



TIC



MAE.GE.AI.04 - Guía para elaborar Modelos de datos / Mapas de información

Ministerio de Tecnologías de la Información y las Comunicaciones

MAE

Carina Murcia Yela – Ministra de Tecnologías de la Información y las Comunicaciones
Giovanny Andrés López Cabezas - Viceministro de Transformación Digital
Lucy Elena Urón Rincón - Directora de Gobierno Digital
Lucy Estella Palacios Valoyes - Subdirectora de Estándares y Arquitectura de TI
Jairo Riascos Muñoz – Contratista Subdirección de Estándares y Arquitectura de TI
Olga Regina Garzón Pupo- Profesional Subdirección de Estándares y Arquitectura de TI
Andrés Mauricio Ruiz Gómez - Profesional Subdirección de Estándares y Arquitectura de TI
Nelson A. Pardo Moreno - Profesional Subdirección de Estándares y Arquitectura de TI
Johan Barrios Rodgers - Club de Gobierno
Ana María Gutiérrez Urquijo - Club de Gobierno
Johana Patricia García Poveda – Universidad Nacional sede Medellín

Ministerio de Tecnologías de la Información y las Comunicaciones
 Viceministerio de Transformación Digital
 Dirección de Gobierno Digital
 Subdirección de Estándares y Arquitectura de TI

Versión	Observaciones
Versión 1 15/12/2025	Guía que orienta la documentación de la arquitectura de información

Comentarios, sugerencias o correcciones pueden ser enviadas al correo electrónico:
sopORTECCC@mintic.gov.co

Guía conformación de equipos de arquitectura en entidades públicas V 1.0
 Este documento de la Dirección de Gobierno Digital se encuentra bajo una Licencia Creative Commons Atribución 4.0 Internacional

Tabla de contenido

Tabla de contenido	3
Listado de Tablas	4
Tabla de ilustraciones	4
Guía para elaborar Modelos de Datos / Mapas de Información.....	5
1. Derechos De Autor.....	5
2. Audiencia	5
3. Introducción	5
4. Justificación	6
5. Objetivo.....	6
6. Alcance	7
7. Conceptos de modelado de datos / mapa de información	7
7.1. Modelo conceptual.....	7
7.2. Modelo lógico.....	8
7.3. Modelo físico.....	9
7.4. Modelo Ontológico.....	9
7.5. Modelo metadatos.	10
8. Herramientas que apoyan los modelos.....	11
9. Buenas prácticas para el modelado.....	14
10. Estándar de industria para los modelos.....	17
11. Ejemplos modelos de datos	19
12. Referencias.....	25

Listado de Tablas

Tabla 1. Modelo metadatos – Elementos descriptivos	23
Tabla 2. Modelo metadatos – Ejemplo parte 1	24
Tabla 3. Modelo metadatos – Ejemplo parte 2	24

Tabla de ilustraciones

Ilustración 1. Modelo Conceptual	20
Ilustración 2. Modelo lógico.....	21
Ilustración 3. Modelo físico.....	22
Ilustración 4. Modelo ontológico.....	22

Guía para elaborar Modelos de Datos / Mapas de Información

1. Derechos De Autor

Para el desarrollo de esta guía lineamiento, se recogieron aspectos importantes de mejores prácticas y documentos de uso libres tales como el Marco de Referencia de Arquitectura Empresarial (MRAE), la Guía de Arquitectura de Información del MINTIC, Lenguaje Común de Intercambio de Información (LCCI), TOGAF (The Open Group Architecture Framework), Estándares de Metadatos y Datos Abiertos - ISO/IEC 11179, DAMA-DMBOK (Data Management Body of Knowledge) y material divulgativo sobre modelado de datos, referencias sobre ontologías en gobierno electrónico, guías de metadatos para datos abiertos, listados de herramientas líderes en modelado.

2. Audiencia

Entidades públicas de orden nacional y entidades públicas del orden territorial, así como proveedores de servicios de Gobierno Digital, y terceros que deseen adoptar el Marco de Referencia de Arquitectura Empresarial V3.0.

3. Introducción

El Ministerio de Tecnologías de la Información y las Comunicaciones – MINTIC presenta esta nueva “Guía para elaborar modelos de datos / mapas de información” con el fin de proporcionar orientación a todas las entidades públicas de Colombia en la elaboración de los modelos de datos / mapas de información, para transformar datos crudos en información valiosa, desde el entendimiento de necesidades de negocio hasta la creación de modelos conceptuales, lógicos y físicos, garantizando sistemas escalables, eficientes y confiables, y abordando tanto la estructura como las relaciones, crucial para toma de decisiones.

En una Entidad pública, la arquitectura de datos es fundamental para mapear y gestionar toda la información relacionada con sus procesos misionales, estratégicos, de apoyo y de atención al ciudadano. Esto abarca desde datos operativos y maestros hasta información documental (expedientes, registros) y datos abiertos publicados al público. Para los arquitectos de datos, es vital construir modelos de datos y mapas de información que brinden una visión única e integrada de todos estos datos, sirviendo como fuente de verdad para la institución.

Esta guía ayuda a que la arquitectura de datos construya exactamente lo que el Negocio necesita. Evita el costoso error de construir un sistema para luego descubrir que faltaron datos clave. Convierte 'datos sueltos' en respuestas útiles para la entidad.

4. Justificación

Las orientaciones expuesta en este documento apoya a las Entidades en la transformación de los datos brutos en conocimiento estratégico, la alineación del negocio y TI mediante un lenguaje común, reduciendo errores, optimizando el rendimiento, eliminando la redundancias mejorando la coherencia y la eficiencia, y la creación de una "fuente única" para la toma de decisiones y el desarrollo de sistemas, al ser un plano esencial para bases de datos y análisis, garantizando escalabilidad y calidad. A través de los modelos de datos las Entidades identifican qué datos existen, cómo se relacionan y cómo se usarán, desde un nivel conceptual hasta la implementación física en una base de datos.

