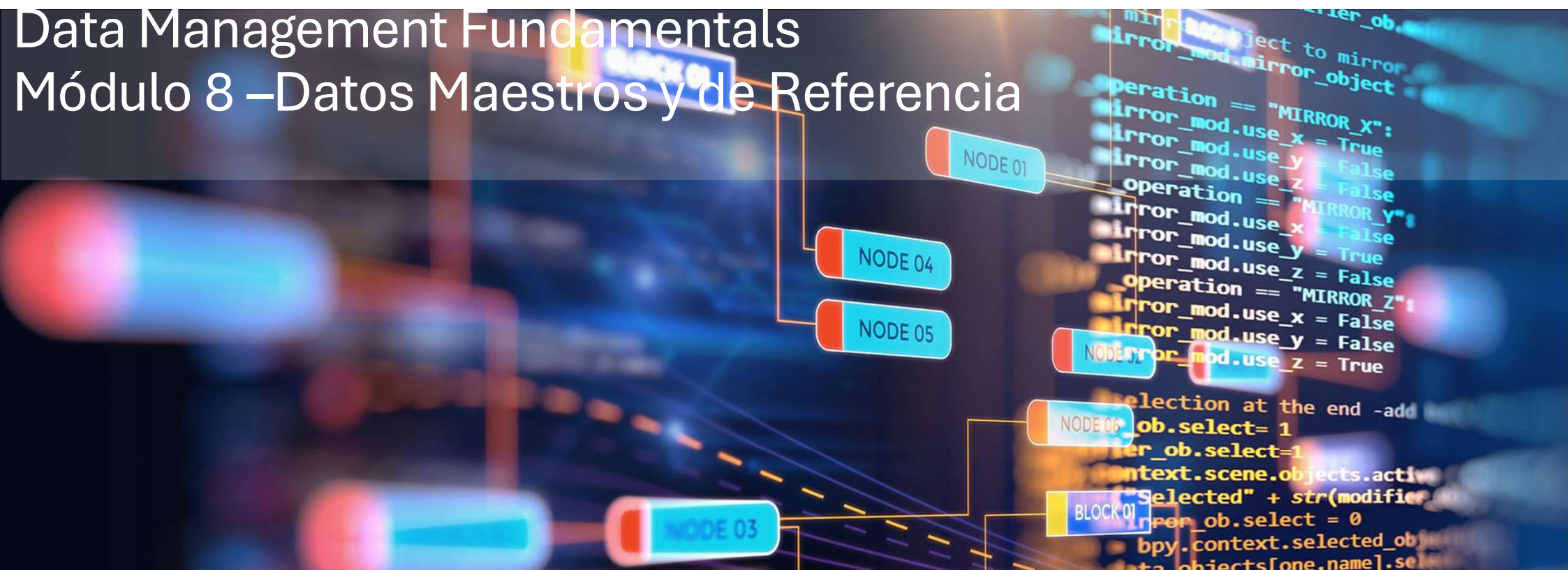


Data Management Fundamentals

Módulo 8 – Datos Maestros y de Referencia



GOLD SPONSORS



SILVER SPONSORS



Entidades académicas



Datos Maestros y de Referencia

- 1. Introducción**
- 2. Actividades**
- 3. Herramientas y Técnicas**
- 4. Guías de Implementación**
- 5. Organización y Cambio Cultural**
- 6. Gobierno de Datos Maestros y de Referencia**

Datos Maestros y de Referencia

Definición:

DATOS MAESTROS: son los datos fundamentales y permanentes que describen las entidades más importantes de una organización.

DATOS DE REFERENCIA: Valores estandarizados que clasifican o validan otros datos.



Datos Maestros y de Referencia



1. Manufactura / Producción

DM	Productos terminados, materias primas, máquinas, proveedores, operarios
DR	Unidades de medida (kg, m ²), códigos de calidad, tipos de materiales, códigos de proceso



3. Salud

DM	Pacientes, médicos, medicamentos, equipos médicos
DR	Códigos CIE, tipos de sangre, especialidades médicas, unidades médicas



5. Logística / Cadena de suministro

DM	Centros de distribución, rutas, transportistas, vehículos
DR	Códigos de país, unidades de peso/volumen, estados de envío



2. Retail / Comercio

DM	Catálogo de productos, clientes, proveedores, tiendas físicas
DR	Categorías de producto (ropa, tecnología), tipos de pago, monedas, niveles de impuestos



4. Banca y Finanzas

DM	Clientes, cuentas bancarias, productos financieros, sucursales
DR	Tipos de moneda, tipos de interés, tipos de cuenta

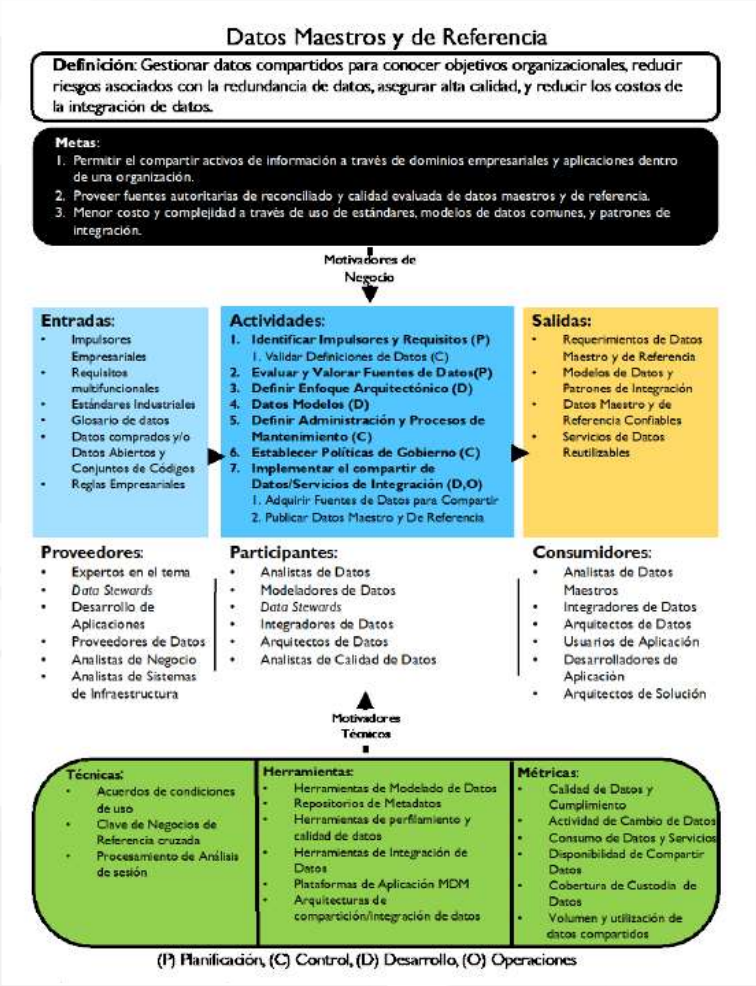


6. Recursos Humanos

DM	Empleados, puestos de trabajo, departamentos
DR	Niveles salariales, tipos de contrato, códigos de país, clasificaciones laborales

Introducción - Diagrama de Contexto

En cualquier organización, se requieren datos a través de sus áreas de negocio, procesos y sistemas. En global la organización y sus clientes se benefician si estos datos son **compartidos**.



