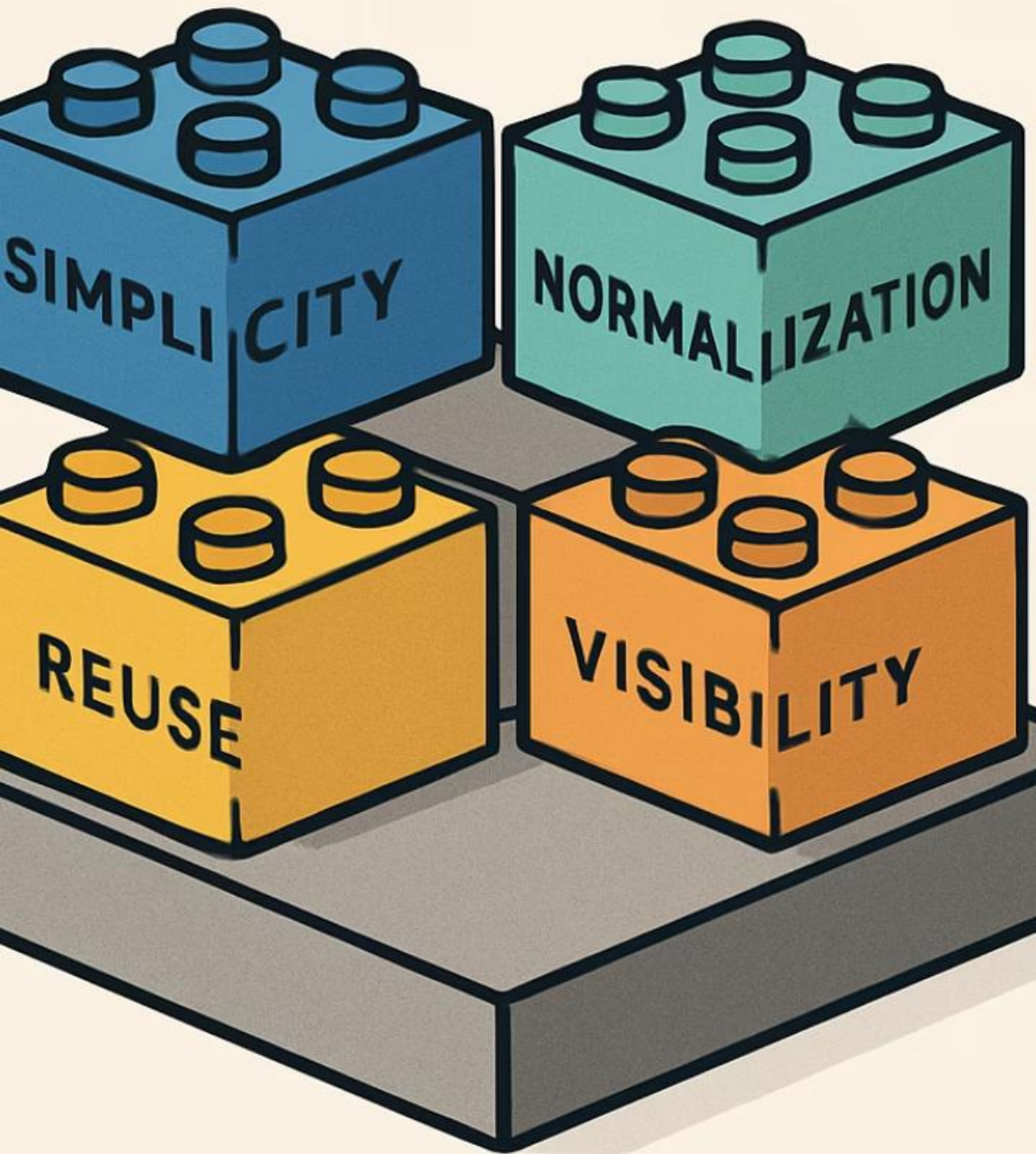
Consecuencias

Duplicidad de datos, errores operativos al no coincidir facturas con destinatarios reales, y análisis fragmentado por la imposibilidad de hacer reportes cruzados entre áreas.

Solución

Un modelo de datos bien diseñado y documentado que establezca términos estandarizados, proporcionando consistencia, autonomía y eficiencia en la gestión de datos.