Esta guía proporciona recomendaciones y pautas que deben ser adoptadas por las Entidades para gestionar la complejidad de los datos modernos, convirtiéndolos en un activo medible y accionable para la Entidad, asegurando que todos hablen el mismo "lenguaje de datos"

5. Objetivo

El objetivo principal del presente documento es orientar la creación de una representación visual clara, estructurada y consistente de la información de la Entidad, definiendo entidades, atributos y relaciones para optimizar su almacenamiento, comprensión, análisis y gestión eficiente en bases de datos, lo que facilita la comunicación entre áreas de negocio y dirección de TI, y reduce errores en el desarrollo de sistemas.

Los objetivos específicos del presente documento son:

- Establecer un lenguaje común y un diseño unificado para los datos de la empresa.
- Diseñar bases de datos para un procesamiento y consulta eficientes.
- Proporcionar una visión integral de la información para análisis de negocio.
- Servir como base para que los equipos técnicos construyan sistemas de información precisos.
- Detectar redundancias o datos faltantes que causan gastos o problemas.

6. Alcance

La presente guía documenta los conceptos clave de modelado (modelo conceptual, lógico, físico, ontológico y metadatos), herramientas sugeridas, buenas prácticas y estándares de la industria, alineados principalmente con el Marco de Referencia de Arquitectura Empresarial (MRAE) del MinTIC, complementado con las mejores prácticas internacionales.

7. Conceptos de modelado de datos / mapa de información

Los mapas de información definen el “**qué**”, define los conceptos de negocio y el flujo de información de forma abstracta para que todos lo entiendan y el modelo de datos define el “**como**” se estructuran y almacenan los datos en un sistema.

El proceso de modelado de datos suele dividirse en tres niveles clásicos – modelo conceptual, lógico y físico – con unos componentes adicionales como el modelo ontológico para capturar la semántica y el modelo metadatos que proporcionando información contextual sobre los datos. Cada nivel tiene un propósito específico y un grado de detalle distinto, que debe ser validado, optimizado y documentado. A continuación, se describen estos modelos:

7.1. Modelo conceptual

A través del modelo de datos conceptual se representa la esencia del negocio y se enfoca en el “**qué**” de la información, sin entrar en los detalles de la implementación. El modelo conceptual es una representación abstracta e independiente de la tecnología que define los conceptos o entidades importantes y sus relaciones a alto nivel.

El objetivo principal del modelo es lograr una comprensión común de la información del negocio en términos entendibles para los interesados no técnicos, sirviendo como un mapa conceptual del dominio. Un modelo conceptual correctamente elaborado ayuda a alinear a todos (usuarios del negocio, analistas y técnicos) sobre los datos relevantes y sus definiciones, convirtiéndose en una herramienta de apoyo al proyecto.

Este modelo suele representarse mediante diagramas de entidades y relaciones (ER) u otros mapas de información de alto nivel, y no incluye detalles como tipos de datos o claves. Se recomienda elaborarlo colaborativamente con los expertos del negocio, asegurando que refleje correctamente las necesidades misionales y estratégicas de la Entidad.

Un ejemplo, es identifica entidades tales como *Ciudadano*, *Servicio*, *Trámite* o *Documento* y las relaciones entre ellas, junto con sus atributos más relevantes a nivel de negocio.

Actividades para generar le modelo.

- Definir el objetivo y alcance.
- Identificar entidades.
- Identificar atributos.
- Establecer relaciones.

7.2. Modelo lógico.

El modelo de datos lógico se acerca más hacia el diseño, enfocándose en el “**cómo**” se organizarán los datos de forma aún independiente de una tecnología específica. A partir del modelo conceptual definido, el modelo lógico detalla las entidades con todos sus atributos, claves primarias y foráneas, y las relaciones con sus cardinalidades, aplicando reglas de negocio y de integridad.

El objetivo principal del modelo es actuar como un puente entre la visión conceptual y la implementación física, permitiendo entender cómo se estructurará la información para su posterior construcción. En este nivel se normalizan los datos (evitando redundancias) y se especifican estructuras más formales que eventualmente soportarán bases de datos u otros almacenamientos.

Este modelo suele ser elaborado por analistas de datos o arquitectos de datos, y sirve para revisar con mayor detalle las estructuras antes de comprometernos con una plataforma tecnológica.

Aunque sigue siendo independiente del motor de base de datos, el modelo lógico ya puede incluir decisiones como la división en tablas, la identificación de dominios de datos (ejemplo listas de valores válidos), reglas de negocio (ejemplo “un ciudadano puede tener múltiples solicitudes”) y restricciones lógicas. Este nivel asegura que el diseño cumpla requisitos de negocio y minimizará problemas en la siguiente fase física.

Actividades para generar el modelo.

- Refinar atributos y claves.
- Definir tipos de datos.
- Detallar relaciones.
- Crear un borrador de Diagrama Entidad-Relación (ERD)

7.3. Modelo físico.

La materialización del modelo lógico se logra a través del modelo de datos físico, mediante un sistema de gestión de bases de datos (DBMS) u otro medio de almacenamiento. Se Concentra en el “**cómo**” de la implementación en un entorno tecnológico específico.

El objetivo principal del modelo físico es optimizar el almacenamiento y acceso a la información, incorporando consideraciones de rendimiento, seguridad y capacidad.

En resumen, es la traducción del modelo lógico al lenguaje propio de la base de datos o sistema destino. Este modelo suele plasmarse en forma de diagramas físicos o directamente en scripts DDL (lenguaje de definición de datos). Los administradores de bases de datos (DBAs) y desarrolladores usarán el modelo físico para crear las bases de datos reales. Es importante que el modelo físico preserve la fidelidad con el modelo lógico, a la vez que aproveche las características del gestor de datos seleccionado.

Se definen en detalle las tablas, columnas y sus tipos de datos concretos, longitudes, índices, particiones, esquemas, vistas, procedimientos almacenados, y cualquier otro aspecto dependiente de la plataforma elegida (por ejemplo, Oracle, SQL Server, PostgreSQL, etc.). También se establecen las claves primarias y foráneas con su sintaxis específica, índices para optimización, restricciones de integridad referencial, etc.

Actividades para generar el modelo.

- Traducir a tablas y columnas.
- Definir índices y restricciones.
- Diseñar el esquema físico.