Introducción - Motivadores de Negocio

Los principales motivadores para empezar un programa de MDM son:

- 1. Cumplir con los requisitos de datos de la Organización** y garantizar que son completos, actuales y consistentes
- 2. Gestionar la Calidad de los Datos** Inconsistencias y mala Calidad llevan a mala decisiones
- 3. Gestionar los costes de la Integración de Datos** sin MDM integrar nuevas fuentes es mucho más complicado
- 4. Reducir el Riesgo.** Simplifica la arquitectura y por lo tanto riesgos integración



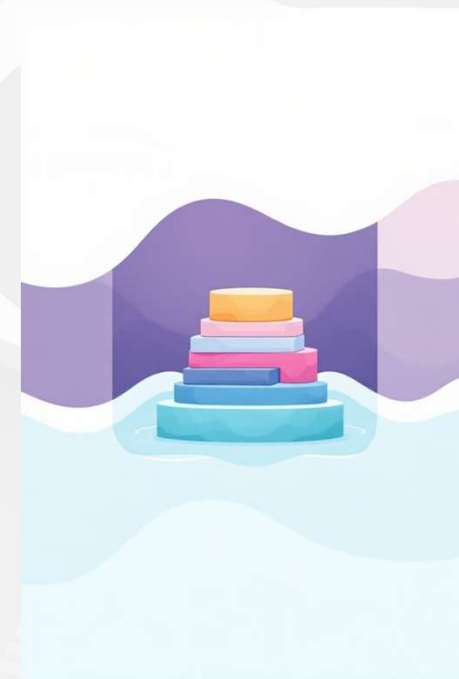
Los principales motivadores para empezar un programa de RDM son similares al MDM.

Mejorar la Disponibilidad y Calidad de los Datos Maestros y de Referencia tiene un impacto en la Calidad general de los Datos y en la confianza del Negocio en los datos

Introducción - Metas y Principios

Las Metas de un programa de Gestión de Datos Maestros y de Referencia son.

1. Asegurar que la Organización cuenta con DM y DR **Completos, Consistentes, Actuales y con autoridad**
2. Permitir que los **DM y DR se compartan** entre todas las unidades de negocio y aplicaciones
3. **Reducir el Costo y la complejidad de integracion** (estándares, datos comunes, patrones)



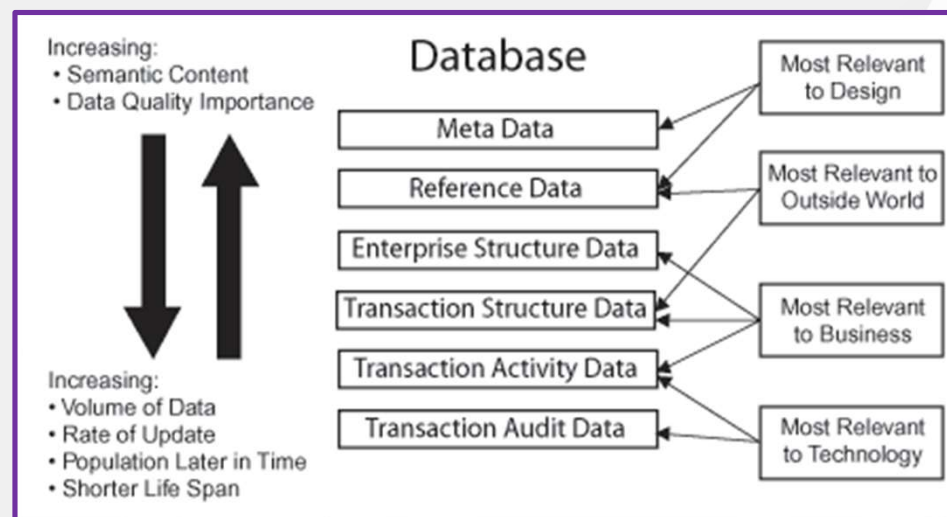
Introducción - Metas y Principios



Introducción - Conceptos Básicos

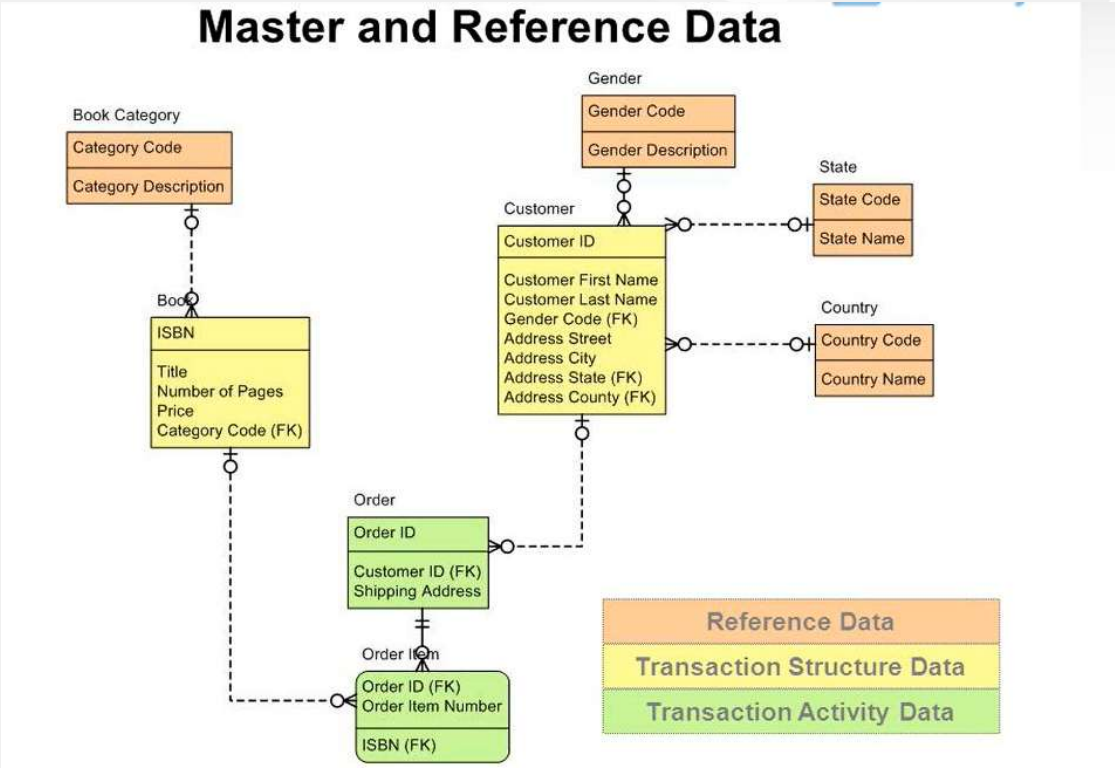
Diferencias entre Datos Maestros y Datos de Referencia

- Diferentes tipos de datos desempeñan diferentes funciones dentro de una organización.
- También tienen diferentes requisitos de gestión.
- Malcolm Chisholm ha propuesto una taxonomía de seis capas de datos:



Introducción - Diagrama de Contexto

Diferencias entre Datos Maestros y Datos de Referencia



Introducción - Conceptos Básicos

Diferencias entre Datos Maestros y Datos de Referencia

La definición de Chisholm es similar a la definición del diccionario de DAMA: Los Datos Maestros son "los datos que proporcionan el contexto para los datos de la actividad del negocio en forma de conceptos comunes y abstractos que se relacionan con la actividad. Incluye los detalles (definiciones e identificadores) de objetos internos y externos involucrados en transacciones del negocio, como clientes, productos, empleados, vendedores y dominios controlados (valores de código) "(DAMA, 2009).

