

Gobierno de Datos

Módulo #1

¿Qué es Gobierno de Datos?

¿Qué es Gobierno de Datos?

“

*Gobierno de Datos (Data Governance) es un conjunto de **políticas, procesos, roles y tecnologías** que garantizan la **disponibilidad, integridad, confidencialidad y usabilidad de los datos** dentro de una organización. Su objetivo es **asegurar que los datos sean confiables, estén bien gestionados y se utilicen de manera efectiva para la toma de decisiones***

¿Qué es Gobierno de Datos?

Para lograr una definición, podemos centrarnos en el Framework de referencia en esta materia, que es el que plantea DAMA International, en el que a través de la representación de una rueda, se muestran 10 dimensiones de la gestión de datos que son parte de Gobierno de Datos



¿Qué es Gobierno de Datos?

Otra definición, es abordarlo como la gestión y administración de todas las actividades basadas en Datos, implementando Procesos de Gestión Operacional, Metodologías y Roles



Calidad de Datos



**Protección y Acceso
Ágil**



**Fuentes Oficiales y
Catálogo**



**Ciclo de Vida de los
Datos**



**Asignación de
responsables**



**Instancias
Decisionales y
Análisis de Impacto**



Políticas, Lineamientos y estándares



Agilidad



Plataformas de Gestión



Gestión de Cambio

¿Cuál es su importancia y rol en las organizaciones modernas?

Gobierno de Datos es clave para que las organizaciones gestionen, protejan y aprovechen sus datos de manera eficiente. Su importancia radica en varios aspectos estratégicos y operativos:



Mejora en la toma de decisiones

Datos confiables y de calidad permiten que las decisiones sean más precisas y basadas en información real, reduciendo el riesgo de errores



Cumplimiento normativo y seguridad

Garantiza que la organización cumpla con regulaciones como **GDPR, CCPA, ISO 27001** o leyes locales de protección de datos, evitando sanciones y riesgos legales



Protección de la privacidad y seguridad de la información

Evita filtraciones, accesos no autorizados y mal uso de datos, reduciendo riesgos de fraude y ciberataques



Optimización de la eficiencia operativa

Reduce la duplicidad de datos, mejora la integración entre sistemas y evita la inconsistencia en los reportes, ahorrando tiempo y recursos



Potencia la analítica avanzada y la inteligencia artificial

Facilita la implementación de **machine learning, big data y business intelligence**, asegurando que los modelos trabajen con datos confiables



Alineación con la estrategia del negocio

Permite que los datos estén alineados con los objetivos de la empresa, impulsando la innovación, la monetización y la competitividad



Facilita la colaboración y la gestión del conocimiento

Define roles y responsabilidades claras (Data Owners, Data Stewards), promoviendo una cultura de datos dentro de la organización

Gobierno de Datos

Módulo #2

Principios clave de Gobierno de Datos

¿Qué son los Principios Clave de Gobierno de Datos?

“

Los principios clave del **Gobierno de Datos** son los fundamentos que guían la gestión de datos dentro de una organización, asegurando su calidad, seguridad y alineación con los objetivos de negocio. Estos principios pueden variar según el marco adoptado (como DAMA-DMBOK)

¿Cuáles son los Principios de Gobierno de Datos?



Responsabilidad



**Calidad de
Datos**



**Disponibilidad y
Usabilidad**



**Seguridad y
Privacidad**



**Transparencia y
Trazabilidad**



**Valor y Alineación
con el Negocio**



**Cumplimiento y
Regulación**



**Interoperabilidad y
Estándares**



**Gestión del Cambio
y Cultura de Datos**



**Adaptabilidad y
Mejora Continua**



Responsabilidad

Cada dato debe tener un **dueño** o **steward** responsable de su gestión, calidad y seguridad. Esto implica definir roles claros y líneas de autoridad dentro de la organización



Calidad de Datos

Los datos deben cumplir con estándares de **precisión, completitud, consistencia, confiabilidad y actualidad**. Esto requiere establecer reglas de calidad y procesos de monitoreo continuo



Disponibilidad y Usabilidad

Los datos deben estar **accesibles** para los usuarios autorizados cuando los necesiten, en formatos adecuados y con la documentación suficiente para su interpretación



Seguridad y Privacidad

Se deben implementar controles para **proteger los datos** de accesos no autorizados, pérdida o manipulación indebida, respetando regulaciones como GDPR o CCPA

Principios de Gobierno de Datos



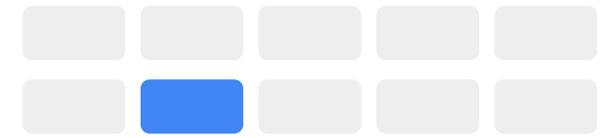
Transparencia y Trazabilidad

Las decisiones sobre datos y sus cambios deben ser **auditables**. Se requiere un seguimiento claro del linaje de los datos, desde su origen hasta su uso final



Valor y Alineación con el Negocio

El gobierno de datos debe generar **valor** para la organización, alineando sus iniciativas con los objetivos estratégicos y operacionales del negocio



Cumplimiento y Regulación

Los datos deben ser gestionados en conformidad con **leyes y normativas** aplicables, evitando riesgos legales y financieros



Interoperabilidad y Estándares

Se deben establecer estándares y modelos para que los datos sean **compatibles y reutilizables** en diferentes sistemas y procesos dentro de la organización



Gestión del Cambio y Cultura de Datos

El éxito del gobierno de datos depende de una **cultura organizacional** que promueva la importancia de los datos y fomente buenas prácticas a nivel transversal



Adaptabilidad y Mejora Continua

El gobierno de datos no es estático. Debe ser un proceso **iterativo y ágil**, ajustándose a nuevas tecnologías, necesidades de negocio y regulaciones emergentes

Gobierno de Datos

Módulo #3

Beneficios de Gobierno de Datos para las Organizaciones

Beneficios de Gobierno de Datos para las Organizaciones



Estrategia de Datos alineada con la Estrategia Corporativa

- Sponsor de la Alta Dirección es fundamental
- Hacerlo bottom up es muy complicado, porque es necesario priorizar agendas de muchas personas del negocio
- Dotar de Recursos y presupuesto



Visión Global, pero Realista

- Una de Corto (Casos de Uso) y Mediano Plazo, tener despierto el interés de ir mostrando resultados



No es un Proyecto, llega para quedarse

- Es un proceso continuo, por lo que llega a la Cadena de Valor de la Compañía



La Gestión de los Datos es un tema cultural, transversal a toda la compañía

- Todos deben adaptarse a las mejores prácticas, por lo que la Gestión del Cambio juega un rol fundamental



Entender que la gobernanza y la gestión de datos son visiones diferentes, pero complementarias

- Gobierno tiene que ver con entender los activos de datos
- La Gestión de Datos tiene que ver con cómo administramos el ciclo operacional de los datos, de acuerdo a las buenas prácticas de gestión



No dejar fuera, dentro del Programa de Gobierno de Datos, la Estrategia de Data Compliance

Beneficios de Gobierno de Datos para las Organizaciones

El despliegue de un programa de Gobierno de Datos responde preguntas de negocio cruciales acerca de los datos, impactando la toma de decisiones, la eficiencia operativa y el cumplimiento normativo



1. ¿Qué datos tiene la organización y dónde están ubicados?