7.4. Modelo Ontológico.

El modelo ontológico (u ontología de datos) es una representación formal de la semántica de la información: describe los conceptos, categorías o clases y las relaciones entre ellos en un dominio específico, de manera que tanto humanos como sistemas puedan interpretarlos de forma unificada. A diferencia de los modelos anteriores (que suelen enfocarse en estructura de bases de datos), una ontología utiliza lenguajes formales (como RDF/OWL del W3C) para definir vocabularios comunes, taxonomías y reglas lógicas.

El objetivo principal del modelo ontológico es proporciona una capa semántica que enriquece la comprensión de los datos más allá de su estructura física, habilitando un vocabulario común y la interoperabilidad semántica en el ecosistema de información público. Las ontologías

permiten establecer un significado común para los datos, evitando confusiones terminológicas y ambigüedades, y facilitando la interoperabilidad entre sistemas.

Si bien no todas las organizaciones desarrollan ontologías completas, adoptar principios ontológicos aporta beneficios: mejor comunicación semántica entre áreas y sistemas, reusabilidad de definiciones y alineación con estándares internacionales (por ejemplo, ontologías de Gobierno Abierto, vocabularios de datos públicos, etc.).

Construir una ontología implica identificar las clases (conceptos) del dominio, sus propiedades (atributos) y relaciones (jerarquías es un, asociaciones tiene, etc.), así como instancias si se requiere. En el contexto gubernamental, un modelo ontológico puede ayudar a integrar datos de múltiples fuentes o entidades: por ejemplo, definiciones estándar para “Ciudadano”, “Empresa” o “Proyecto” que varias agencias comparten. También son útiles para datos abiertos y gobierno digital, ya que permiten enlazar datos (linked data) y habilitar razonamiento automático (ejemplo inferir relaciones no directas a partir de reglas).

Herramientas como Protégé permiten crear y gestionar ontologías de forma amigable. Vale destacar que, en el MRAE y la literatura de arquitectura de datos, un modelo ontológico puede considerarse complementario al modelo conceptual – ambos buscan modelar la realidad del negocio –, pero la ontología lo hace con más formalismo matemático/lógico.

Actividades para generar el modelo.

- Identificar los elementos (conceptos, atributos y relaciones)
- Traducir a un lenguaje formal.
- Derivar nuevos conceptos

7.5. Modelo metadatos.

Los metadatos son, en términos simples, “datos sobre los datos”. Es decir, son elementos descriptivos que proveen contexto y detalles acerca de un conjunto de datos o un recurso informativo. Los metadatos permiten entender, ubicar y usar correctamente la información, ya que responden preguntas como *¿qué es este dato? ¿de dónde proviene? ¿qué significa? ¿con qué frecuencia se actualiza?*

El objetivo de los metadatos es mejorar la calidad y gobierno de los datos, facilitando su descubrimiento, evitando duplicidades y asegurando su correcta interpretación a lo largo del tiempo.

Desde el punto de vista de arquitectura de datos, la gestión de metadatos implica mantener un Catálogo de Información institucional o Diccionario de Datos, donde se documenta cada

elemento de datos manejado por la entidad: su nombre, definición, formato, origen, dueño, relaciones con otros datos, clasificación de seguridad (público, reservado, confidencial) y también su clasificación documental si aplica. En entornos de datos abiertos, el uso de perfiles de metadatos estandarizados (por ejemplo, DCAT-AP en la Unión Europea o perfiles locales) es imprescindible para que terceros puedan descubrir y reutilizar los datos fácilmente. En el caso de entidades públicas de Colombia, se debe alimentar el catálogo de datos abiertos (datos.gov.co) con metadatos exhaustivos de cada conjunto publicado, siguiendo las guías nacionales.

Dado que en entidades públicas la información documental se rige por tablas de retención documental (TRD), es buena práctica incluir en los metadatos de cada tipo de información su código o serie documental correspondiente y el tiempo de retención, para así enlazar la gestión de datos con la gestión documental.

Por ejemplo, para un conjunto de datos de ciudadanos atendidos en una ventanilla única, sus metadatos incluirían: título (ejemplo “Atenciones por Ventanilla Única 2025”), descripción del contenido, fecha de última actualización, responsable o custodio de la información, formato, campos que contiene con sus definiciones, nivel de confidencialidad, entre otros. Un título y una breve descripción son metadatos mínimos que todo conjunto de datos debe tener, pero existen muchos otros según buenas prácticas (palabras clave, categorías temáticas, georreferencia, normas de calidad, etc.).

Herramientas de *data catalog* o repositorios de metadatos (soluciones como CKAN para datos abiertos o plataformas comerciales tipo Collibra) pueden emplearse para centralizar esta información.

Actividades para generar el modelo.

- Identificar los elementos
- Definir la semántica de los elementos
- Definir el formato de cada elemento
- Identificar las relaciones entre los elementos

8. Herramientas que apoyan los modelos.

Existen diversas herramientas que apoyan la construcción de modelos conceptuales, lógicos, físicos, ontológico y la gestión de metadatos. La elección dependerá de los recursos, la profundidad requerida y preferencias de la Entidad. Entre las herramientas más populares. No es necesario comprar el software más costoso para empezar a organizar sus datos. La mejor

herramienta es la que su equipo sepa usar y le permita comunicarse claramente. Hemos clasificado las herramientas disponibles según el objetivo que usted tenga.

1. Para "Dibujar y Comunicar" (Nivel Estratégico y de Negocio)

Ideal para: Líderes de procesos, gerentes y primeras versiones del Modelo Conceptual.

Si su objetivo es hacer un diagrama rápido para que todos entiendan qué es un 'Ciudadano' o un 'Trámite', no necesita software complejo. Busque herramientas que funcionen como una pizarra digital.

- **Opciones Gratuitas o de Bajo Costo:**

- **Draw.io / Diagrams.net:** Gratuita y funciona en el navegador. Es excelente para bocetos rápidos.
- **Excel o PowerPoint:** Aunque no son herramientas de modelado, son perfectas para iniciar diccionarios de datos o diagramas simples que todos pueden abrir.