Introducción - Conceptos Básicos

Diferencias entre Datos Maestros y Datos de Referencia

Los Datos Maestros y de Referencia comparten objetivos conceptualmente similares. Ambos proporcionan un **contexto crítico** para la creación y el uso de datos transaccionales

El enfoque de la gestión de datos difiere entre los Datos de Referencia y los Maestros:

Gestión de Datos Maestros (MDM, Master Data Management) implica el control sobre los valores y los identificadores de los Datos Maestros que permiten el uso consistente, a través de sistemas, de los datos más precisos y oportunos sobre entidades de negocio esenciales. Los objetivos del MDM incluyen garantizar la disponibilidad de valores precisos y actuales al mismo tiempo que se reduce el riesgo asociado con identificadores ambiguos

Gestión de Datos de Referencia (RDM, Reference Data Management) implica el control sobre los valores de dominio definidos y sus definiciones. El objetivo de RDM es asegurar que la organización tenga acceso a un conjunto completo de valores precisos y actuales para cada concepto representado.

Introducción - Diagrama de Contexto

Diferencias entre Datos Maestros y Datos de Referencia

Structure: Differences

Master Data
Many Rows per Table
Many Columns per Table
Few Tables per Database
Changes More Quickly
Highly Shared within Enterprise

Reference Data
Few Rows per Table
Few Columns per Table
Many Tables per Database
Changes More Slowly
Highly Shared within Enterprise and between Enterprises

Introducción - Datos de Referencia

Datos de Referencia

Datos de Referencia son cualquier dato utilizado para caracterizar o clasificar otros datos, o para relacionar datos con información externa a una organización (Chisholm, 2001).

Las técnicas comunes de almacenamiento utilizan:

- **Tablas de códigos** en bases de datos relacionales vinculadas mediante claves externas para mantener integridad
- **Sistemas de Gestión de Datos de Referencia**, reglas de negocio, estado, tablas permitidas, etc.
- **Metadatos de atributos** para especificar valores permitidos.

El objetivo de la Gestión de Datos de Referencia es asegurar que los valores son consistentes y actuales en diferentes funciones y que los datos son accesibles para la organización.

Introducción - Datos de Referencia

Estructura de los Datos de Referencia

Dependiendo de la granularidad y complejidad de lo que representan los Datos de Referencia, pueden estructurarse como

- **Listas**

Valor de Código	Descripción
US	United States of America
GB	United Kingdom (Great Britain)

- **Listas Ampliadas**

Código	Descripción	Definición
1	Nuevo	Indica un <i>ticket</i> creado sin un recurso asignado
2	Asignado	Indica que el <i>ticket</i> que tiene asignado un recurso nombrado
3	Trabajo en proceso	Indica que el recurso asignado inicio el trabajo indicado en el <i>ticket</i> .

- **Listas de Referencias Cruzadas**

Código de estado USPS	Código de estado ISO	Código numérico de estado FIPS	Abreviación del Estado	Nombre de estado	Nombre formal del Estado
CA	US-CA	06	Calif.	California	State of California
KY	US-KY	21	Ky.	Kentucky	Commonwealth of Kentucky
WI	US-WI	55	Wis.	Wisconsin	State of Wisconsin

- **Taxonomías**

Valor de Código	Descripción	Código Padre
10161600	Floral plants	10160000
10161601	Rose plants	10161600
10161602	Poinsettias plants	10161600
10161603	Orchid plants	10161600
10161700	Cut flowers	10160000
10161705	Cut roses	10161700

- **Ontologías**

Uno de los principales casos de uso de las ontologías es la gestión de contenidos

Introducción - Datos de Referencia

Tipos de Datos de Referencia – Dependiendo de granularidad y complejidad

Datos de Referencia Propios o Internos

Muchas organizaciones crean Datos de Referencia para soportar procesos y aplicaciones internas.

Datos Geográficos o Geoestadísticos

La referencia geográfica o geoestadística permite la clasificación o el análisis basado en geografía

Datos de Referencia de la Industria

Conjuntos de datos que son creados y mantenidos por asociaciones de la industria u organismos gubernamentales con el fin de proporcionar un estándar común para la codificación de conceptos importantes.

Datos de Referencia Computacional

Muchas actividades del negocio dependen del acceso a cálculos comunes y consistentes.

Introducción - Datos de Referencia

Metadatos del conjunto de Datos de Referencia estándar

Conjunto de Datos de Referencia Información Clave	Descripción
Nombre Formal	Oficial, en especial si es el nombre externo del conjunto de Datos de Referencia (por ejemplo: ISO 3166-1991 Lista de Códigos de países)
Nombre Interno	Nombre asociado con el conjunto de datos dentro de la organización (por ejemplo: Códigos de países – ISO)
Proveedor de Datos	La parte que provee y mantiene el conjunto de Datos de Referencia. Puede ser externo (ISO), interno (un departamento específico), o externo – extendido (obtenido de una parte externa pero después extendido y modificado internamente).
Fuente del Conjunto de Datos del Proveedor de Datos	Descripción de dónde se pueden obtener los conjuntos de datos del proveedor de datos. Puede ser un <i>Universal Resource Identifier</i> (URI) dentro o fuera de la red de la organización.
Último Número de Versión del Proveedor de Datos	Si está disponible y se mantiene, describe la última versión del conjunto de datos del proveedor de datos externo a partir de la cual la organización puede agregar o rechazar información
Última fecha de Versión del Proveedor de Datos	Si está disponible y se mantiene, describe cuándo fue la última vez que se actualizó una lista estandarizada
Número Interno de Versión	Número de versión del conjunto de Datos de Referencia actual o número de versión de la última actualización que fue aplicada en ese conjunto de datos
Fecha de Reconciliación de la Versión Interna	Fecha en la que el conjunto de datos se actualice basado en una Fuente externa
Fecha de la Última Actualización de la Versión Interna	La fecha en la que el conjunto de datos se cambió por última vez. Esto no implica la reconciliación con una versión externa

Introducción - Datos Maestros

Datos Maestro

- Datos Maestros son **datos sobre las entidades de negocio** que proporcionan contexto para las transacciones de negocio y el análisis
- Los Datos Maestros **deben representar los datos autorizados y más exactos** disponibles sobre las principales entidades del negocio.
- Los Datos Maestros organizacionales comunes incluyen datos sobre:
 - **Partes (*parties*)** compuestas por individuos y organizaciones, y sus roles
 - **Productos y Servicios** tanto internos como externos
 - **Estructuras financieras** cuentas del libro mayor, centros de coste
 - **Ubicaciones** como direcciones

Introducción - Datos Maestros

Sistema de Registro, Sistema de Referencia

Cuando hay versiones potencialmente diferentes de 'la verdad', es necesario distinguirlas

- **Un Sistema de Registro (SOR)** es un sistema con autoridad donde los datos son creados/capturados y/o mantenidos a través de un conjunto definido de reglas y expectativas.
 - Fuente Oficial autorizada
 - Dato original
 - Auditable
 - Integridad
- **Un Sistema de Referencia (SORf)** es un sistema con autoridad donde los consumidores de datos pueden obtener datos confiables para soportar transacciones y análisis, incluso si la información no se originó en el sistema de referencia.
 - Datos Consolidados
 - Datos Armonizados
 - Actuales y Completos
 - Accesible

Introducción - Datos Maestros

Fuente de Confianza, Registro Dorado

- Una **Fuente de Confianza** es reconocida como la "mejor versión de la verdad", vista única, vista de 360°. Dentro de Fuente de confianza, los datos que representan los datos mas precisos pueden ser referidos como los Registros Dorados.
- El término **Registro Dorado** puede ser engañoso. Tech Target define un Registro Dorado como "la única versión de la verdad"