- Responde a la necesidad de identificar, catalogar y mapear los activos de datos, incluidas bases de datos, aplicaciones, y sistemas.



2. ¿Quién es Responsable de los Datos y cuáles son sus roles?

- Define los propietarios, custodios y usuarios de los datos, asegurando que cada rol entienda sus responsabilidades en la gestión de los datos.



3. ¿Qué políticas y estándares regulan el uso de los datos?

- Garantiza que exista un marco normativo que incluya políticas de acceso, seguridad, privacidad y calidad.



4. ¿Cómo se garantiza la calidad de los datos?

- Aborda la implementación de procesos y métricas para asegurar la precisión, completitud, consistencia y actualidad de los datos.



5. ¿Cómo se gestionan los riesgos asociados a los datos?

- Identifica riesgos como el acceso no autorizado, pérdida de datos, incumplimiento normativo o uso indebido, y detalla cómo se mitigan.



6. ¿Cómo se asegura la privacidad y seguridad de los datos?

- Responde cómo se establece quién tiene acceso a qué datos, y bajo qué circunstancias, asegurando el cumplimiento del principio de menor privilegio.



7. ¿Cómo se controla el acceso a los datos?

- Responde cómo se establece quién tiene acceso a qué datos, y bajo qué circunstancias, asegurando el cumplimiento del principio de menor privilegio.



8. ¿Cómo se garantiza el cumplimiento de regulaciones externas e internas?

- Establece cómo el programa monitorea y asegura el cumplimiento con leyes, estándares y políticas específicas de la industria.



9. ¿Cómo se mide el éxito del programa de Gobierno de Datos?

- Define métricas clave, como la calidad de los datos, reducción de riesgos, cumplimiento normativo, y satisfacción de los usuarios.



10. ¿Cómo se gestiona el ciclo de vida de los datos?

- Explica cómo se administra la creación, almacenamiento, uso, archivo y eliminación de los datos, asegurando que sean gestionados de manera eficiente y segura durante todo su ciclo de vida.

Gobierno de Datos

Módulo #4

Conceptos Fundamentales de Gobierno de Datos

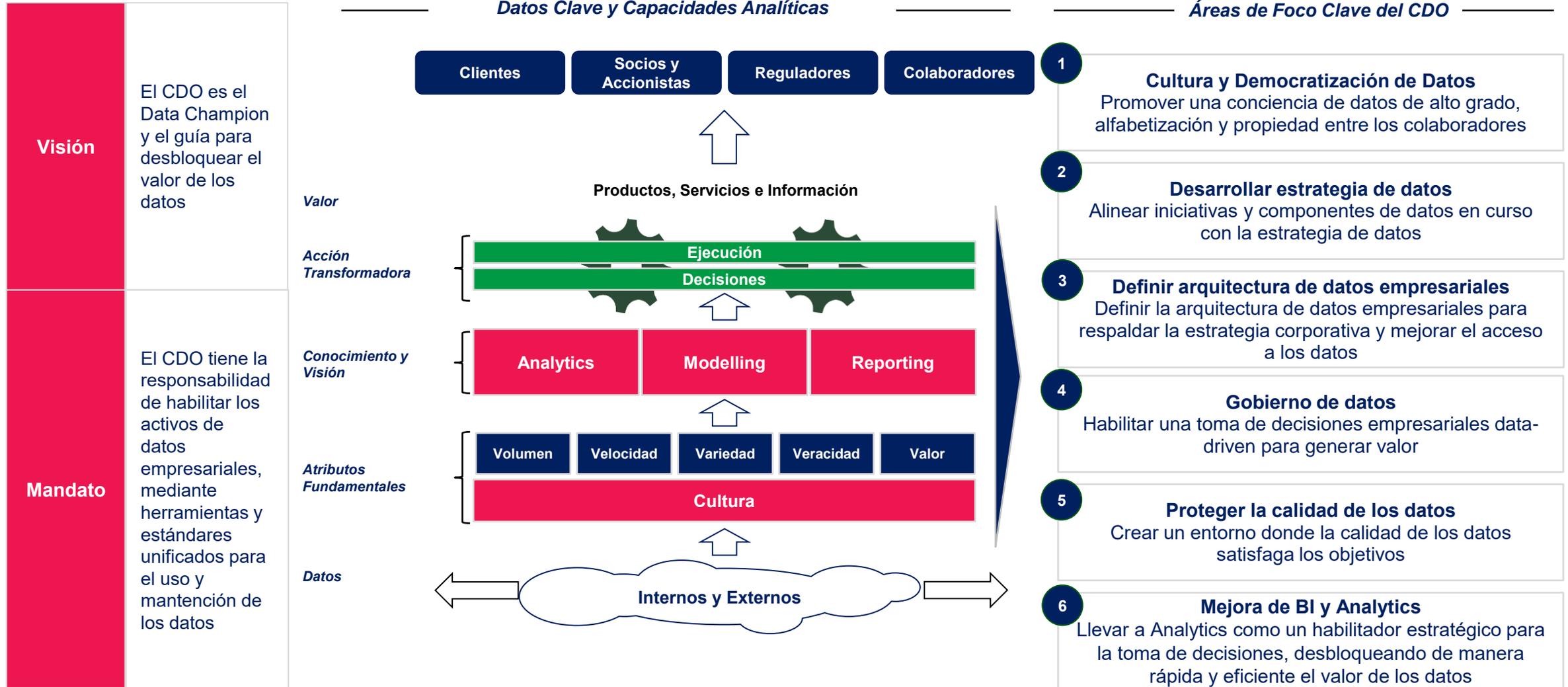
Conceptos Fundamentales de Gobierno de Datos

Personas, procesos, tecnología y datos: el marco básico para estructurarlo



Conceptos Fundamentales de Gobierno de Datos

Modelo de Gobierno de Datos



Gobierno de Datos

Módulo #5

Regulación y Cumplimiento en Gobierno de Datos

Regulación y Cumplimiento en Gobierno de Datos

¿Por qué es importante la regulación en Gobierno de Datos?

- ✓ Protección de la privacidad y derechos de los usuarios
- ✓ Reducción de riesgos legales y financieros
- ✓ Generación de confianza en clientes y stakeholders
- ✓ Asegurar la calidad y seguridad de los datos



Regulación y Cumplimiento en Gobierno de Datos

¿Qué es GDPR?

Definición breve:

- Reglamento General de Protección de Datos (GDPR) de la UE
- Establece reglas sobre recopilación, almacenamiento y procesamiento de datos personales

Principales requerimientos:

- ✓ Consentimiento explícito del usuario
- ✓ Derecho al acceso, rectificación y olvido de los datos
- ✓ Notificación de brechas de seguridad en menos de 72 horas
- ✓ Limitaciones en transferencia de datos fuera de la UE



Regulación y Cumplimiento en Gobierno de Datos

¿Cuál es el alcance de GDPR?