- **Opciones de Oficina:**

- **Microsoft Visio:** El estándar clásico para diagramas en muchas oficinas. Útil para mapas visuales, aunque no valida reglas técnicas.
- **Lucidchart:** Permite que varias personas dibujen al mismo tiempo a través de internet, ideal para trabajo colaborativo.

2. Para "Diseñar y Construir" (Nivel Técnico e Ingeniería)

Ideal para: Ingenieros de datos, Arquitectos de TI y Administradores de Bases de Datos (DBA).

Cuando ya vamos a crear la base de datos real (Modelo Físico) o necesitamos asegurar reglas estrictas (Modelo Lógico), necesitamos herramientas de "ingeniería" que puedan generar código automáticamente.

- **Opciones Robustas:**

- **Erwin Data Modeler / SAP PowerDesigner:** Son las herramientas "profesionales" de la industria. Permiten manejar modelos gigantes y conectar con muchos tipos de bases de datos.

- **Oracle SQL Developer Data Modeler:** Una opción potente y gratuita si su entidad trabaja con tecnologías Oracle.
- **DBeaver / MySQL Workbench / pgModeler:** Herramientas técnicas específicas para motores de base de datos populares y de uso libre.

3. Para "Organizar el Vocabulario" (Semántica y Metadatos)

Ideal para: Arquitectos de Información, gestores de calidad y equipos de Datos Abiertos.

Si su objetivo es crear el "diccionario" de la entidad (Ontologías) o catalogar qué datos tienen (Metadatos), use herramientas que permitan clasificar y etiquetar.

- **Gestión de Significados (Ontologías):**

- **Protégé:** Es el estándar gratuito para crear ontologías. Funciona como un editor avanzado de vocabularios.

- **Catálogos de Datos:**

- **CKAN:** La herramienta estándar mundial para publicar datos abiertos (usada en datos.gov.co).
- **Excel (Nuevamente):** Para muchas entidades, una hoja de cálculo bien organizada es el mejor punto de partida para su Diccionario de Datos.

4. Para "Ver Todo el Panorama" (Arquitectura Empresarial)

Ideal para: Arquitectos Empresariales.

Si necesita conectar los datos con los procesos y las aplicaciones de la entidad.

- **Archi (con ArchiMate):** Gratuita y de código abierto. Permite ver cómo un dato afecta a un trámite y a un software.
- **Sparx Enterprise Architect:** Una herramienta completa para gestionar toda la arquitectura de la entidad.

Existen otras herramientas y soluciones: **pgModeler** (especializada en PostgreSQL), **MySQL Workbench**, **Toad Data Modeler**, entre otras., para la gestión de metadatos y catálogo de datos, se pueden evaluar plataformas como **CKAN** (orientada a datos abiertos), **DataHub**, **Apache Atlas**, o soluciones comerciales tipo **Collibra** y

Informática EDC que facilitan el gobierno de datos a gran escala. Lo importante es que la herramienta elegida permita documentar adecuadamente los modelos y compartirlos con los interesados.

Si no se cuenta con herramientas sofisticadas, incluso plantillas en Excel para diccionarios de datos y diagramas en Visio/Draw.io pueden servir como punto de partida, aunque a largo plazo es recomendable consolidar un repositorio central de arquitectura como sugiere el MRAE.

9. Buenas prácticas para el modelado.

Para asegurar la calidad y utilidad de los modelos de datos, las Entidades deben tener en cuenta una serie de buenas prácticas, reconocidas tanto en el MRAE como en estándares internacionales:

- **Alineación con los objetivos de negocio.**

Todo modelado debe partir de las necesidades de la organización. Identifique la información clave para los procesos misionales, servicios al ciudadano y decisiones estratégicas, y priorice modelar esos dominios inicialmente. Un modelo de datos es valioso en la medida que refleja la realidad del negocio y ayuda a resolver sus problemas.

- **Enfoque iterativo y evolutivo.**

No intente modelar *todo* de una sola vez. Construya el modelo de información de forma iterativa, agregando áreas temáticas o nivel de detalle en sucesivas fases. Por ejemplo, en un primer ciclo modele las entidades relacionadas con un proceso crítico; en iteraciones posteriores amplíe a otros procesos o detalle más atributos. Este enfoque incremental coincide con la recomendación del MRAE de desarrollar el Modelo de Información Institucional por iteraciones sucesivas, permitiendo aprender y ajustar sobre la marcha.

- **Participación de las partes interesadas.**

Involucre a los dueños de proceso, expertos de negocio y usuarios clave en la definición del modelo conceptual. Validar tempranamente con los usuarios que las entidades y relaciones reflejan correctamente la realidad evita retrabajos. Un modelo conceptual debe ser comprensible para personal no técnico; use el lenguaje del negocio (términos que ellos usan) en los nombres de entidades y atributos. Las sesiones de modelado

colaborativo (talleres de blancos) son muy útiles para lograr consenso en las definiciones de datos.

- **Documentación exhaustiva y centralizada.**

Mantenga un catálogo o diccionario de datos institucional donde se registre cada entidad, atributo y relación con su definición, responsable, origen y otros metadatos relevantes. Este catálogo debe actuar como la “única fuente” sobre los datos para toda la entidad, evitando duplicidad de definiciones o versiones no alineadas. Cada vez que se cree o modifique un modelo (sea conceptual, lógico o físico), se requiere actualizar el catálogo y viceversa. La trazabilidad es crucial: por ejemplo, si “*Ciudadano*” es una entidad en el modelo conceptual, en el catálogo debe estar su descripción y si aparece en múltiples sistemas, estas relaciones deben ser conocidos.

- **Estándares de nomenclatura y modelado.**

Adopte convenciones consistentes para nombrar entidades, atributos y otros objetos. Por ejemplo, decidir si los nombres serán singulares o plurales, en español o inglés, uso de prefijos/sufijos, etc., y aplicarlos uniformemente. Asimismo, se requiere utilizar notaciones estándar para los diagramas (ER clásico, UML, etc.) según convenga, de forma que otros arquitectos/analistas puedan interpretarlos fácilmente. Siempre que sea posible, aplique estándares comunes de datos ya existentes: por ejemplo, el Lenguaje Común de Intercambio de Información definido por MinTIC para datos compartidos entre entidades, que provee significado y estructura unificada para los datos en el gobierno. Esto asegurará que sus modelos sean interoperables con otras entidades y entendidos en un contexto más amplio.