Introducción - Datos Maestros

Gestión de Datos Maestros

Gartner define Gestión de Datos Maestros como "una disciplina habilitada por la tecnología en la que las empresas y TI trabajan juntas para asegurar la uniformidad, exactitud, administración, consistencia semántica y rendición de cuentas de los activos compartidos oficialmente de Datos Maestros de la empresa. Datos Maestros es el conjunto consistente y uniforme de identificadores y atributos ampliados que describen las entidades centrales de la empresa, incluyendo clientes, prospectos, ciudadanos, proveedores, sitios, jerarquías y un mapa de cuentas "

La definición anterior enfatiza que MDM es una disciplina compuesta de PERSONAS, PROCESOS y TECNOLOGÍA. No es una solución de una herramienta específica

Introducción - Datos Maestros

Gestión de Datos Maestros

La **evaluación de los requerimientos MDM** de una organización incluye la identificación de:

- Roles, organizaciones, lugares y cosas a las que se hacen referencia repetidamente
- Datos que se utilizan para describir personas, organizaciones, lugares y cosas
- Definición y estructura de los datos, incluyendo la granularidad de los datos
- Dónde los datos son creados/adquiridos, almacenados, puestos a disposición y accedidos
- Cambios en los datos a medida que se mueve a través de los sistemas dentro de la organización
- Quién utiliza los datos y para qué fines
- Criterios que se utilizan para comprender la calidad y fiabilidad de los datos y sus fuentes

Introducción - Datos Maestros

Gestión de Datos Maestros

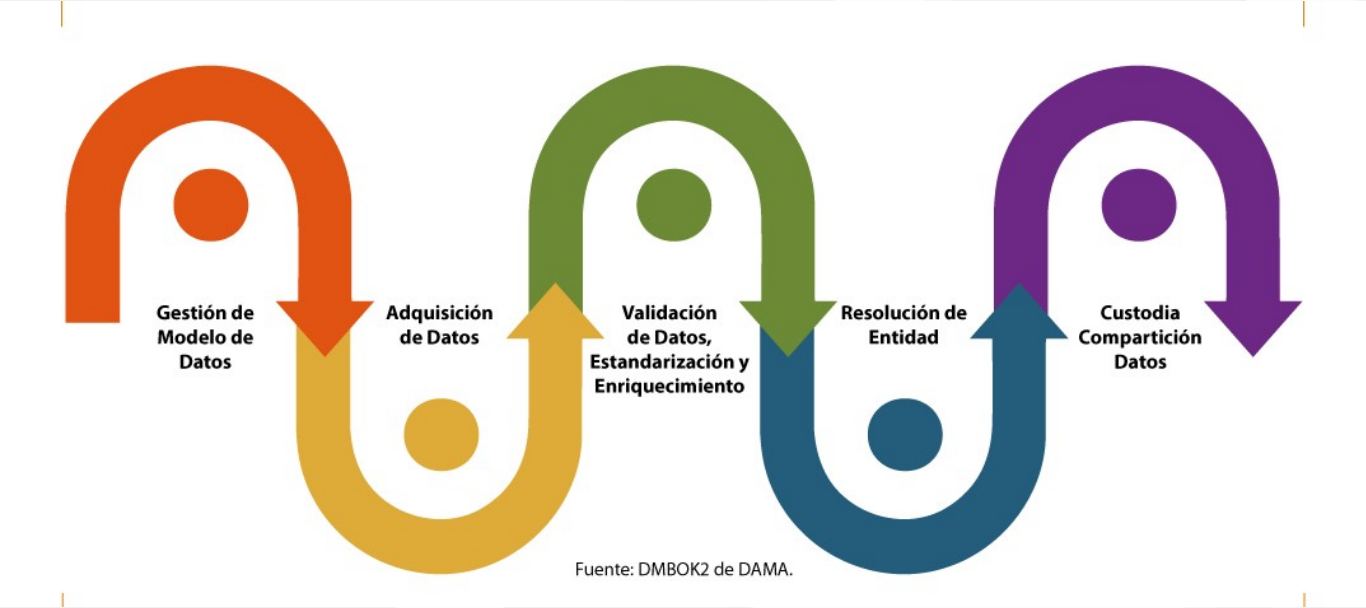
La **planificación de la Gestión de Datos Maestros incluye varios pasos básicos**. Dentro de un dominio:

- Identificar las fuentes candidatas que proporcionarán una visión integral de las entidades de los Datos Maestros
- Desarrollar reglas para llevar a cabo con exactitud la correspondencia y fusión de instancias de entidades
- Establecer un enfoque para identificar y restaurar datos inadecuadamente correspondidos y fusionados
- Establecer un enfoque para distribuir datos confiables a los sistemas de toda la empresa

Ejecutar el proceso, sin embargo, no es tan simple como implican estos pasos, ya que MDM es un proceso de gestión del ciclo de vida.

Introducción - Datos Maestros

Pasos Clave de Proceso para la Gestión de Datos Maestros



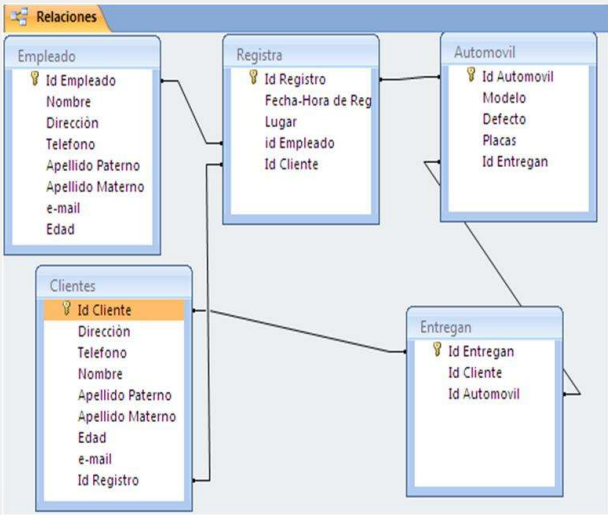
Introducción - Datos Maestros

Gestión del Modelo de Datos



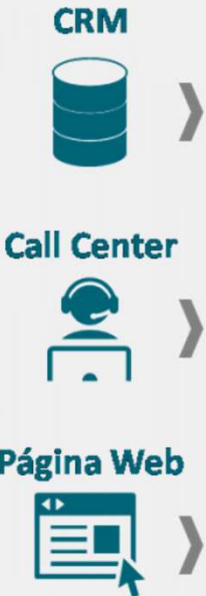
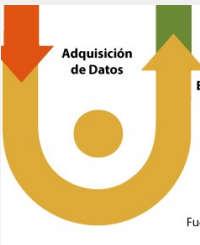
El modelo debe ayudar a la organización a superar la premisa de '**sistema manda**'.

Para los Datos Maestros, los términos y las definiciones utilizados en nivel de empresa deben estar en el contexto de la actividad llevada a cabo en toda la organización y no necesariamente depender del sistema fuente que aporte valores de datos.