DATA MODELING PRINCIPLES



Principios Clave del Modelado



Simplicidad

Un modelo de datos no debe ser más complicado de lo necesario. Si puedes representar algo en una tabla clara, evita estructuras rebuscadas.



Normalización Útil

La normalización mejora la integridad y evita duplicaciones. Sin embargo, si la desnormalización mejora el rendimiento en consultas frecuentes, puede aplicarse conscientemente.



Reutilización

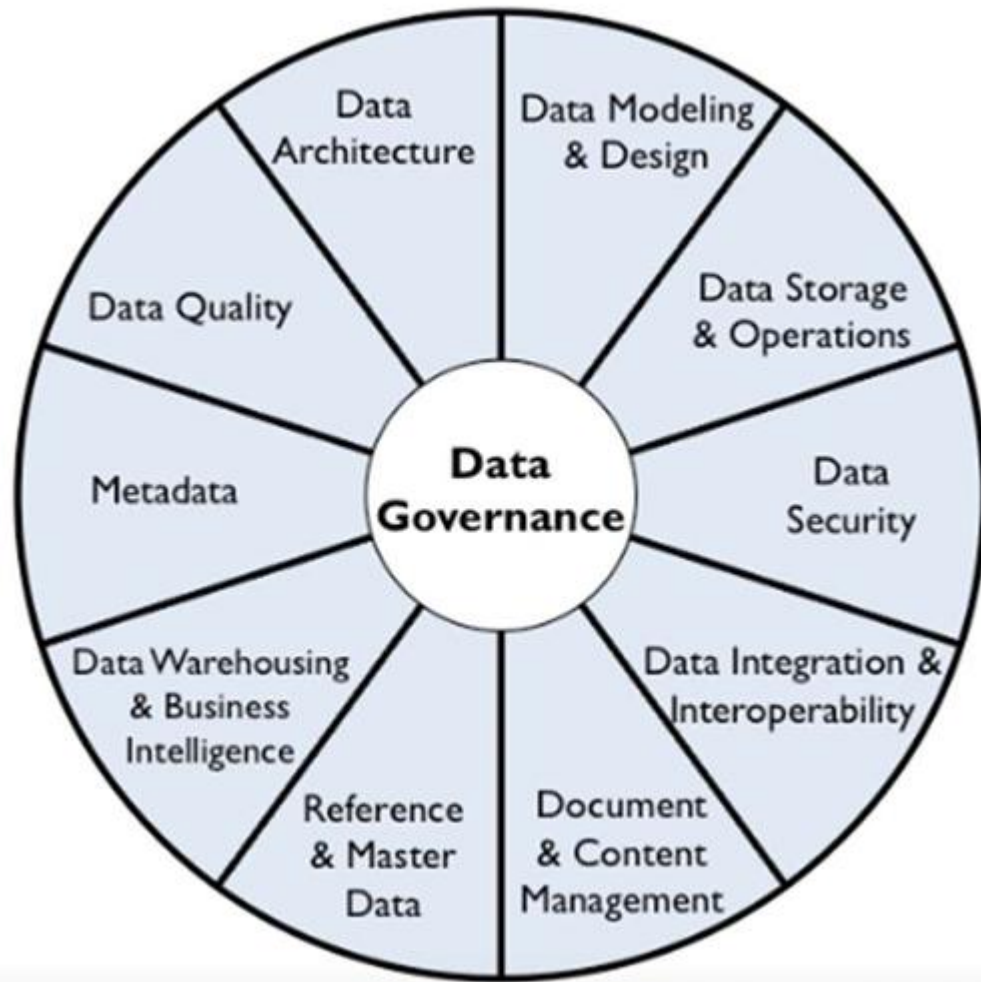
Si varios sistemas manejan "Clientes", usa una definición común. Esto facilita la interoperabilidad y el mantenimiento del modelo.



Visibilidad

El modelo debe estar documentado y disponible. No sirve de nada un modelo "perfecto" si nadie lo entiende ni lo puede revisar.