- **Calidad y normalización de los datos.**

Durante el modelado lógico, busque eliminar redundancias e inconsistencias mediante técnicas de normalización (1FN, 2FN, 3FN, etc.), a menos que existan razones de rendimiento para no normalizar. Cada entidad debe representar un concepto único y cada atributo un hecho atómico sobre esa entidad. Verifique la calidad del modelo realizando revisiones: por ejemplo, comprobar que cada entidad tenga una clave primaria apropiada y que las relaciones muchos-a-muchos se resuelvan con entidades intermedias, evitando estructuras difíciles de manejar. Asegúrese también de capturar reglas de negocio (constraints) en el modelo; ejemplo, si un campo es obligatorio, si cierto valor tiene un formato específico, etc., idealmente como parte del diccionario de datos.

- **Consideraciones de seguridad, privacidad y cumplimiento.**

Incorpore desde el diseño metadatos de clasificación de la información. Identifique qué datos son personales, sensibles o reservados por ley, y márcalos en el modelo. Esto ayudará a implementar controles de acceso luego. Asimismo, se debe reflejar en el modelo las políticas de retención: por ejemplo, indicar qué tablas o entidades contienen datos sujetos a la TRD y por cuánto tiempo deben conservarse. Asegurar el cumplimiento normativo (Ley de Protección de Datos Personales) es parte del buen gobierno de datos. También, planifique la calidad de datos incorporando indicadores o requisitos de integridad.

- **Mapeo entre datos, procesos y sistemas.**

Un modelo de datos no vive aislado; debe estar conectado con la arquitectura de procesos y de aplicaciones. Una buena práctica de arquitectura empresarial es construir *matrices de trazabilidad* o “mapas de información” que vinculen qué procesos generan o usan qué entidades de datos, y en qué sistemas o bases de datos residen. Por ejemplo, saber que la entidad “*QuejaCiudadana*” se crea en el proceso de “Gestión de PQRS” y se almacena en el Sistema X. Esto permite identificar redundancias y oportunidades de integración, pensar en compartir datos en lugar de re-capturarlos. El MRAE promueve esta visión integrada a través de su dominio de Arquitectura de Información e Interoperabilidad, fomentando el uso de servicios para compartir datos entre entidades.

- **Mantenimiento y gobierno continuo.**

Establezca procedimientos para que los modelos de datos se mantengan actualizados con la evolución de la Entidad. Cada nuevo proyecto de sistema o cambio importante en procesos debe gatillar una revisión del modelo conceptual/lógico para ver si surgen nuevas entidades o atributos.

El MRAE exige que la Arquitectura de Información se mantenga actualizada y habilite la generación de información de valor, lo cual implica actualizar el Modelo de Información Institucional periódicamente. Defina roles de gobierno de datos que aseguren que los nuevos datos se documentan correctamente y que los obsoletos se eliminan o archivan según corresponda.

- **Comunicación y difusión.**

Una vez elaborados, haga que los modelos sean fácilmente accesibles para quienes los necesiten. Puede publicar el modelo conceptual o partes de él, en forma de diagramas

entendibles para alto nivel y el modelo lógico en repositorios técnicos para desarrolladores.

Capacite a los equipos en la lectura e importancia del modelo de datos. Una cultura de datos se refuerza cuando todos comprenden el mapa de la información de la entidad y saben dónde acudir para definir un término o verificar la fuente oficial de un dato.

10. Estándar de industria para los modelos.

Se recomienda alinear el trabajo no solo con el MRAE local sino también considerar estándares internacionales de arquitectura y gestión de datos, para incorporar las mejores prácticas globales:

- **Marco de Referencia de Arquitectura Empresarial (MRAE) – MinTIC.**

Es el marco local (versión 3.0 a 2023) que guía a las entidades públicas colombianas en la adopción de Arquitectura Empresarial. Dentro del MRAE, el Dominio de Arquitectura de Información establece lineamientos para modelar, describir y mantener la arquitectura de datos institucional. Por ejemplo, exige contar con un Modelo de Información Institucional validado con los interesados, que brinde una vista común de la información como base para todos los proyectos de datos. También promueve productos como el Catálogo de Información y diagramas de flujos de información. Seguir el MRAE implica entregar los artefactos sugeridos (diagramas, catálogos, documentos) y asegurar que la arquitectura de datos soporte la misión institucional. Esta guía está alineada con esos lineamientos: por ejemplo, la idea de que el modelo de datos institucional condensa los modelos conceptuales, lógicos y físicos de la entidad, garantizando consistencia desde la visión de alto nivel hasta la implementación.

- **TOGAF (The Open Group Architecture Framework).**

Marco de arquitectura empresarial reconocido internacionalmente. TOGAF define la arquitectura de datos como uno de sus dominios fundamentales dentro de la Arquitectura de Información, y recomienda desarrollar arquitecturas en niveles Conceptual, Lógico y Físico de forma secuencial. La metodología ADM de TOGAF sugiere abordar primero una arquitectura de datos de alto nivel (conceptual) durante las fases de arquitectura, antes de diseñar la solución concreta. Alinearse con TOGAF asegura que el modelado de datos esté integrado con la visión global de arquitectura empresarial (junto con arquitectura de aplicaciones, tecnología, etc.).

- **Marco Zachman.**

Ofrece una matriz de vistas para arquitectura empresarial. En la columna de “Datos”, Zachman identifica distintas representaciones desde la perspectiva contextual del negocio hasta la especificación detallada física. Aunque Zachman es más teórico, subraya la necesidad de tener distintas vistas (ej. Conceptual = “modelo de nivel de Negocio”, Lógico = “modelo de Sistemas”, Físico = “modelo Tecnológico”). Esto respalda la separación de modelos que hemos descrito y refuerza la trazabilidad entre ellos.