Introducción - Datos Maestros

Adquisición de Datos



Nombre	DNI	Dirección	Email
Frodo Bolsón	49886351D	La Comarca, Bolsón Cerrado	frodo.2968@mail.com

Nombre	DNI	Dirección	Email
Frodho Bolson	-	La Comarca	fbolsón@mail.com

Nombre	DNI	Dirección	Email
Frodo B	49886351D	La Comarca, Bolsón Cerrado	fbolsón@mail.com

Nombre
Frodo Bolsón

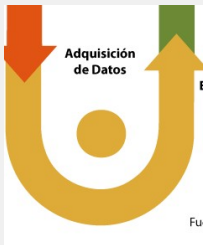
DNI
49886351D

Dirección
La Comarca, Bolsón Cerrado

Email
fbolsón@mail.com

Introducción - Datos Maestros

Adquisición de Datos



Planear, evaluar e incorporar nuevas fuentes de datos en la solución de Gestión de Datos Maestros debe ser un proceso confiable y repetible. Las actividades de adquisición de datos incluyen:

- Recibir y responder a nuevos requerimientos de adquisición de fuentes de datos
- Realizar evaluaciones rápidas, *ad-hoc*, de correspondencia y calidad de datos de alto nivel utilizando herramientas de limpieza de datos y creación de perfilados de datos.
- Evaluar y comunicar la complejidad de la integración de datos a los solicitantes para ayudarles en su análisis de costo-beneficio.
- Pilotaje de la adquisición de datos y su impacto en las reglas de concordancia.
- Finalización de métricas de calidad de datos para la nueva fuente de datos.
- Determinar quién será responsable de monitorear y mantener la calidad de los datos de una nueva fuente.
- Completar la integración en el ambiente general de gestión de datos.

ID fuente	Nombre	Dirección	Teléfono
123	John Smith	123 Main, Dataland, SQ 98765	
234	J. Smith	123 Main, Dataland, DA	2345678900
345	Jane Smith	123 Main, Dataland, DA	234-567-8900

Introducción - Datos Maestros

Validación, Normalización y Enriquecimiento de Datos



Los datos de entrada consistentes reducen la posibilidad de errores en la asociación de registros. Los procesos de preparación incluyen:



Validación
Busco posibles errores



Estandarización
Aseguro que datos se ajustan a estándares



Enriquecimiento
Agrego atributos para facilitar resolución entidades

Tabla 25 Datos de Entrada Normalizados y Enriquecidos

ID Fuente	Nombre	Dirección (limpiada)	Teléfono (limpiado)
123	John Smith	123 Main, Dataland, SQ 98765	
234	J. Smith	123 Main, Dataland, SQ 98765	+1 234 567 8900
345	Jane Smith	123 Main, Dataland, SQ 98765	+1 234 567 8900

Introducción - Datos Maestros

Resolución de Entidades y Gestion de Identificadores



La *resolución de entidad* es el proceso de determinar si dos referencias a objetos del mundo real se refieren al mismo objeto o a objetos diferentes.

Tabla 26 Identificación de Candidatos y Resolución de Identidad

ID Fuente	Nombre	Dirección (limpiada)	Teléfono (limpiado)	ID Candidato	ID de Parte
123	John Smith	123 Main, Dataland, SQ 98765		XYZ	1
234	J. Smith	123 Main, Dataland, SQ 98765	+1 234 567 8900	XYZ, ABC	2
345	Jane Smith	123 Main, Dataland, SQ 98765	+1 234 567 8900	ABC	2

- Correspondencia
- Resolución de Identidad
- Flujos de Trabajo de Correspondencia y Tipos de Reconciliación.
- Gestión de ID de Datos Maestros.
- Gestión de la Afiliación

Introducción - Datos Maestros

Resolución de Entidades y Gestion de Identificadores



Correspondencia o Identificación de Candidato

Tabla 26 Identificación de Candidatos y Resolución de Identidad

ID Fuente	Nombre	Dirección (limpiada)	Teléfono (limpiado)	ID Candidato	ID de Parte
123	John Smith	123 Main, Dataland, SQ 98765		XYZ	1
234	J. Smith	123 Main, Dataland, SQ 98765	+1 234 567 8900	XYZ, ABC	2
345	Jane Smith	123 Main, Dataland, SQ 98765	+1 234 567 8900	ABC	2

Riesgos

- **Falsos positivos:** Dos referencias que no representan la misma entidad están vinculadas con un único identificador.
- **Falsos negativos:** Dos referencias representan la misma entidad, pero no están enlazadas con un solo identificador.

Análisis de Similitud



Dos enfoques básicos

Los algoritmos **Determinísticos**. se basan en patrones y reglas definidos para asignar pesos y puntuaciones para determinar la similitud. Son predecibles.

Los algoritmos **Probabilísticos**. dependen de técnicas estadísticas para evaluar la probabilidad de que cualquier par de registros representan la misma entidad.

Introducción - Datos Maestros

Resolución de Entidades y Gestion de Identificadores



 Resolución de Identidad.

Tabla 26 Identificación de Candidatos y Resolución de Identidad

ID Fuente	Nombre	Dirección (limpiada)	Teléfono (limpiado)	ID Candidato	ID de Parte
123	John Smith	123 Main, Dataland, SQ 98765		XYZ	1
234	J. Smith	123 Main, Dataland, SQ 98765	+1 234 567 8900	XYZ, ABC	2
345	Jane Smith	123 Main, Dataland, SQ 98765	+1 234 567 8900	ABC	2

Algunas Correspondencia ocurren con gran confianza, otras son sugeridas con menos confianza.

- Es esencial mantener la historia de las correspondencias de modo que se puedan deshacer cuando se descubre que son incorrectos
- Es importante disponer de métricas de tasa de correspondencia para poder supervisar el impacto y la efectividad de las reglas de correspondencia.

Introducción - Datos Maestros

Resolución de Entidades y Gestion de Identificadores



Flujos de Trabajo de Correspondencia y Tipos de Reconciliación.

Tabla 26 Identificación de Candidatos y Resolución de Identidad

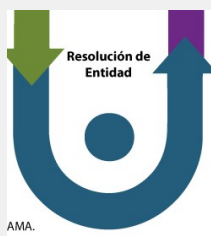
ID Fuente	Nombre	Dirección (limpiada)	Teléfono (limpiado)	ID Candidato	ID de Parte
123	John Smith	123 Main, Dataland, SQ 98765		XYZ	1
234	J. Smith	123 Main, Dataland, SQ 98765	+1 234 567 8900	XYZ, ABC	2
345	Jane Smith	123 Main, Dataland, SQ 98765	+1 234 567 8900	ABC	2

Las reglas de correspondencia requieren diferentes flujos de trabajo

- **Reglas de correspondencia de identificación duplicadas** identifican de forma única una entidad e identifican oportunidades de combinación sin tomar acción automática. Los Data stewards pueden revisar y decidir.
- **Reglas de ligas de correspondencia** identifican y hacen referencia cruzada de los registros que parecen relacionarse con un registro maestro sin actualizar el contenido del registro con referencias cruzadas. Las reglas de ligas de correspondencia son fáciles de implementar y mucho más fáciles de revertir
- **Reglas de unión de correspondencia** corresponden registros y unen los datos de estos registros en un registro único, unificado, reconciliado y exhaustivo.

Introducción - Datos Maestros

Resolución de Entidades y Gestion de Identificadores



Flujos de Trabajo de Correspondencia y Tipos de Reconciliación.

Las reglas de correspondencia **son complejas** y buscan proporcionar la versión unificada y reconciliada de la información a través de múltiples registros y fuentes de datos.

La complejidad **se debe a la necesidad de identificar qué campo y de qué fuente puede ser confiable** a partir de una serie de reglas.

La introducción de cada **nueva fuente puede cambiar las reglas** a través del tiempo.

Introducción - Datos Maestros

Resolución de Entidades y Gestion de Identificadores



Gestión de ID de Datos Maestros. Ids globales y x-Ref

Tabla 26 Identificación de Candidatos y Resolución de Identidad

ID Fuente	Nombre	Dirección (limpiada)	Teléfono (limpiado)	ID Candidato	ID de Parte
123	John Smith	123 Main, Dataland, SQ 98765		XYZ	1
234	J. Smith	123 Main, Dataland, SQ 98765	+1 234 567 8900	XYZ, ABC	2
345	Jane Smith	123 Main, Dataland, SQ 98765	+1 234 567 8900	ABC	2

Hay dos tipos de identificadores que necesitan ser gestionados a través de fuentes de datos en un entorno MDM:

IDs Globales: Un ID Global es el identificador único asignado y mantenido por la solución MDM adjunto a los registros reconciliados. Su propósito es identificar de forma única la instancia de la entidad. (ID de Parte tabla 26).