Alcance

Aplicación:	Todos los sectores
Protección:	Datos personales
Penalización:	Hasta 20M€ o 4% de ingresos
Notificación:	72 horas

Impacto en la gestión de datos:

- ✓ Procesos de **seguridad y auditoría** más estrictos
- ✓ **Retención y eliminación** de datos conforme a regulaciones
- ✓ Mayor inversión en **tecnología de protección de datos**
- ✓ Necesidad de **equipos especializados en cumplimiento**

Regulación y Cumplimiento en Gobierno de Datos

Conclusión

- ✓ GDPR es clave para proteger los datos y evitar sanciones
- ✓ Su impacto va más allá del cumplimiento: mejora seguridad y confianza
- ✓ Las empresas deben adaptarse continuamente a los cambios normativos

Gobierno de Datos

Módulo #6

Roles y Responsabilidades en Gobierno de Datos

Conceptos Fundamentales de Gobierno de Datos

¿Qué es un Dominio de Datos?

Imaginemos un terreno, en el que se encuentran todos los datos de una Empresa



Estos datos, al igual que el terreno, pueden ser delimitados según distintos objetivos, por ejemplo: la propiedad



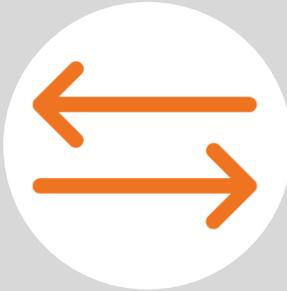
Conceptos Fundamentales de Gobierno de Datos

¿Qué es un Dominio de Datos?

Los Dominios de Datos son pilares fundamentales de Gobierno de Datos, agrupando y organizando la información según áreas de negocio o temas específicos, facilitando una gestión más eficiente, segura y alineada con los objetivos estratégicos de la organización. Usualmente se incluyen todos los tipos de datos: estructurados, no estructurados, flujos de datos (tiempo real), data histórica, proyecciones, y metadatos

Un Dominio de Datos ES...

- Una agrupación lógica de datos con propósito de propiedad.
- Cada Dominio de Datos debe ser propiedad de un rol con autoridad jerárquica
- Cada Dominio de Datos tiene un Owner (Propietario)
- Algunos roles pueden ser propietarios de múltiples dominios de datos



Un Dominio de Datos NO ES...

- Una estructura organizacional (aunque puede coincidir)
- Grupo de Sistemas de TI
- Data storage or Sistemas Fuente
- Modelo de Capacidad de Negocio
- Cadena de Suministro o Modelo de Proceso

Conceptos Fundamentales de Gobierno de Datos

Personas y Roles de Gobierno de Datos



Chief Data Officer (CDO):

Ejecutivo responsable de la estrategia de datos en la organización. Define políticas, impulsa el gobierno de datos y alinea la gestión de datos con los objetivos del negocio



Data Owner:

Responsable de la calidad, integridad y seguridad de los datos dentro de su dominio. Define reglas de uso y asegura el cumplimiento de políticas de gobierno de datos



Data Steward:

Rol operativo encargado de aplicar las políticas de calidad, integridad y gestión de datos en el día a día. Monitorea y corrige problemas en los datos

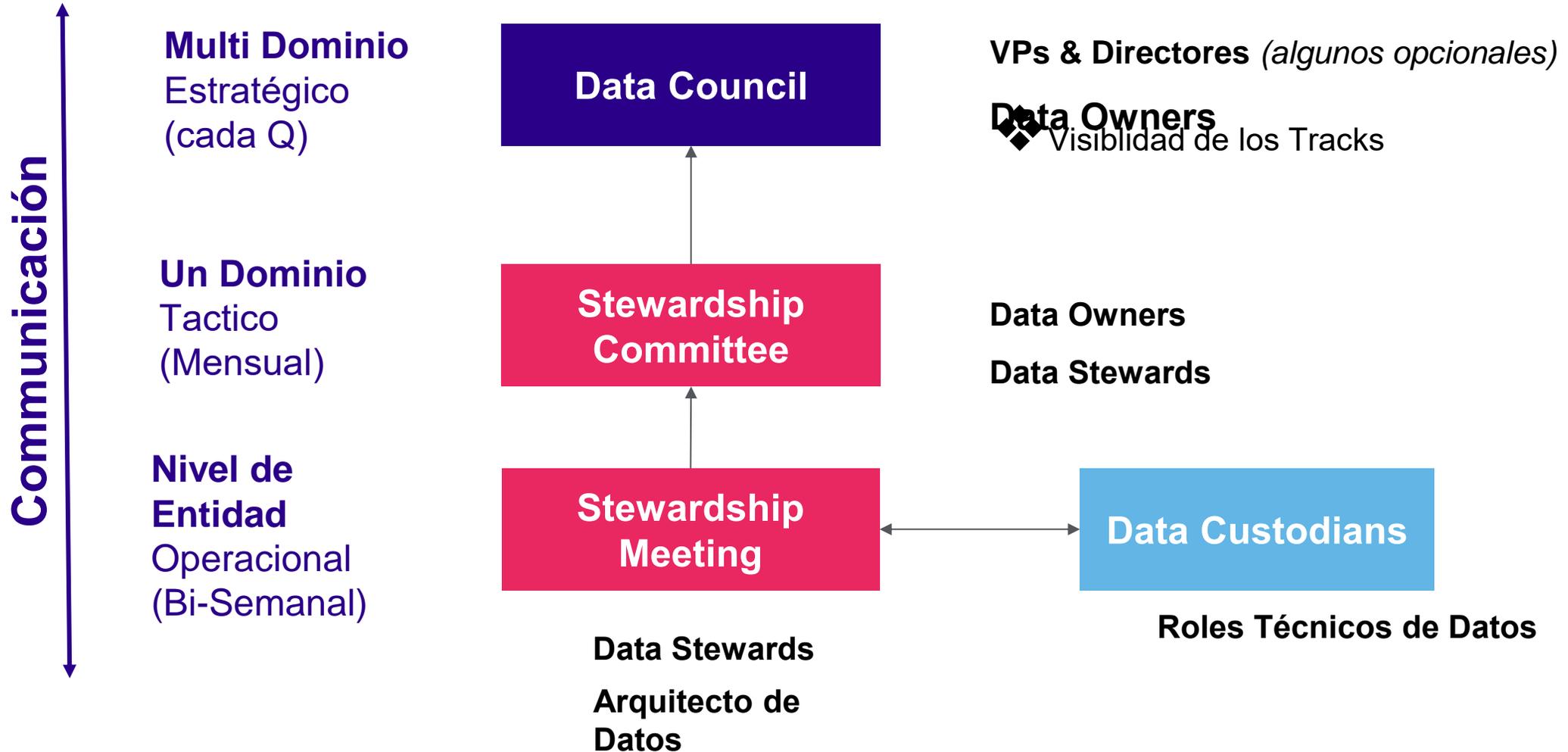


Data Custodian:

Responsable de la administración técnica de los datos. Gestiona su almacenamiento, seguridad y acceso dentro de las plataformas tecnológicas

Conceptos Fundamentales de Gobierno de Datos

Instancias de Gobierno de Datos



Gobierno de Datos

Módulo #7

Data Governance Framework

Data Governance Framework: DAMA

¿Qué es DAMA?

Data Management Association International (DAMA International) es una organización profesional sin fines de lucro dedicada al desarrollo de estándares internacionales para profesionales de la gestión de datos. Fundada en 1980 en Los Ángeles, DAMA ha crecido hasta convertirse en una comunidad global que abarca capítulos en todo el mundo.