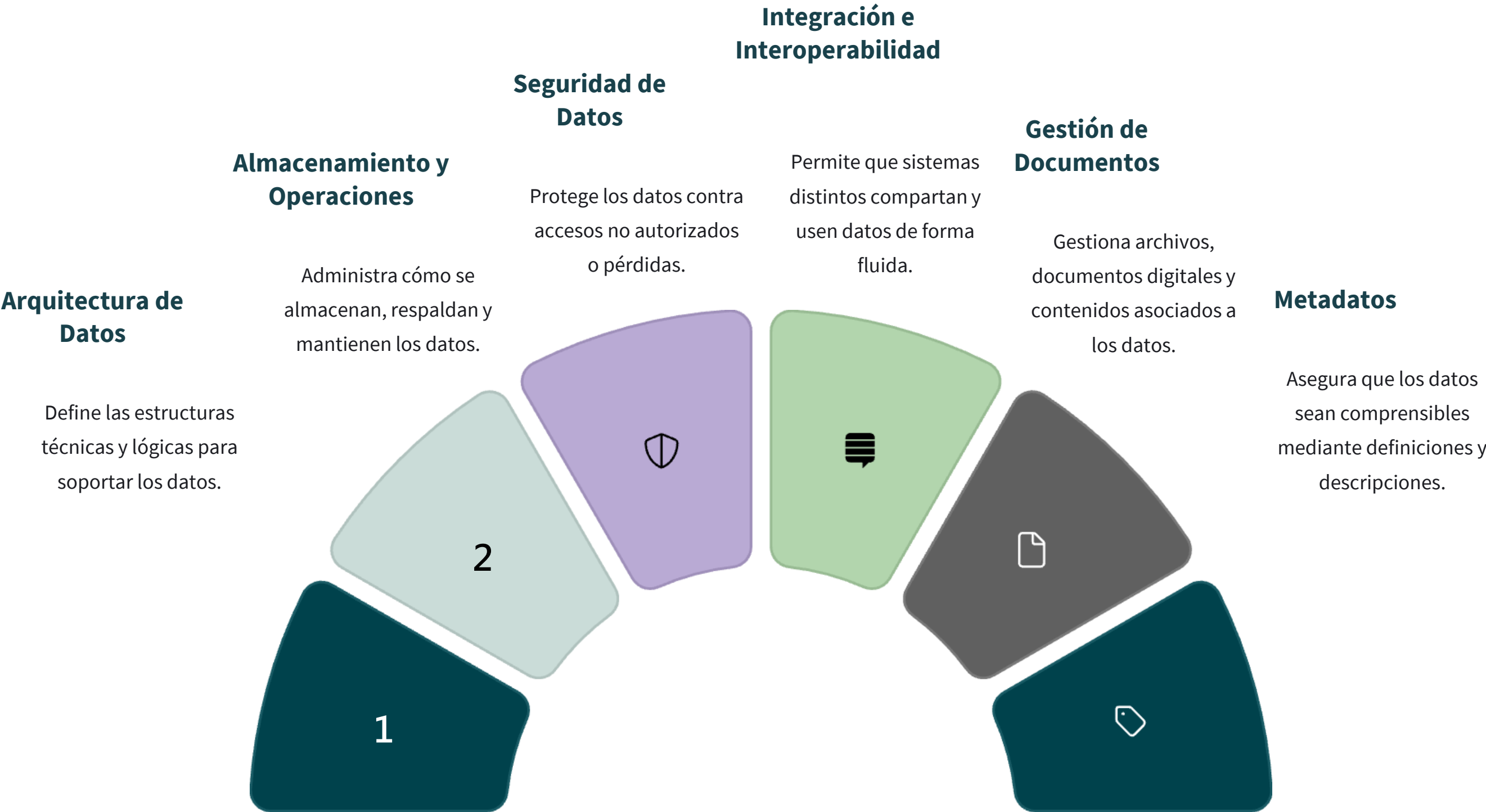
Gobernanza del Modelo de Datos



La gobernanza del modelo de datos garantiza que los modelos sean coherentes, documentados, versionados, reutilizables y validados en toda la organización.

La rueda DAMA-DMBOK muestra que la gobernanza no es un área aislada, sino una función transversal que coordina, regula y da coherencia a todas las demás disciplinas de la gestión de datos, incluyendo la arquitectura, el modelado, el almacenamiento, la seguridad, la integración, la gestión documental, los datos maestros, el data warehousing, los metadatos y la calidad.

Áreas Críticas de la Gobernanza de Datos



Importancia de la Gobernanza de Modelos

Sin una estrategia de gobernanza, los modelos de datos pueden volverse inconsistentes, obsoletos o contradictorios entre áreas.