Información de Referencias Cruzadas (X-Ref): es la gestión de la relación entre los ID de origen y los ID Globales. La gestión de X-Ref debe incluir capacidades para mantener el historial de tales asignaciones para soportar métricas de tasa de coincidencia

Introducción - Datos Maestros

Resolución de Entidades y Gestion de Identificadores



 **Gestión de la Afiliación.** Relaciones de Afiliación y Relaciones Padre-Hijo

Tabla 26 Identificación de Candidatos y Resolución de Identidad

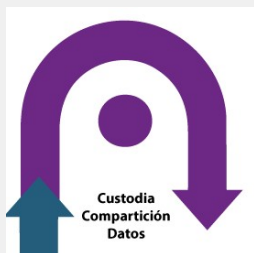
ID Fuente	Nombre	Dirección (limpiada)	Teléfono (limpiado)	ID Candidato	ID de Parte
123	John Smith	123 Main, Dataland, SQ 98765		XYZ	1
234	J. Smith	123 Main, Dataland, SQ 98765	+1 234 567 8900	XYZ, ABC	2
345	Jane Smith	123 Main, Dataland, SQ 98765	+1 234 567 8900	ABC	2

La Gestión de Afiliación es establecer y mantener relaciones entre los registros de entidades de Datos Maestros que tienen relaciones en el mundo real.

- **Las Relaciones de Afiliación.** El tipo de relación se puede utilizar para exponer dichos datos en una jerarquía padre-hijo. (por ejemplo, la Compañía X es una subsidiaria de la Compañía Y, una relación padre-hijo)
- Las **Relaciones Padre-hijo** requieren menos lógica de programación, ya que la estructura de navegación está implícita.

Introducción - Datos Maestros

Intercambio de Datos y Gestoría



Aunque gran parte del trabajo de Gestión de Datos Maestros puede ser automatizado a través de herramientas que permiten el procesamiento de un gran número de registros, todavía se requiere de gestoría **para resolver situaciones en las que los datos se corresponden incorrectamente.**

Introducción - Datos Maestros

Tipos de Datos Maestros

Datos Maestros de Parte (Party Data)

Datos sobre individuos, organizaciones y los roles que desempeñan.

Presentan diferentes desafíos como la identificación única, múltiples canales, múltiples fuentes...

Son particularmente desafiantes cuando las partes desempeñan diferentes roles. **un empleado que también es un cliente**

Datos Maestros Financieros

Datos sobre unidades de Negocio, centros de coste, cuentas libro mayor, presupuestos, etc. En ocasiones pueden simular cambios y proyecciones

Datos Maestros Legales

Datos sobre contratos, regulaciones.
Permiten análisis de contratos para permitir una mejor negociación.

Introducción - Datos Maestros

Tipos de Datos Maestros

Datos Maestros de Producto

Datos de Productos y Servicios internos.

Soportan diferentes funciones de negocio:

Datos Maestros de Ubicación

Proporcionan capacidad de rastrear y compartir información geográfica y crear relaciones y jerarquías.

- **Datos de Referencia de Ubicación** (países, provincias, etc) rara vez cambian.
- **Datos Maestros de Ubicación.** Direcciones de las partes de negocio y su localización.

Datos Maestros de la Industria - Directorios de Referencia

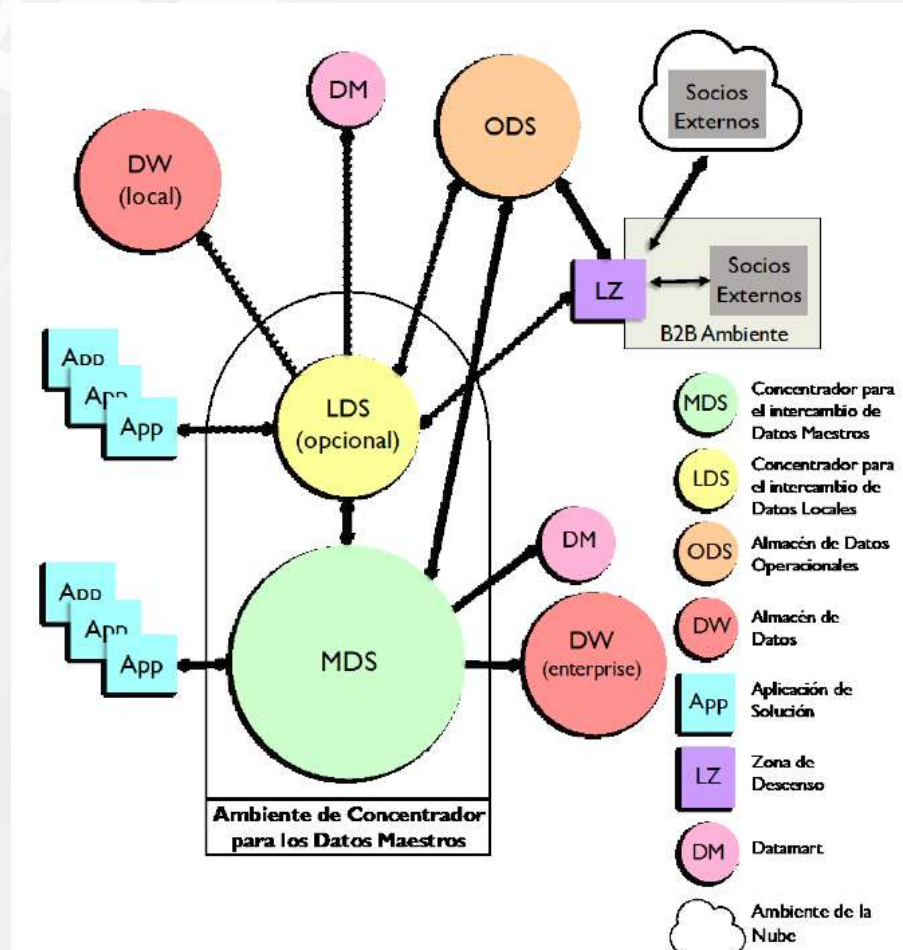
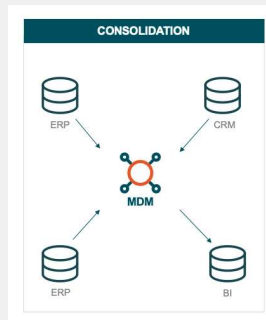
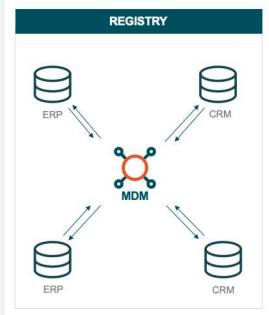
Directorios de Referencia son listas de datos maestros autorizadas que se pueden comprar. Son creados externamente, pero se pueden consolidar internamente.

Introducción - Datos Maestros

Arquitectura de Compartición de Datos

Tres enfoques básicos para implementar un *hub* de Datos Maestros:

- Un **Registro**. Índices a DM de SOF.
- En un **Nodo de Transacciones** DM en el concentrador y no en aplicaciones.
- Un enfoque **Consolidado**. Híbrido. Repositorio común no conlleva cambios en SOR pero duplica los datos



Actividades - Actividades del MDM

Definir Motivadores y Requerimientos del MDM

Los motivadores a menudo incluyen oportunidades para mejorar el servicio al cliente y/o la eficiencia operacional, así como para reducir los riesgos relacionados con la privacidad y el cumplimiento.