DAMA International ofrece recursos educativos, certificaciones y oportunidades de desarrollo profesional en el campo de la gestión de datos.

Uno de sus aportes más destacados es la publicación del “Data Management Body of Knowledge” (DMBOK), una guía integral que establece estándares y mejores prácticas en áreas como arquitectura de datos, seguridad, calidad, modelado y gobernanza de datos.

Además, DAMA otorga la **certificación “Certified Data Management Professional” (CDMP)**, reconocida a nivel mundial, que valida los conocimientos y habilidades de los profesionales en gestión de datos.

Promueve la colaboración y el intercambio de conocimientos entre profesionales, fortaleciendo la comunidad de gestión de datos a nivel local e internacional.



Data Governance Framework

La Rueda (Wheel) de DAMA

La rueda de **DAMA** es un modelo conceptual desarrollado por **DAMA International** para representar las principales disciplinas de la gestión de datos. Se basa en el **Data Management Body of Knowledge (DMBOK)** y sirve como una guía estructurada para profesionales que trabajan en **Data Management**

La rueda de DAMA está compuesta por **11 áreas de conocimiento**



Data Governance Framework

La Rueda (Wheel) de DAMA



registros.

clave.

mediante la analítica.

1. **Data Governance** (Gobernanza de Datos): Establece políticas, procesos y roles para la gestión de datos.
2. **Data Architecture** (Arquitectura de Datos): Diseña la estructura de los datos dentro de la organización.
3. **Data Modeling & Design** (Modelado y Diseño de Datos): Crea modelos conceptuales, lógicos y físicos para organizar los datos.
4. **Data Storage & Operations** (Almacenamiento y Operaciones de Datos): Administra bases de datos y almacenes de datos.
5. **Data Security** (Seguridad de Datos): Protege la confidencialidad, integridad y disponibilidad de los datos.
6. **Data Integration & Interoperability** (Integración e Interoperabilidad de Datos): Garantiza la conectividad y fluidez entre sistemas.
7. **Documents & Content Management** (Gestión de Documentos y Contenido): Administra datos no estructurados como archivos y registros.
8. **Reference & Master Data Management** (Gestión de Datos Maestros y de Referencia): Asegura consistencia y unicidad en los datos clave.
9. **Data Warehousing & Business Intelligence** (Almacenes de Datos e Inteligencia de Negocios): Facilita la toma de decisiones mediante la analítica.
10. **Metadata Management** (Gestión de Metadatos): Administra información descriptiva sobre los datos.
11. **Data Quality** (Calidad de Datos): Garantiza la precisión, completitud y confiabilidad de los datos.

Gobierno de Datos

Módulo #8

Políticas y Estándares de Datos

Políticas y Estándares de Datos

Introducción



¿Por qué son importantes las políticas y estándares en Gobierno de Datos?



Aseguran el **cumplimiento normativo**
(GDPR, HIPAA, ISO 27001)



Mejoran la **calidad de datos** y la **seguridad**



Facilitan la **toma de decisiones basada en datos confiables**



Reducen riesgos operativos y financieros

Políticas y Estándares de Datos

Cómo definir políticas claras y establecer estándares organizacionales

Definición breve:

Políticas de Datos:

Directrices generales que establecen cómo deben gestionarse los datos en la organización.

Estándares:

Normas y procedimientos específicos para garantizar consistencia y cumplimiento.

Políticas y Estándares de Datos

Componentes Claves de una Política de Datos

1. **Objetivo** (Propósito de la política)
2. **Alcance** (A qué datos y áreas afecta)
3. **Principios** (Reglas generales a seguir)
4. **Responsabilidades** (Roles clave: Data Owners, Data Stewards, IT)
5. **Mecanismos de Cumplimiento** (Monitoreo y auditoría)



Políticas y Estándares de Datos

Pasos para Definir Políticas de Datos

- 1. Identificar Necesidades del Negocio** (Requisitos regulatorios y operacionales)
- 2. Definir Objetivos Claros** (Ej: Protección de datos sensibles, integridad de datos, accesibilidad)
- 3. Involucrar Stakeholders** (Gobernanza de datos, IT, legal, negocio)
- 4. Desarrollar el Documento de Política**
- 5. Aprobar y Comunicar** (Capacitaciones y divulgación)
- 6. Monitorear y Actualizar Periódicamente**



Políticas y Estándares de Datos

¿Qué son los Estándares Organizacionales?

- ¿Qué son los estándares?
 - Reglas técnicas y procedimientos para gestionar datos.
- Ejemplos:
 - **Calidad de Datos:** Reglas para garantizar consistencia y exactitud.
 - **Seguridad de Datos:** Control de accesos y encriptación.
 - **Integración de Datos:** Formatos y protocolos para interoperabilidad.



Políticas y Estándares de Datos

Beneficios de los Estándares de Datos

- Homogeneidad en la gestión de datos
- Reducción de errores y redundancias
- Facilita la colaboración entre áreas
- Asegura que los datos sean reutilizables y confiables



Políticas y Estándares de Datos

Ejemplo de Política y Estándar en una Empresa

Caso: Empresa de Seguros

Política:

- "Los datos de clientes deben estar actualizados cada 6 meses."

Estándar:

- "Formato de identificación: RUT con validación numérica y sin caracteres especiales."

Políticas y Estándares de Datos

Claves para el Éxito



**Alineación con los
objetivos del negocio**



**Compromiso de la alta
dirección**



**Educación y
concienciación
organizacional**



**Uso de tecnología para
automatizar
cumplimiento**

Gobierno de Datos

Módulo #9

Gobierno de Datos vs Gestión de Datos

Gobierno de Datos vs Gestión de Datos

Diferencias clave y cómo se complementan

🌟 Gobierno de Datos (Data Governance)

◆ Enfoque: Estratégico y normativo

El **Gobierno de Datos** establece las **políticas, estándares y responsabilidades** para garantizar el uso adecuado de los datos en una organización. Se centra en la toma de decisiones y en la definición de reglas que regulan la gestión de los datos.

◆ Principales responsabilidades:

- ✓ Definir políticas y normativas de datos
- ✓ Establecer roles y responsabilidades (Data Owners, Data Stewards, etc.)
- ✓ Asegurar el cumplimiento de regulaciones (GDPR, HIPAA, etc.)
- ✓ Garantizar la **calidad, seguridad y privacidad** de los datos
- ✓ Resolver conflictos sobre la propiedad y uso de los datos
- ✓ Facilitar la toma de decisiones basada en datos confiables

◆ Ejemplo:

El equipo de **Gobierno de Datos** define una política de calidad de datos que establece que el 95% de los registros de clientes deben tener un número de teléfono válido antes de ser utilizados en campañas de marketing.

🌟 Gestión de Datos (Data Management)

◆ Enfoque: Operacional y técnico

La **Gestión de Datos** se encarga de las **actividades diarias** relacionadas con la administración, almacenamiento, procesamiento y uso de los datos. Se centra en la implementación de las políticas establecidas por el Gobierno de Datos.