Esto trae consecuencias como:

- Retrabajo entre equipos (cada uno “reinventa” el modelo).
- Fallos de interpretación (los datos significan cosas distintas según el área).
- Dificultades para auditar o mantener sistemas de análisis o reporting.
- La falta de gobernanza no solo afecta a los datos, **afecta a la toma de decisiones.**

Buenas Prácticas de Gobernanza

Repositorio Centralizado

Una fuente única de verdad para acceder, consultar y modificar modelos, garantizando que todos los equipos trabajen con la misma versión.

Control de Versiones

Cada cambio queda registrado, documentado y puede revertirse si es necesario, manteniendo la trazabilidad de las modificaciones.

Validaciones Colaborativas

Los modelos se revisan en conjunto por equipos técnicos y de negocio, asegurando que cumplan tanto requisitos técnicos como funcionales.

Estándares de Nomenclatura

Las entidades y atributos siguen reglas claras y consistentes, facilitando la comprensión y evitando ambigüedades.



Roles Clave en la Gobernanza del Modelo



Data Architect

Diseña la arquitectura de datos y valida modelos de alto nivel, asegurando que se alineen con la estrategia tecnológica de la organización.



Data Modeler

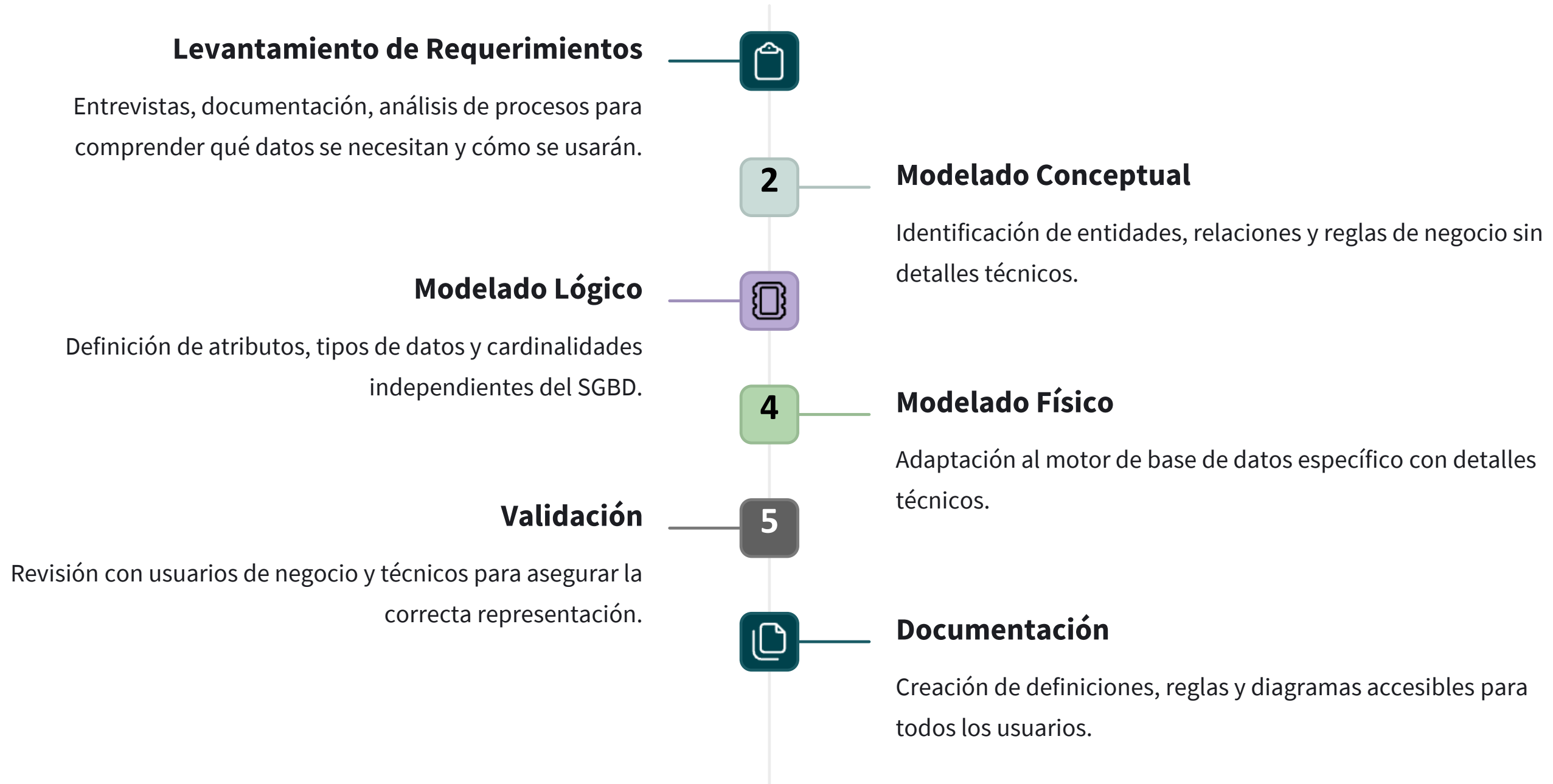
Implementa y documenta los modelos conceptuales, lógicos y físicos, traduciendo los requisitos del negocio en estructuras de datos.