Los obstáculos incluyen diferencias en el significado de los datos y la estructura entre sistemas. Éstos a menudo están ligados a barreras culturales

Es relativamente fácil definir los requerimientos para los Datos Maestros dentro de una aplicación. Es más difícil definir los requerimientos estándar a través de aplicaciones

Actividades - Actividades del MDM

Evaluar las Fuentes de Datos

Es importante comprender la estructura y el contenido de los datos de las aplicaciones fuente y los procesos a través de los cuales se recopilan o crean.

Uno de los objetivos es comprender cómo se encuentran los datos completos respecto a sus Atributos.

Otro objetivo es entender la calidad de los datos. El mayor desafío la disparidad de las fuentes.

Los datos pueden ser de alta calidad dentro de una fuente, pero aun así no se ajustan con datos de otras fuentes, debido a diferencias estructurales y diferencias en los valores por los cuales se representan atributos similares. Las iniciativas de Datos Maestros brindan la oportunidad de definir e implementar estándares en aplicaciones en las que se crean o recopilan datos.

Actividades - Actividades del MDM

Definir el Enfoque Arquitectónico

Depende de estrategia de negocio, las plataformas existentes, su linaje y volatilidad y la latencia alta o baja.

Las pequeñas organizaciones pueden utilizar efectivamente un concentrador (*hub*) de transacciones, mientras que una organización global con múltiples sistemas es más probable que utilice un registro.

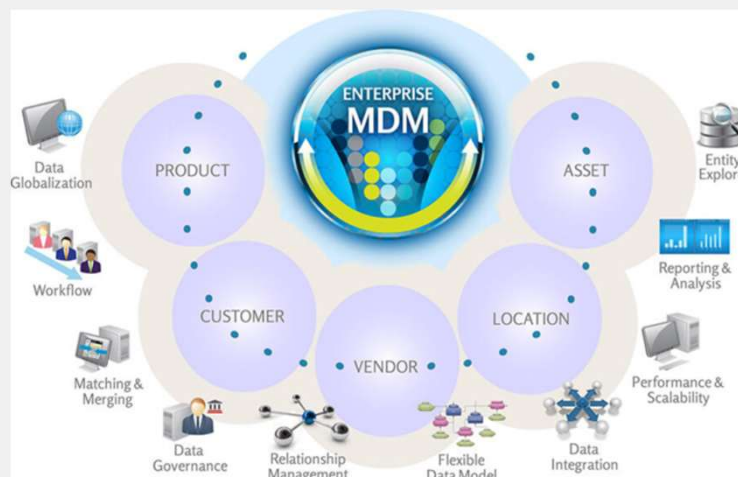


Actividades - Actividades del MDM

Modelar Datos Maestros

La Gestión de Datos Maestros es un proceso de integración de datos.

Para lograr resultados consistentes y para gestionar la integración de nuevas fuentes, a medida que una organización se expande, es necesario modelar los datos dentro de las áreas temáticas.



Actividades - Actividades del MDM

Definir Procesos de Custodia y Mantenimiento

Las soluciones técnicas pueden hacer trabajos excepcionales, de coincidencia, combinación y administración de identificadores para registros maestros. Sin embargo, **el proceso también requiere custodia**

Los proyectos MDM deben considerar los recursos necesarios para apoyar la calidad continua de los Datos Maestros.



Actividades - Actividades del MDM

Establecer Políticas de Gobierno para Reforzar el Uso de Datos Maestros

El lanzamiento inicial de un esfuerzo de Datos Maestros es un reto y requiere enfoque.

Los beneficios no se ven hasta que los sistemas y las personas comienzan a usar los Datos reales.

El esfuerzo debe incluir una hoja de ruta para que los sistemas adopten valores maestros e identificadores como entrada a los procesos.

Actividades - Actividades del RDM

Definir Motivadores y Requerimientos del RDM

Los principales motivadores de la Gestión de Datos de Referencia son la **eficiencia operativa y la mayor calidad** de los datos.

Manejar centralizadamente los Datos de Referencia es más rentable que tener múltiples unidades de negocios que mantienen sus propios conjuntos de datos.

Actividades - Actividades del RDM

Evaluar las Fuentes de Datos

La mayoría de los conjuntos de Datos de Referencia estándar de la industria se pueden obtener de las organizaciones que los crean y los mantienen.

La mayoría de las organizaciones también se basan en Datos de Referencia que se crean y se mantienen internamente

Actividades - Actividades del RDM

Definir el Enfoque Arquitectónico

Antes de adquirir o construir una herramienta para gestionar los Datos de Referencia, es fundamental tener en cuenta los requisitos y retos planteados por la gestión de los Datos de Referencia. Frecuencia de actualizaciones, modelos de consumo, si es necesario mantener histórico, etc.

El enfoque arquitectónico debe reconocer que, invariablemente, algunos Datos de Referencia deberán actualizarse manualmente.

Los *data stewards* deben programar las actualizaciones conocidas para que se alineen con la publicación de nuevos códigos. Los consumidores de datos deben ser informados de cada los cambios.

Actividades - Actividades del RDM

Modelar los Conjuntos de Datos de Referencia

Con el fin de permitir el uso a largo plazo y establecer Metadatos precisos, así como para el propio proceso de mantenimiento, es valioso crear modelos de datos de conjuntos de Datos de Referencia. Los modelos ayudan a los consumidores de datos a comprender las relaciones dentro del conjunto de Datos de Referencia y pueden utilizarse para establecer reglas de calidad.

Actividades - Actividades del RDM

Definir Procesos de Custodia y Mantenimiento

Los Datos de Referencia requieren una administración responsable para asegurar que los valores estén completos y actualizados y que las definiciones sean claras y comprensibles.

Como parte del proceso de administración, es útil capturar Metadatos básicos sobre cada conjunto de Datos de Referencia.

Muchas herramientas de Gestión de Datos de Referencia incluyen flujos de trabajo para gestionar la revisión y la aprobación de cambios. Estos flujos de trabajo dependen de identificar quién dentro de una organización es responsable del contenido de los Datos de Referencia.

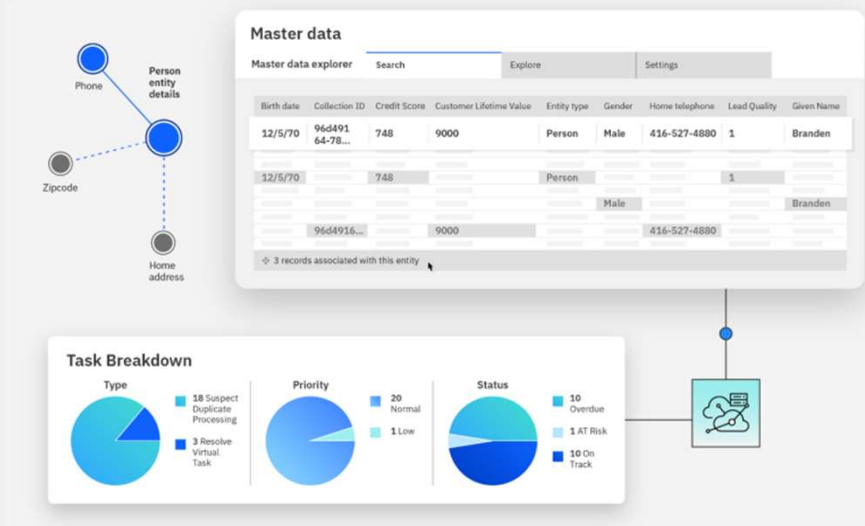
Actividades - Actividades del RDM

Establecer Políticas de Gobierno de Datos de Referencia

Una organización solo obtiene el valor de un repositorio de Datos de Referencia centralizada si las personas realmente usan los datos de ese repositorio.