◆ Principales responsabilidades:

- ✓ Modelado y diseño de datos
- ✓ Almacenamiento y operaciones de bases de datos
- ✓ Integración e interoperabilidad de datos
- ✓ Administración de datos maestros y metadatos
- ✓ Implementación de controles de calidad de datos
- ✓ Seguridad y acceso a los datos

◆ Ejemplo:

El equipo de **Gestión de Datos** desarrolla un sistema automatizado que valida los números de teléfono antes de cargarlos en la base de datos de clientes, asegurando que cumplan con la política definida por el Gobierno de Datos.

Gobierno de Datos vs Gestión de Datos

Comparación Resumida

Aspecto	Gobierno de Datos (Data Governance)	Gestión de Datos (Data Management)
Enfoque	Estratégico y normativo	Operacional y técnico
Objetivo	Definir reglas y políticas	Ejecutar y administrar datos
Responsabilidad	Control y supervisión	Implementación y mantenimiento
Ejemplo de tarea	Crear una política de calidad de datos	Implementar reglas de validación en una base de datos

Gobierno de Datos vs Gestión de Datos

Conclusión

“

*El **Gobierno de Datos** define “el qué” y “el por qué”, mientras que la **Gestión de Datos** ejecuta “el cómo”. Ambos son esenciales para lograr una administración efectiva de los datos y asegurar que sean confiables, seguros y útiles para la organización.*

Gobierno de Datos

Módulo #10

Cultura de Datos: Promoviendo la alfabetización de datos

Cultura de datos: Promoviendo la alfabetización de datos

Introducción: ¿Qué es la cultura de datos y su importancia en la organización?

📌 Definición de Cultura de Datos

La **Cultura de Datos** es el conjunto de valores, comportamientos y prácticas dentro de una organización que promueven el uso de los datos como un activo estratégico. Una organización con una cultura de datos sólida **fomenta la toma de decisiones basada en datos, la alfabetización de datos y la colaboración entre equipos.**

📌 Importancia en la Organización

- ✅ **Mejor Toma de Decisiones** → Los datos permiten decisiones más informadas y basadas en evidencia, reduciendo riesgos y mejorando la eficiencia.
- ✅ **Mayor Innovación** → Empresas con cultura de datos pueden identificar nuevas oportunidades y mejorar procesos.
- ✅ **Optimización de Procesos** → Facilita la automatización y mejora continua en áreas como ventas, marketing y operaciones.
- ✅ **Cumplimiento Normativo y Seguridad** → Ayuda a garantizar el cumplimiento de regulaciones como GDPR, HIPAA, etc.
- ✅ **Ventaja Competitiva** → Empresas que gestionan bien sus datos tienen una ventaja sobre la competencia.

Cultura de datos: Promoviendo la alfabetización de datos

Alfabetización de Datos: Definición y niveles de madurez

Definición de Alfabetización de Datos

La **alfabetización de datos** es la capacidad de una persona o equipo para **comprender, interpretar y utilizar los datos de manera efectiva** en la toma de decisiones. No se trata solo de habilidades técnicas, sino también de la capacidad de cuestionar, analizar y comunicar hallazgos basados en datos.

Niveles de Madurez en la Alfabetización de Datos

1 Nivel Básico – Conciencia

- ◆ Los colaboradores reconocen el valor de los datos, pero no saben cómo usarlos correctamente.
- ◆ Se apoyan en expertos para interpretar la información.
- ◆ Uso limitado de datos en la toma de decisiones.

2 Nivel Intermedio – Aplicación

- ◆ Los usuarios pueden leer, analizar y entender reportes y dashboards.
- ◆ Pueden identificar datos relevantes para su trabajo.
- ◆ Comienzan a hacer preguntas basadas en datos.

3 Nivel Avanzado – Dominio

- ◆ Los datos son la base para la toma de decisiones en toda la organización.
- ◆ Los empleados pueden combinar y analizar datos por sí mismos.
- ◆ Se promueve el pensamiento crítico basado en datos.

4 Nivel Experto – Innovación

- ◆ Uso avanzado de analítica, machine learning y automatización de datos.
- ◆ Equipos capacitados para descubrir nuevas oportunidades mediante el análisis de datos.
- ◆ Cultura de datos arraigada en todos los niveles de la organización.

Cultura de datos: Promoviendo la alfabetización de datos

El Rol del Gobierno de Datos: Cómo facilita la alfabetización

📌 Gobierno de Datos como Facilitador de la Alfabetización de Datos

El **Gobierno de Datos** es un pilar clave para el desarrollo de la alfabetización de datos en una organización, ya que establece las reglas, procesos y responsabilidades que permiten que los datos sean comprensibles, accesibles y confiables para todos los colaboradores.

📌 Principales Acciones del Gobierno de Datos para Fomentar la Alfabetización

✅ Estandarización y Claridad

- ◆ Define un **lenguaje común de datos** para que toda la organización entienda los términos clave.
- ◆ Establece estándares para la gestión, uso y calidad de los datos.

✅ Accesibilidad y Disponibilidad

- ◆ Asegura que los datos estén disponibles para quienes los necesitan, respetando principios de **seguridad y privacidad**.
- ◆ Implementa un **catálogo de datos** para que los empleados encuentren y entiendan los datos fácilmente.

✅ Calidad y Confiabilidad

- ◆ Garantiza que los datos sean **precisos, completos y actualizados**, evitando la desconfianza en la información.
- ◆ Implementa métricas y controles de calidad de datos.