Data Steward

Asegura la calidad del modelo desde el punto de vista del negocio, actuando como puente entre los equipos técnicos y los usuarios finales.

Actividades de Modelado de Datos



Levantamiento de Requerimientos



Entrevistas

Conversaciones con usuarios y áreas funcionales para entender sus necesidades de datos.



Documentación

Revisión de documentos existentes como formularios, reportes y catálogos.



Análisis de Procesos

Estudio de los flujos de trabajo para identificar datos necesarios en cada etapa.



Reglas de Negocio

Identificación de restricciones y políticas que afectan a los datos.



Validación del Modelo



Representación de Procesos

Verificar que el modelo refleja correctamente los procesos reales del negocio y captura todos los datos necesarios.



Claridad de Nomenclatura

Comprobar que los nombres de entidades y atributos son claros, coherentes y comprensibles para todos los usuarios.



Cumplimiento de Reglas

Asegurar que el modelo implementa correctamente todas las reglas de negocio identificadas durante el levantamiento de requerimientos.



Viabilidad Técnica

Evaluar si el modelo es implementable desde el punto de vista técnico y si cumple con los requisitos de rendimiento.



Documentación y Publicación

Elementos Esenciales

- Glosario de entidades con definiciones claras
- Diccionario de datos con descripción de atributos
- Diagramas visuales de relaciones
- Información de versionado y autores
- Reglas de negocio implementadas
- Guías de uso y ejemplos

Beneficios

- Facilita la comprensión del modelo por nuevos miembros
- Reduce dependencia de conocimiento tribal
- Permite reutilización en otros proyectos
- Sirve como referencia para resolución de problemas
- Apoya los procesos de auditoría y cumplimiento

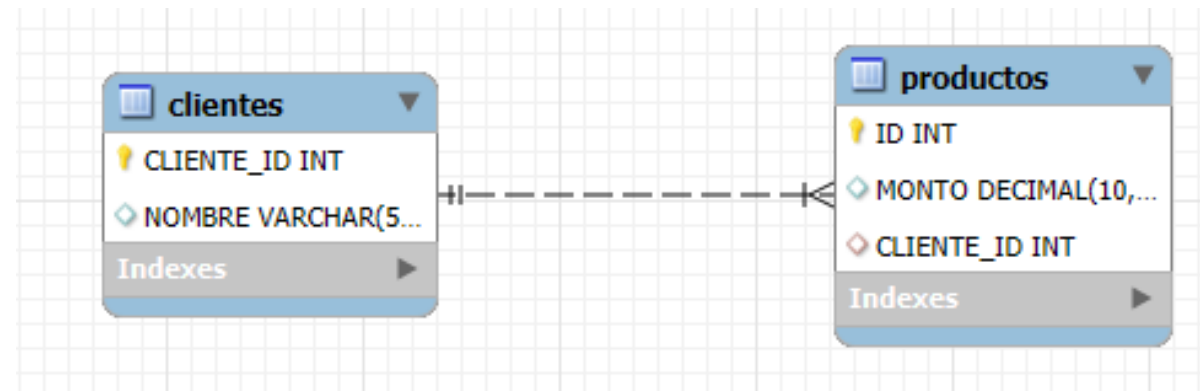
Herramientas de Modelado de Datos



Existen diversas herramientas especializadas para el modelado de datos, desde opciones comerciales robustas como ER/Studio y PowerDesigner, hasta alternativas gratuitas o freemium como Lucidchart, MySQL Workbench, dbdiagram.io y Draw.io. Cada una ofrece diferentes funcionalidades para diseñar, documentar y mantener modelos de datos según las necesidades específicas del proyecto.