- **DAMA-DMBOK (Data Management Body of Knowledge).**

Es un marco de mejores prácticas en gestión de datos desarrollado por DAMA International. Dentro de sus áreas de conocimiento, destaca la Arquitectura de Datos y el Desarrollo de Modelos de Datos como disciplinas clave. DMBOK aconseja documentar claramente los metadatos, establecer gobierno de modelado, y usar modelos de datos para comunicación con el negocio. Seguir estas recomendaciones complementa la perspectiva de MRAE con una visión más amplia de gobierno de datos, incluyendo roles (administradores de datos), procesos de control de cambios a modelos, etc.

- **Estándares de Metadatos y Datos Abiertos.**

En cuanto a metadatos, existen estándares internacionales como **ISO/IEC 11179** (modelo para registros de metadatos) que brindan pautas para definir datos de forma consistente (ej. formato de nombres, definiciones de elementos de datos). Para datos abiertos y catalogación, el estándar DCAT (Data Catalog Vocabulary) del W3C es ampliamente utilizado para describir conjuntos de datos y sus distribuciones, asegurando interoperabilidad entre catálogos. Adoptar estos estándares en la publicación de datos abiertos (por ejemplo, usar DCAT-AP en el portal de datos abiertos) incrementa la facilidad de intercambio y entendimiento por parte de la comunidad internacional. En el caso local, los perfiles de metadatos recomendados (la *Guía para publicación de metadatos* de la Estrategia de Gobierno en Línea) deben observarse para cumplir con la normativa de Datos Abiertos.

- **Lenguaje Común de Intercambio de Información (LCCI).**

Ya mencionado anteriormente, este es un estándar colombiano para asegurar que cuando distintas entidades intercambien datos, utilicen definiciones y estructuras uniformes. Por ejemplo, define estructuras estándar para datos de persona, datos de ubicación geográfica, etc. Asegurarse de que el modelo de datos institucional incorpore los elementos del LCCI cuando corresponda (campos como tipo de identificación,

nomenclatura de direcciones, etc.) es crucial para habilitar la interoperabilidad a nivel país. Esto alinea la arquitectura de datos con el Marco de Interoperabilidad gubernamental, evitando esfuerzos aislados de definición de datos

En conclusión, elaborar modelos de datos robustos implica no **“inventar la rueda”**, sino apoyarse en estos marcos y estándares probados.

La guía aplicada a una entidad pública nacional permite estructurar un Mapa de Información completo: desde los conceptos de datos más abstractos hasta los detalles físicos y semánticos, pasando por buenas prácticas y herramientas concretas para lograrlo. Siguiendo este enfoque alineado con el MRAE y las mejores prácticas internacionales, la entidad podrá gestionar sus datos como activos estratégicos, mejorar la calidad de la información que ofrece a ciudadanos y grupos de interés, y facilitar tanto la interoperabilidad con otras instituciones como la toma de decisiones basada en datos.

11. Ejemplos modelos de datos

El ejemplo de esta guía se centra en la necesidad del negocio de los flujos de información de órdenes de compra hasta la afectación del inventario, pueden ser utilizados para ilustrar la construcción de los modelos de datos y mapas de información dentro de la Arquitectura de Información. La herramienta utilizada para el ejercicio es draw.io.

Caso del negocio.

El proceso inicia con la Orden de Compra, emitida a un Proveedor, la cual puede contener múltiples detalles de los materiales requeridos, definiendo el material como la unidad mínima de información, trazabilidad y control de los datos a lo largo del proceso.

Una vez el proveedor dispone de los materiales, se documenta una Remisión, que da origen al registro de una o varias Entradas de Almacén, las cuales son objeto de validación y aprobación conforme a las reglas del negocio. Una vez aprobadas, se genera el Movimiento de Inventario, encargado de actualizar las existencias del material. Este conjunto de actividades permite evidenciar la transformación, uso y relación de la información, sirviendo como insumo para la definición de los modelos.

11.1. Modelo Conceptual

El modelo conceptual presenta una visión de alto nivel del proceso de órdenes de compra e inventarios, identificando las principales entidades del negocio y sus relaciones, como

proveedor, orden de compra, detalle, material, remisión, entrada de almacén y movimiento de inventario, sin considerar aspectos técnicos.

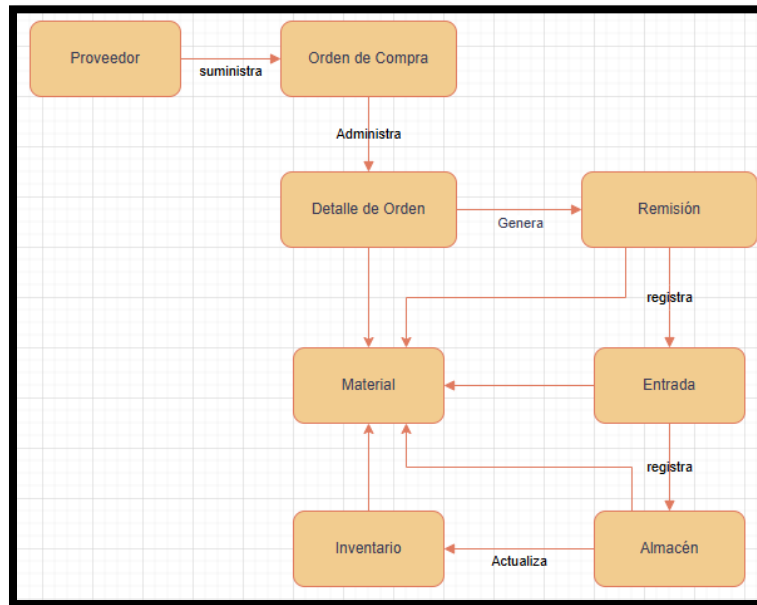


Ilustración 1. Modelo Conceptual

11.2. Modelo lógico

El modelo lógico detalla la estructura de la información del proceso de órdenes de compra e inventarios, definiendo entidades, atributos, claves y relaciones, y representando que una orden de compra puede incluir múltiples detalles asociados a materiales, unidad mínima de control del inventario.