✅ Formación y Capacitación

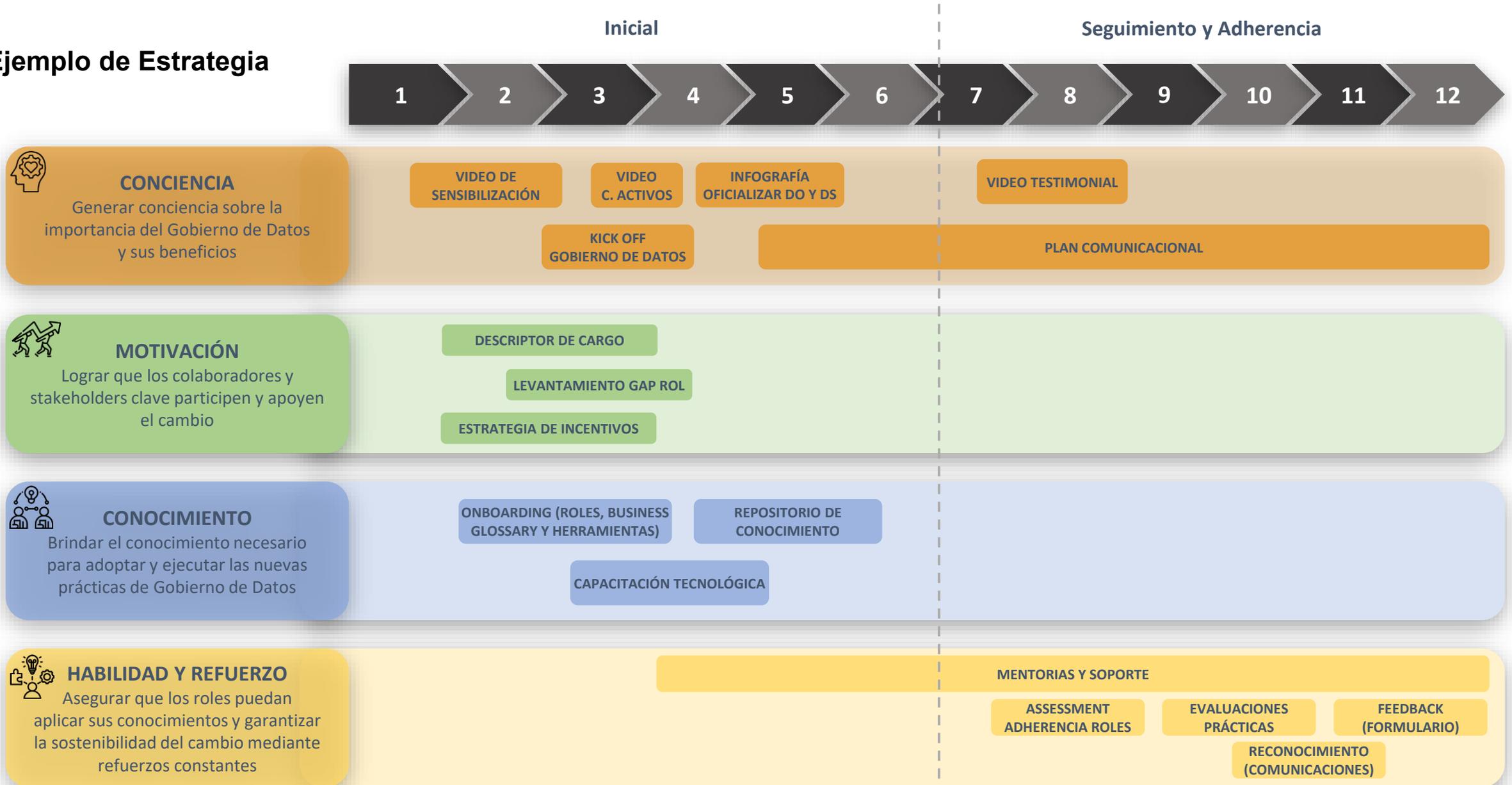
- ◆ Desarrolla programas de alfabetización de datos para distintos niveles de la organización.
- ◆ Crea materiales educativos, workshops y certificaciones internas sobre el uso de datos.

✅ Fomento de la Cultura de Datos

- ◆ Impulsa el uso de datos en la toma de decisiones, promoviendo un **enfoque basado en evidencia**.
- ◆ Incentiva la participación de equipos en iniciativas de mejora de datos.

Cultura de datos: Promoviendo la alfabetización de datos

Ejemplo de Estrategia



Gobierno de Datos

Módulo #11

Evaluación de madurez de Gobierno de Datos

Evaluación de madurez de Gobierno de Datos

Niveles de Madurez de Gobierno de Datos

Nivel: Descripción

- 1. Inicial (Caótico/Reactivo):** No hay una estrategia formal de Gobierno de Datos. Se trabaja de manera reactiva y sin estándares definidos.
- 2. Repetible (Emergente/Ad Hoc):** Existen iniciativas aisladas de Gobierno de Datos, pero sin un marco estructurado. Algunas áreas aplican prácticas básicas.
- 3. Definido (Estandarizado):** Se establecen políticas y procesos formales. El Gobierno de Datos es reconocido y comienza a integrarse en la organización.
- 4. Gestionado (Medible y Controlado):** Se miden los resultados del Gobierno de Datos. Se implementan métricas y mecanismos de mejora continua.
- 5. Optimizado (Maduro/Innovador):** El Gobierno de Datos está completamente integrado en la cultura organizacional y se optimiza continuamente con tecnologías avanzadas.

Dimensión DAMA	Nivel 1 - Inicial	Nivel 2 - Repetible	Nivel 3 - Definido	Nivel 4 - Gestionado	Nivel 5 - Optimizado
1. Data Governance	No hay roles ni políticas de gobierno.	Se definen responsables, pero sin procesos formales.	Existen políticas y comités de gobierno.	Se aplican métricas y mejoras continuas.	El gobierno es parte de la cultura organizacional.
2. Data Architecture	Datos dispersos y sin estructura clara.	Se comienzan a definir modelos de datos.	Arquitectura de datos documentada y estandarizada.	Monitoreo y mejoras basadas en análisis de uso.	Arquitectura optimizada y adaptable a cambios.
3. Data Modeling & Design	Modelos de datos inconsistentes o inexistentes.	Se crean modelos en algunos proyectos clave.	Estándares de modelado definidos y aplicados.	Modelos gestionados centralmente con control de calidad.	Modelos optimizados con automatización y mejores prácticas.
4. Data Storage & Operations	Almacenamiento sin gestión formal.	Se implementan bases de datos con reglas básicas.	Gestión estructurada con políticas de retención.	Monitoreo activo del rendimiento y seguridad.	Infraestructura optimizada y automatizada.
5. Data Security	No hay controles de acceso ni seguridad de datos.	Se implementan controles básicos de acceso.	Se aplican normativas de seguridad en toda la empresa.	Monitoreo continuo y auditorías de seguridad.	Seguridad proactiva con IA y análisis predictivo.
6. Data Integration & Interoperability	Integraciones manuales y ad hoc.	Se establecen conexiones básicas entre sistemas.	Se implementan estándares de integración.	Procesos de integración optimizados y monitorizados.	Interoperabilidad avanzada con automatización y APIs.
7. Document & Content Management	Documentación dispersa y sin control.	Se crean repositorios básicos de documentación.	Gestión centralizada de documentos con estándares.	Control de versiones y gestión automatizada.	Contenido optimizado con AI y machine learning.
8. Reference & Master Data Management	Datos maestros inconsistentes y duplicados.	Se identifican entidades clave sin una gestión formal.	Se implementan soluciones de MDM con reglas de calidad.	Gestión activa de datos maestros con métricas de calidad.	Datos maestros optimizados y gestionados en tiempo real.
9. Data Warehousing & Business Intelligence	Informes manuales y poco confiables.	Se crean dashboards aislados en algunas áreas.	Implementación formal de BI y Data Warehouses.	BI gobernado con métricas y control de calidad.	Inteligencia de datos optimizada con analítica avanzada.
10. Metadata Management	No se gestionan metadatos.	Metadatos documentados parcialmente en algunos sistemas.	Repositorio central de metadatos implementado.	Metadatos gestionados activamente con estándares.	Metadatos enriquecidos con automatización y AI.
11. Data Quality	Datos inconsistentes y sin controles de calidad.	Se identifican problemas de calidad, pero sin soluciones automatizadas.	Se implementan reglas de calidad y validaciones.	Gestión proactiva de calidad con indicadores.	Calidad de datos optimizada con IA y autoservicio.

Gobierno de Datos

Módulo #12

Desarrollo de un Data Catalog

Desarrollo de un Data Catalog

📌 ¿Qué es un Catálogo de Datos?

Un **Catálogo de Datos** es un repositorio centralizado que organiza y documenta los activos de datos de una organización. Su propósito es facilitar la **búsqueda, comprensión y gestión** de los datos disponibles, proporcionando **metadatos estructurados** sobre ellos.