Mini Actividad

En la siguiente imagen identifica los elementos que describen el esquema, reconoces las entidades, las relaciones, etc.



Mejores Prácticas según DAMA-DMBOK

Revisión Constante con Stakeholders

Involucrar a usuarios de negocio, analistas y técnicos en iteraciones del modelo mediante reuniones periódicas para asegurar que cumple con las necesidades reales.

Versionamiento Controlado

Registrar cada cambio con fecha, autor y propósito utilizando herramientas como Git o repositorios internos de modelos para mantener la trazabilidad.

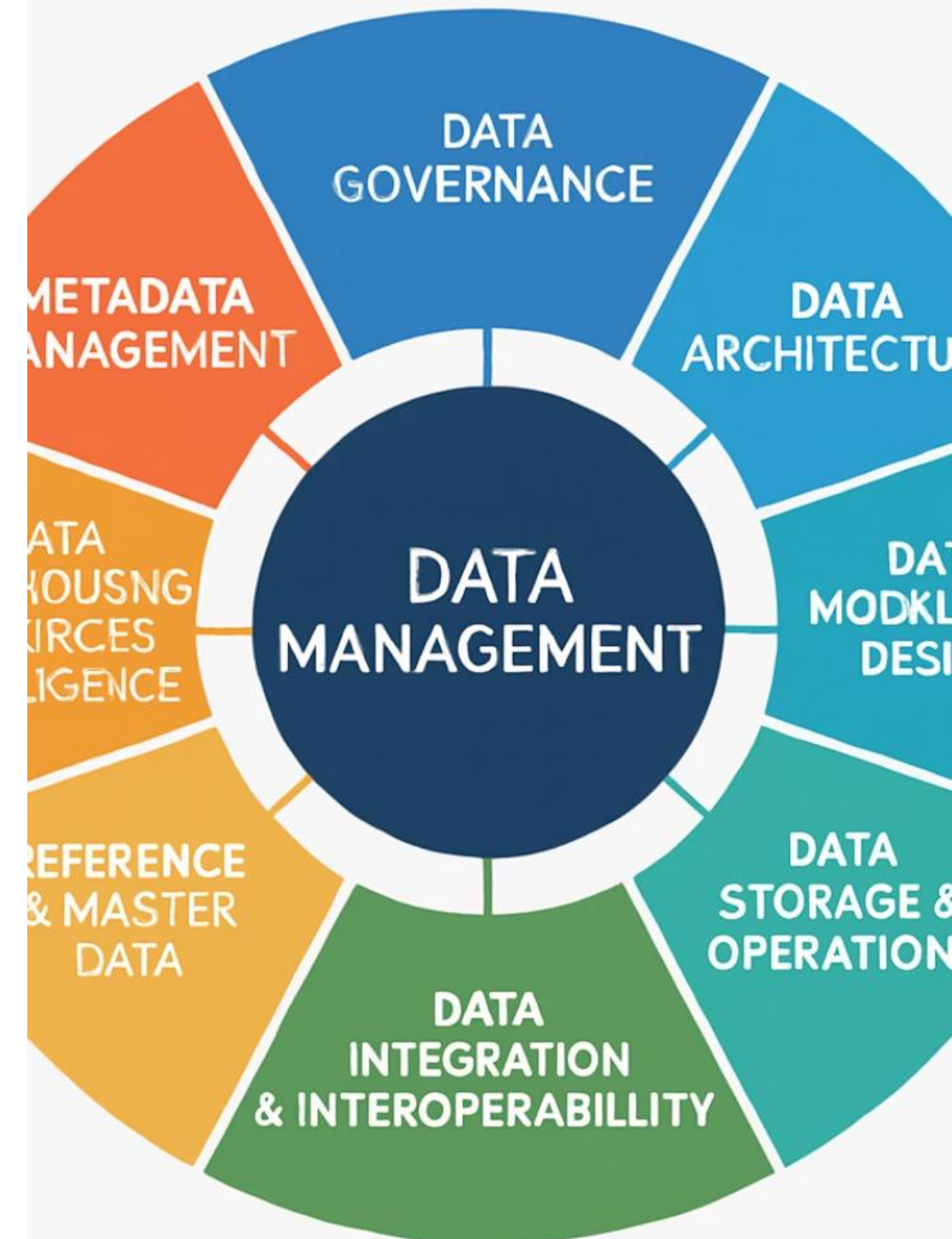
Nomenclatura Estándar


Establecer un lenguaje común en nombres de entidades, atributos y relaciones, con reglas claras como el uso de prefijos o la consistencia en singular/plural.