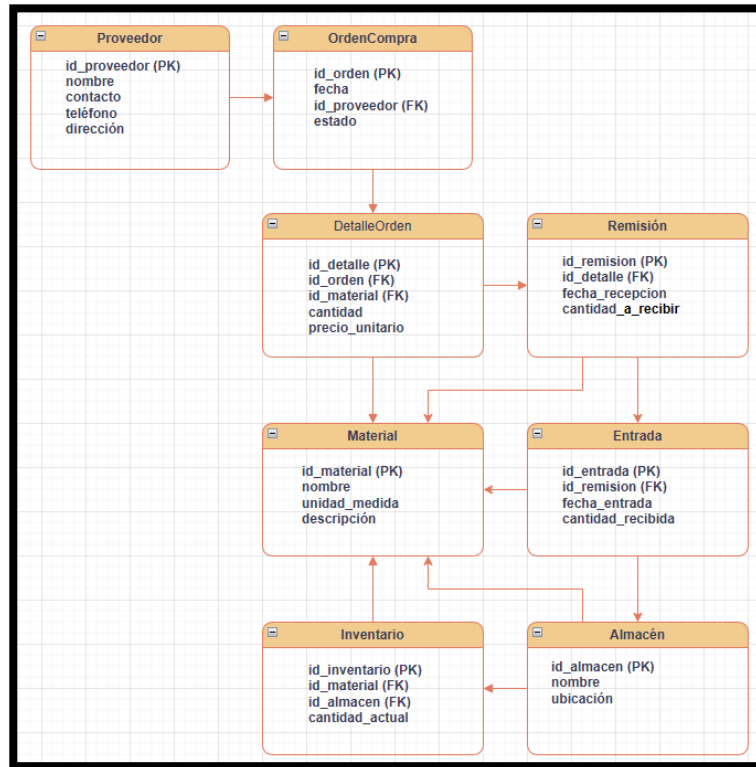


Ilustración 2. Modelo lógico

11.3. Modelo físico

El modelo físico implementa el modelo lógico en una plataforma tecnológica específica, definiendo tablas, tipos de datos, índices y restricciones que permiten gestionar de forma eficiente y segura la información de órdenes de compra, materiales, entradas de almacén y movimientos de inventario.

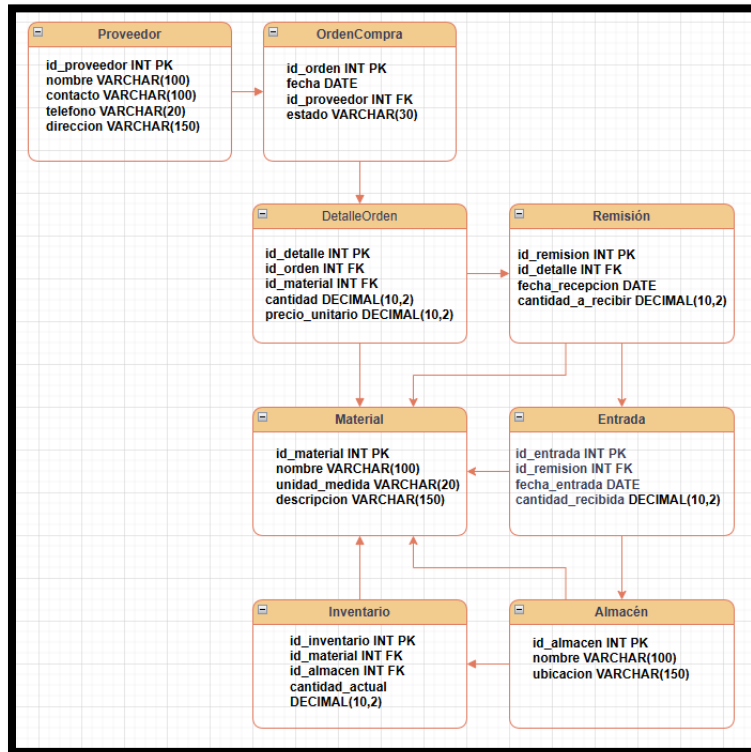


Ilustración 3. Modelo físico

11.4. Modelo ontológico

El modelo ontológico define el significado común de la información tanto para interpretación humana como de máquina, en este caso del proceso de órdenes de compra e inventarios, estableciendo relaciones semánticas entre proveedor, orden de compra, detalle, material, remisión, entrada de almacén y movimiento de inventario, donde el material constituye la unidad mínima de control.

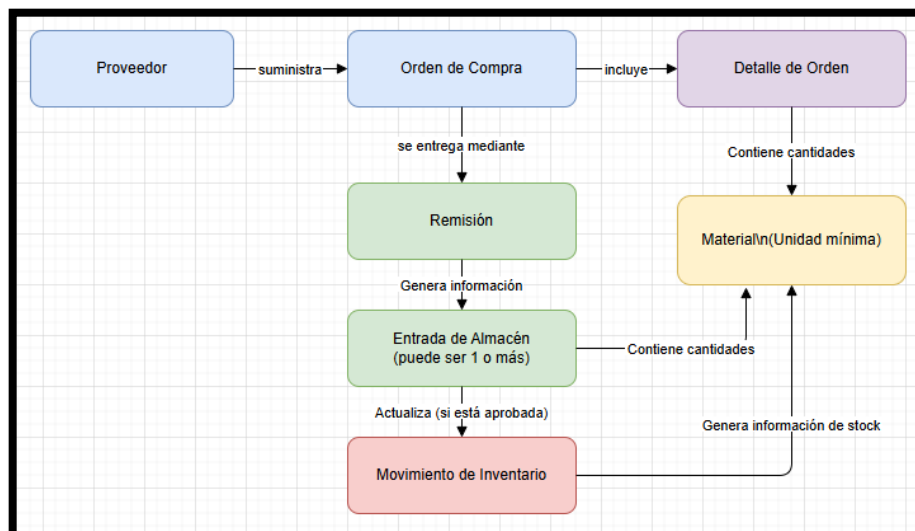


Ilustración 4. Modelo ontológico

11.5. Modelo metadatos

El modelo de metadatos describe las características, origen, responsables y reglas de uso de la información del proceso de órdenes de compra e inventarios, permitiendo la gestión, trazabilidad y calidad de los datos asociados a materiales, órdenes de compra, remisiones, entradas de almacén y movimientos de inventario.