📌 Elementos Claves de un Catálogo de Datos:

- ✅ **Inventario de Datos** → Lista de datasets, tablas, archivos, APIs, etc.
- ✅ **Metadatos Técnicos** → Información sobre estructura, origen, formato y relaciones de los datos.
- ✅ **Metadatos de Negocio** → Definiciones de términos, reglas y contexto de los datos.
- ✅ **Metadatos Operacionales** → Información sobre calidad, uso, accesos y transformaciones de los datos.
- ✅ **Control de Accesos y Seguridad** → Permisos y trazabilidad de los datos para cumplimiento normativo.
- ✅ **Búsqueda y Descubrimiento** → Herramientas para localizar datos según palabras clave, categorías y relaciones.

Desarrollo de un Data Catalog

¿Cómo ayuda un Catálogo de Datos en la Gestión de Datos Maestros?

El Catálogo de Datos es una herramienta clave para gestionar los metadatos dentro de una organización, ya que permite:

- ✓  **Documentación Estructurada** → Centraliza y mantiene actualizada la información sobre los datos.
- ✓  **Descubrimiento de Datos** → Facilita la búsqueda eficiente de datos disponibles dentro de la organización.
- ✓  **Definición Clara y Contexto** → Evita ambigüedades al proporcionar definiciones de negocio claras.
- ✓  **Gobernanza y Cumplimiento** → Ayuda a cumplir normativas como GDPR, HIPAA y regulaciones internas.
- ✓  **Trazabilidad de Datos** → Permite rastrear el linaje de los datos desde su origen hasta su consumo.
- ✓  **Colaboración** → Facilita la comunicación entre equipos de negocio, TI y analistas de datos.
- ✓  **Mejora de la Calidad de Datos** → Proporciona indicadores de calidad y alertas sobre datos inconsistentes.

Desarrollo de un Data Catalog

Ejemplo de Uso del Catálogo de Datos

🚀 Ejemplo 1 – Analista de Negocio

- ◆ Un analista de marketing necesita conocer la definición de “*Cliente Activo*”.
- ◆ Consulta el Catálogo de Datos y encuentra la definición aprobada por el equipo de gobierno de datos.
- ◆ Verifica qué datasets contienen información sobre clientes y quién tiene acceso.

✂ Ejemplo 2 – Ingeniero de Datos

- ◆ Un ingeniero de datos necesita integrar una nueva fuente de información.
- ◆ Usa el Catálogo de Datos para identificar datasets existentes y evitar redundancia.
- ◆ Consulta la documentación de metadatos técnicos para conocer la estructura de los datos.

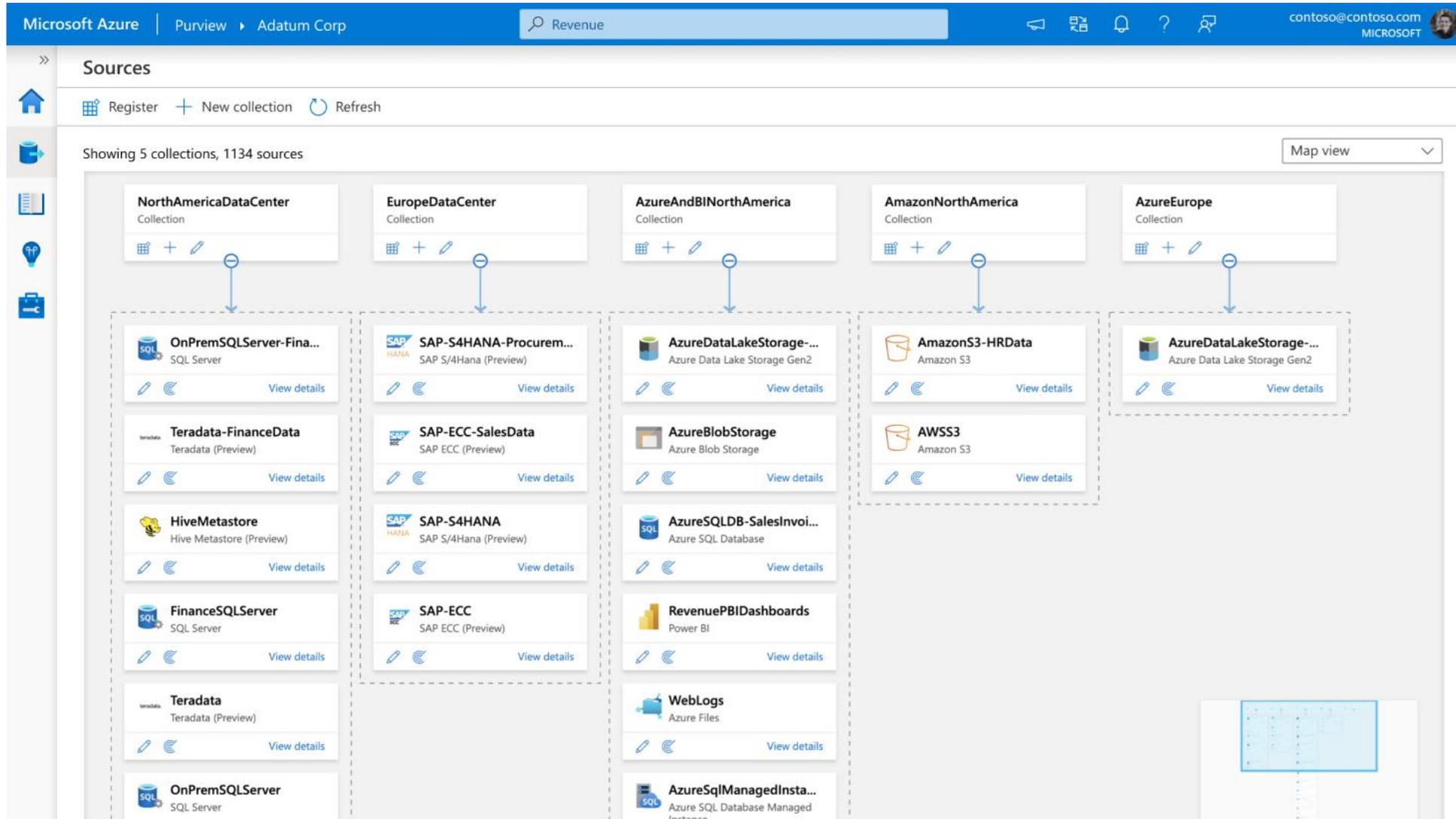
🚀 Ejemplo 3 – Auditoría de Cumplimiento

- ◆ Un auditor revisa qué datos personales se almacenan y cómo se protegen.
- ◆ A través del Catálogo de Datos, obtiene información sobre permisos y accesos.
- ◆ Confirma que los datos cumplen con las regulaciones de privacidad.