Documentación Centralizada

Mantener un diccionario de datos accesible a todos los equipos, con definiciones, formatos, orígenes y reglas de validación claramente especificados.

DATA MANAGEMENT FRAMEWORK





Customer

CRM

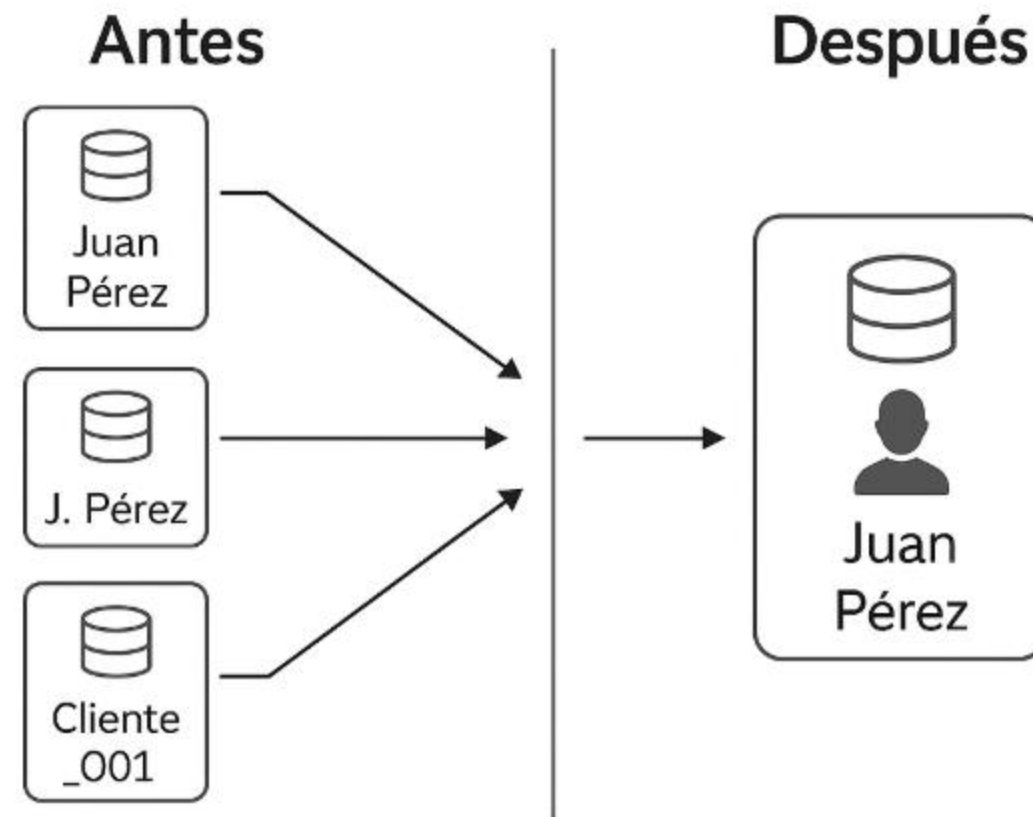
Datos Maestros (MDM)

La gestión de datos maestros (Master Data Management, MDM) busca garantizar que los datos fundamentales compartidos entre sistemas estén uniformes, precisos, accesibles y gobernados.

Los datos maestros representan entidades clave del negocio como clientes, productos, proveedores, ubicaciones y empleados, que deben ser consistentes en toda la organización. Sin una estrategia MDM, las organizaciones enfrentan problemas como facturación duplicada, errores de inventario, problemas de cumplimiento legal y análisis poco confiables.

Dominios Comunes de Datos Maestros

Visión unificada con MDM: del caos a la coherencia



MASTER DATA MANAGEMENT

Data

Data

Model

Recomendaciones de Modelado para MDM

Definir Entidades Maestras como Tablas Clave

Las tablas como Cliente, Producto y Ubicación deben ser diseñadas para ser reutilizadas en múltiples sistemas, no duplicadas en cada aplicación.