Es importante contar con políticas que gobiernen la calidad y exijan el uso de Datos de Referencia de ese repositorio.

Herramientas y Técnicas



MDM requiere herramientas específicamente diseñadas para permitir la administración de identidades.

La incorporación de estos servicios puede permitir a las organizaciones utilizar las mejores soluciones, integrándolas a su arquitectura empresarial global para satisfacer necesidades específicas.

Guías de implementación

Guía de Implementación

Implantar programa de MDM y RDM requiere perfiles de negocio y de TI.

Las organizaciones deben esperar implementar soluciones de Referencia y de Datos Maestros de forma **incremental** a través de una serie de proyectos definidos en una hoja de ruta de implementación, priorizados basados en las necesidades del negocio y guiados por una arquitectura general. Sin gobierno fallará.

Adhiérase a la Arquitectura de Datos Maestros

Establecer y seguir una arquitectura de referencia apropiada es fundamental para gestionar y compartir Datos Maestros en una organización.

Monitoree el Movimiento de Datos

Dentro de un entorno de referencia o de Datos Maestros, el flujo de datos debe ser monitoreado con el fin de:

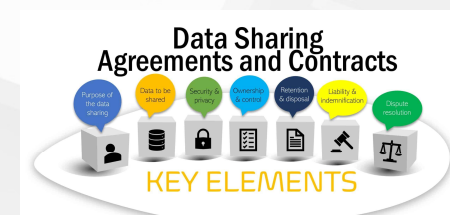
- Mostrar **cómo se comparten** y utilizan los datos
- Identificar **linaje de Datos**
- Ayudar a analizar **las causas de los problemas**
- Demostrar la efectividad de las técnicas de integración
- Denotar la **latencia** de valores de datos de los sistemas fuente a través del consumo
- Determinar la **validez de las reglas de negocio** y las transformaciones ejecutadas

Gestione el Cambio de Datos de Referencia



Acuerdos de Compartición de Datos

Para asegurar el acceso y uso adecuados, se deben establecer acuerdos de compartición que estipulen qué datos pueden ser compartidos y en qué condiciones.



Organización y Cambio Cultural

- La Gestión de Datos Maestros y de Referencia requiere que **las personas renuncien al control** de algunos de sus datos y procesos para crear recursos compartidos.
- La mejora de la disponibilidad y la calidad de los Datos de Referencia y Maestros indudablemente requerirá **cambios en los procedimientos y prácticas tradicionales.**
- Las soluciones deben ser delimitadas e implementadas en base a la preparación organizacional actual y las necesidades futuras relacionadas con la misión y la visión de la organización.
- Quizás el cambio cultural más desafiante es fundamental para la gobernabilidad: **Determinar qué individuos son responsables de las decisiones.**

Gobierno de Datos Maestros y de Referencia

➤ Debido a que son recursos compartidos, los Datos Maestros y de Referencia requieren gobierno y custodia

Los procesos de gobernabilidad determinarán:

- Las fuentes de datos a integrar
- Las normas de calidad de datos que se aplicarán
- Las condiciones de uso a seguir
- Las actividades por supervisar y la frecuencia de seguimiento
- Los niveles de prioridad y respuesta de los esfuerzos de custodia de datos
- Cómo se debe representar la información para satisfacer las necesidades de las partes interesadas
- Compuertas de aprobación estándar, expectativas en la implementación de RDM y MDM

Como un proceso continuo, el gobierno de los datos debe tener la habilidad de revisar, recibir y considerar nuevos requisitos y cambios a las reglas existentes, a la vez que los principios, reglas y directrices están disponibles para aquellos que usan regularmente Datos de Referencia y Maestros

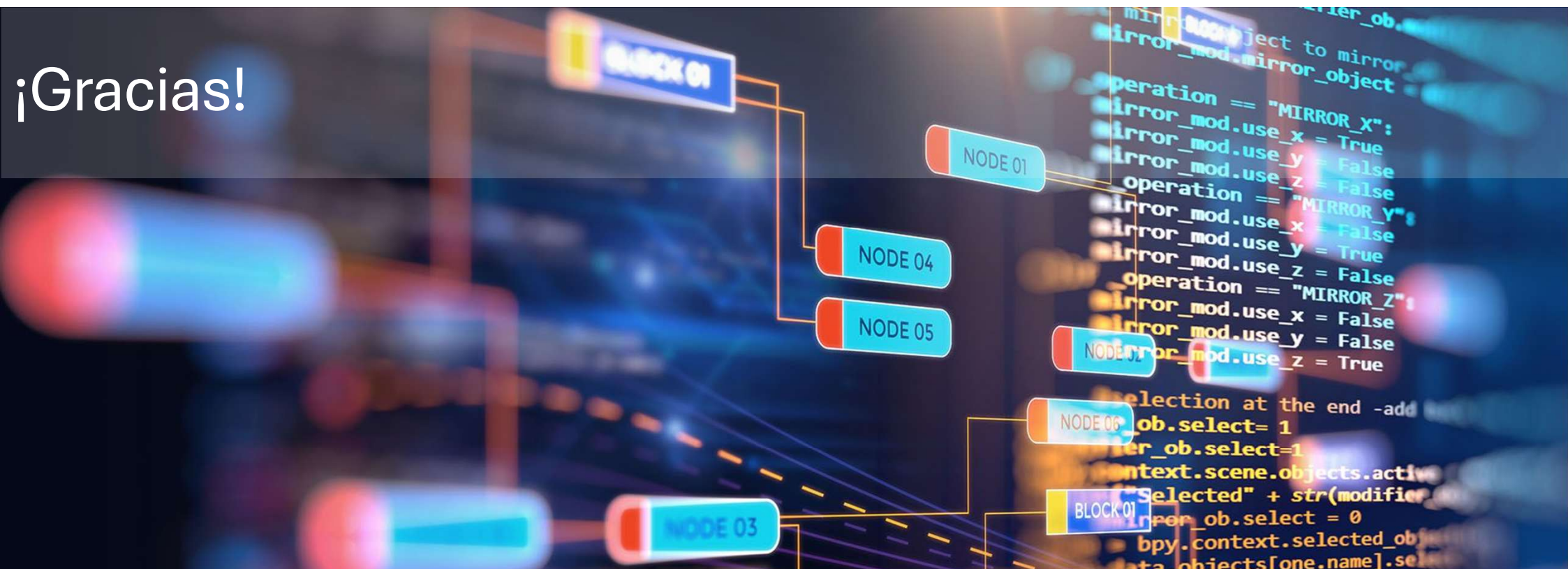
Gobierno de Datos Maestros y de Referencia

Métricas

Ciertas métricas pueden vincularse a la calidad de los Datos de Referencia y Maestros y a los procesos que apoyan estos esfuerzos

- Calidad y cumplimiento de los datos. Según reglas DQ
- Actividad de cambio de datos- tasa de cambio de valores
- Ingesta y consumo de datos. quién provee y quien consume
- Acuerdos de Nivel de Servicio. Nivel de adherencia a los SLA
- Coste Total de Propiedad. Licencias, personal de soporte
- Volumen y uso compartido de datos. Volúmenes de ingestión y consumo

¡Gracias!



GOLD SPONSORS



SILVER SPONSORS



Entidades académicas