Campo	Descripción / Valor Ejemplo
Identificador del sistema	Codificación interna del metadato
Nombre	Nombre formal del elemento, corresponde al título del elemento de metadato Sintaxis: NombreElemento Ejemplo: Fecha
Título	Nombre o con el cual se identifica un elemento de metadato debe ser legible por un humano. Texto Mayúscula sostenida Ejemplo: FECHA INICIO
Subelemento de	Referencia por su Nombre el elemento de metadato del que depende y sin el cual no puede utilizarse
Subelementos	Referencia por Nombre los elemento de metadato que contiene (puede ser uno o varios)
Descripción	Describe breve el contenido del metadato
Aplicabilidad	Describe la entidad para la cual es aplicable el metadato, es decir, de acuerdo con el modelo definido, si aplica para un documento, para el agente o productor o para las relaciones
Obligatoriedad	Especifica si el metadato es Obligatorio, Condicional, Opcional
Repetibilidad	Especifica si el metadato es o no repetible. (si/no)
Tipo de dato	Este elemento sirve para describir el tipo de dato a utilizar para la captura de información (char, int, boolean o string entre otros)
Modo captura	Describe la forma como el metadato debe ser capturado. (Automático/Manual)
Valores	Sintaxis de acuerdo con esquemas y fuente autorizada
Finalidad	Propósito para lograr con el uso
Indicador para conservar o destruir	Sirve para indicar si el elemento debe eliminarse cuando se destruye la entidad a la que pertenece
Notas de Alcance	Información de orientación para la comprensión y uso
Ejemplos	Ejemplos de uso

Tabla 1. Modelo metadatos – Elementos descriptivos

Identificador del sistema	Nombre	Título	Subelemento de	Subelementos	Descripción	Aplicabilidad
INV-MAT-001	id_material	Identificador de Material	Material	Inventario	Identificador único del material (unidad mínima)	Inventarios
INV-MAT-002	nombre_material	Nombre del Material	Material	DetalleOrden	Nombre comercial del material	Compras - Inventarios
INV-OC-001	id_orden_compra	ID Orden de Compra	OrdenCompra	DetalleOrden	Identificador de la orden de compra	Compras
INV-OC-002	fecha_orden	Fecha Orden	OrdenCompra		Fecha de creación de la OC	Compras
INV-DET-001	id_detalle	ID Detalle	DetalleOrden	Remisión	Identificador del detalle	Compras
INV-DET-002	cantidad_solicitada	Cantidad Solicitada	DetalleOrden	Remisión	Cantidad solicitada del material	Compras
INV-REM-001	id_remision	ID Remisión	Remisión	Entrada	Documento de entrega proveedor	Recepción
INV-REM-002	fecha_recepcion	Fecha Recepción	Remisión	Entrada	Fecha recepción material	Recepción
INV-ENT-001	id_entrada	ID Entrada	Entrada	Movimiento	Identificador entrada almacén	Almacén
INV-ENT-002	cantidad_recibida	Cantidad Recibida	Entrada	Movimiento	Cantidad recibida aprobada	Almacén
INV-MOV-001	id_movimiento	ID Movimiento	Movimiento	Inventario	Identificador movimiento inventario	Inventarios
INV-MOV-002	tipo_movimiento	Tipo Movimiento	Movimiento	Inventario	Entrada o Salida	Inventarios

Tabla 2. Modelo metadatos – Ejemplo parte 1

Obligatoriedad	Repetibilidad	Tipo de dato	Modo captura	Valores	Finalidad	Indicador para conservar o destruir	Notas de Alcance	Ejemplos
Sí	No	Entero	Maestro	Catálogo	Identificar material	Conservar	Unidad mínima de control	MAT-001
Sí	No	Texto	Maestro	Libre	Describir material	Conservar	Uso estándar	Cemento gris
Sí	No	Entero	Automático	Secuencial	Trazabilidad OC	Conservar	Generado por sistema	OC-1001
Sí	No	Fecha	Automático	YYYY-MM-DD	Control temporal	Conservar	Fecha sistema	2025-08-01
Sí	No	Entero	Automático	Secuencial	Relación OC-Material	Conservar	Por cada material	DET-01
Sí	Sí	Decimal	Manual	>0	Control cantidades	Conservar	Según unidad	100
Sí	No	Texto	Manual	Documento	Trazabilidad entrega	Conservar	Documento externo	REM-500
Sí	No	Fecha	Manual	YYYY-MM-DD	Control recepción	Conservar	Registro almacén	2025-08-03
Sí	No	Entero	Automático	Secuencial	Control entradas	Conservar	Requiere aprobación	ENT-200
Sí	Sí	Decimal	Manual	>=0	Actualizar inventario	Conservar	Base del movimiento	98
Sí	No	Entero	Automático	Secuencial	Trazabilidad inventario	Conservar	Generado tras aprobación	MOV-900
Sí	No	Texto	Automático	Entrada - Salida	Clasificar movimiento	Conservar	Regla del proceso	Entrada

Tabla 3. Modelo metadatos – Ejemplo parte 2

12. Referencias.

- Marco de Referencia de Arquitectura Empresarial (MRAE) – MinTIC - <https://gobiernodigital.mintic.gov.co/portal/Iniciativas/Marco-de-Arquitectura-Empresarial/>
- Guía de Arquitectura de Información del MINTIC - https://www.mintic.gov.co/arquitecturaempresarial/630/articles-237650_recurso_1.pdf
- Lenguaje Común de Intercambio de Información (LCCI) - <https://lenguaje.mintic.gov.co>
- TOGAF (The Open Group Architecture Framework) - <https://www.opengroup.org/togaf>
- Estándares de Metadatos y Datos Abiertos - ISO/IEC 11179
- DAMA-DMBOK (Data Management Body of Knowledge) - <https://dama-org.translate.goog/learning-resources/dama-data-management-body-of-knowledge-dmbok/? x tr sl=en& x tr tl=es& x tr hl=es& x tr pto=tc>
- Material divulgativo sobre modelado de datos, referencias sobre ontologías en gobierno electrónico, guías de metadatos para datos abiertos, listados de herramientas líderes en modelado.