Establecer Claves Únicas

Utilizar identificadores como RUT o SKU que aseguren la identificación unívoca de los registros a través de todos los sistemas.

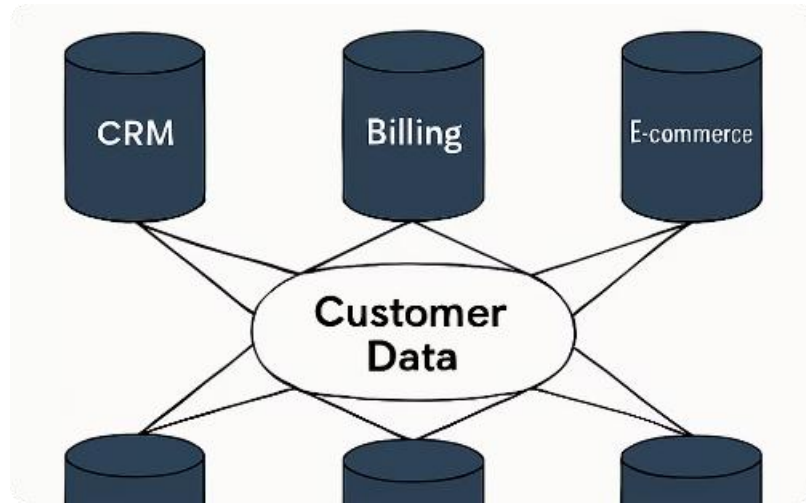
Usar Relaciones Foráneas

Implementar relaciones entre tablas que impidan la creación de datos huérfanos o inconsistentes, manteniendo la integridad referencial.

Validar Atributos con Reglas de Negocio

Aplicar restricciones como formatos específicos para RUT o validación de que una comuna pertenezca a una región determinada.

Unificación de Datos mediante MDM



Antes de MDM

Datos fragmentados en múltiples sistemas, con diferentes formatos, nomenclaturas y niveles de actualización. Cada departamento mantiene su propia versión de la verdad, generando inconsistencias y errores operativos.



Después de MDM

Datos unificados en un repositorio central que actúa como fuente única de verdad. Todos los sistemas consumen información estandarizada, garantizando consistencia en operaciones, análisis y reportes a través de la organización.



Beneficios

Mejora en la calidad de datos, reducción de errores operativos, análisis más precisos, mejor experiencia de cliente y cumplimiento normativo más eficiente gracias a la consistencia y trazabilidad de la información.

STRATEGIC VALUE OF DATA MODELING



Conclusiones: El Valor Estratégico del Modelado



Fundamento para Sistemas Robustos

Un buen modelado de datos es la base para construir sistemas de información coherentes, escalables y alineados con los objetivos del negocio.



Puente entre Negocio y Tecnología

El modelado actúa como lenguaje común entre áreas técnicas y funcionales, facilitando la comunicación y el entendimiento mutuo.



Habilitador de Análisis Avanzados

Modelos bien diseñados permiten análisis más profundos, reportes consistentes y una mejor toma de decisiones basada en datos.



Protección de Activos de Datos

La gobernanza de modelos asegura que los datos, como activo estratégico, mantengan su valor, integridad y utilidad a lo largo del tiempo.

Enlaces de Interés

- ✓ **Video:** [Tutorial - Diagrama Entidad-Relación \(ER\)](#)
- ✓ **Video:** [Video 1 - Modelado de Datos - Introducción](#)
- ✓ Video: [Desnormalización](#)
- ✓ **Video:** [Qué es Master Data Management \(MDM\)?](#)
- ✓ Artículo normalización: [DBMS Normalization: 1NF, 2NF, 3NF Database Example](#)



Preguntas de reflexión



Impacto de la falta de gobernanza

¿Qué impacto tendría la falta de gobernanza en los modelos de datos de una organización que trabaja con múltiples sistemas y usuarios?



Elementos para modelar bases de datos

¿Qué elementos aprendidos hoy aplicarías primero si tuvieras que modelar la base de datos para una nueva aplicación?



Relación con la calidad de datos

¿Cómo se relacionan los conceptos de normalización, nomenclatura y MDM con la calidad de los datos y la toma de decisiones?