Desarrollo de un Data Catalog

Ejemplos de Catálogo de Datos con Diversas Herramientas: Microsoft Purview



Desarrollo de un Data Catalog

Ejemplos de Catálogo de Datos con Diversas Herramientas: Collibra

The screenshot displays the Collibra Data Catalog interface. At the top, there is a navigation bar with a 'Create' button and a search field. Below the navigation bar, the breadcrumb path is 'Business Analysts Community > Schemas > Customer - Prospect Data'. The main header area shows the table name 'Final', its type 'Table', and status 'Candidate'. A 'Business Steward' section identifies 'John Fisher' as the steward. A sidebar on the left contains navigation options: 'Details', 'Add characteristic', 'Data Profiling', 'Sample data', 'Tags', 'Comments', 'Traceability', 'Responsibilities', 'History', and 'Files'. The 'Data Profiling' section is active, showing a table with 17 columns. The table includes columns for Column ID, Name, Data Type, Row Count, Missing Values Count, Number of distinct values, Maximum Value, Minimum Value, Minimum Text Length, Maximum Text Length, and Chart. The data is as follows:

Column P...	Name	Data Type	Row Count	Missing Values Count	Number of distinct values	Maximum Value	Minimum Value	Minimum Text Length	Maximum Text Length	Chart
1	Final > lastname	Text	32,561	0	32,561	Zyskowski	Aamodt	2	13	
2	Final > firstname	Text	32,561	12	5,161	Zulma	Aaron	2	11	
3	Final > Year Born	Whole Number	32,561	0	73	2000.0	1927.0	4	4	
4	Final > age	Whole Number	32,561	0	73	90.0	17.0	2	2	
5	Final > type_employer	Text	32,561	1,836	8	Without-pay	Federal-gov	7	16	
6	Final > fnlwtg	Whole Number	32,561	0	21,648	1484705.0	12285.0	5	7	
7	Final > education	Text	32,561	0	16	Some-college	10th	3	12	
8	Final > education_num	Whole Number	32,561	0	16	16.0	1.0	1	2	
9	Final > marital	Text	32,561	0	7	Widowed	Divorced	7	21	
10	Final > occupation	Text	32,561	1,843	14	Transport-moving	Adm-clerical	5	17	
11	Final > relationship	Text	32,561	0	6	Wife	Husband	4	14	
12	Final > gender	Text	32,561	0	2	Male	Female	4	6	
13	Final > capital_gain	Whole Number	32,561	0	119	99999.0	0.0	1	5	
14	Final > capital_loss	Whole Number	32,561	0	92	4356.0	0.0	1	4	
15	Final > hr_per_week	Whole Number	32,561	0	94	99.0	1.0	1	2	
16	Final > country	Text	32,561	0	4	UK	Australia	2	9	
17	Final > income	Decimal Number	32,561	0	3,123	240000.0	31250.0	5	11	

At the bottom right of the interface, it indicates '17 Relations'.

Desarrollo de un Data Catalog

Ejemplos de Catálogo de Datos con Diversas Herramientas: Databricks (Unity Catalog)

The screenshot shows the Databricks Data Explorer interface. The top navigation bar includes the Databricks logo, a search bar, and user information. The main area is titled "Data Explorer" and shows the path "Catalogs > retail_prod > churn > retail_prod.churn.churn_feature_store". A "Create" button is visible in the top right. Below the path, there are tags, owner information, and a popularity/size indicator. A descriptive note states: "These features are derived from the churn_bronze_customers table in the lakehouse. We created dummy variables for the categorical columns, cleaned up their names, and added a boolean flag for whether the customer churned or not. No aggregations were performed." Below this, there are tabs for "Columns", "Sample Data", "Details", "Permissions", "History", "Lineage", "Insights", and "Quality". The "Columns" tab is active, showing a table with the following columns:

Column	Type	Comment	Tags
user_id	string		
email	string		
creation_date	timestamp		
last_activity_date	timestamp		
age_group	int		
canal	string		
churn	int		

Gobierno de Datos

Módulo #13

Gobierno de Calidad de Datos

Gobierno de Calidad de Datos

¿Qué es la Calidad de Datos?



Es un frente que nos permite medir, monitorear y reportar los niveles de calidad de los datos, para apoyar las tomas de decisiones y así, satisfacer las necesidades del negocio

Gobierno de Calidad de Datos

Objetivos



Mejorar la **precisión** de los datos mediante la reducción de errores de entrada y actualización



Aumentar la **integridad** de los datos asegurando su **completitud** y **coherencia**



Garantizar la **relevancia** de los datos mediante la calidad y priorización de datos críticos



Optimizar la **accesibilidad** de los datos para facilitar su búsqueda, recuperación y uso por parte de los usuarios autorizados

Gobierno de Calidad de Datos

Dimensiones de Calidad de Datos

Formato	Exactitud	Compleitud	Consistencia	Validez	Unicidad	Integridad
						
<p>¿Cumplen registros con los formatos establecidos?</p> <p>Formatos generados por sistemas o reguladoras, etc.</p>	<p>¿Qué tan exactos son nuestros en comparación a realidad?</p> <p>Comparación con fuentes de datos</p>	<p>¿Están presentes todos los datos esperados?</p> <p>Columnas mandatorias y estado</p>	<p>¿Son nuestros consistentes entre diferentes conjuntos?</p> <p>Conjuntos con consistentes y estandarizados</p>	<p>¿Cumplen datos con un conjunto de esperados?</p> <p>Lista de valores posibles, matemáticos y restricciones de dominios</p>	<p>¿Tenemos repetidas?</p> <p>Clave de unicidad que representa a entidad</p>	<p>¿Cuenta el con alguna referencia? ¿se cumple?</p> <p>Integridad referencial</p>

Gobierno de Calidad de Datos

Actividades Clave

Definir
Objetivos de
Calidad de
Datos

Definir Datos de
Alta Calidad

Estrategia de
Calidad

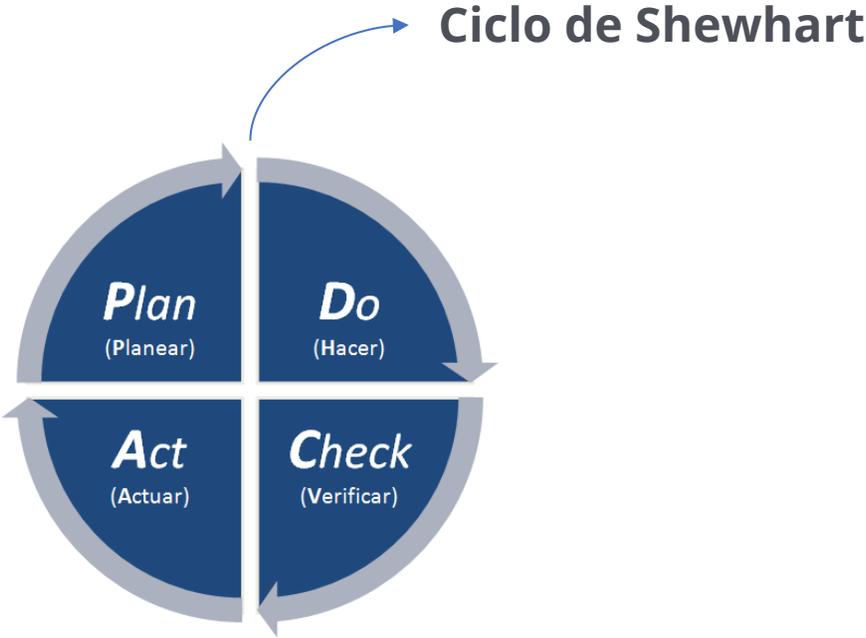
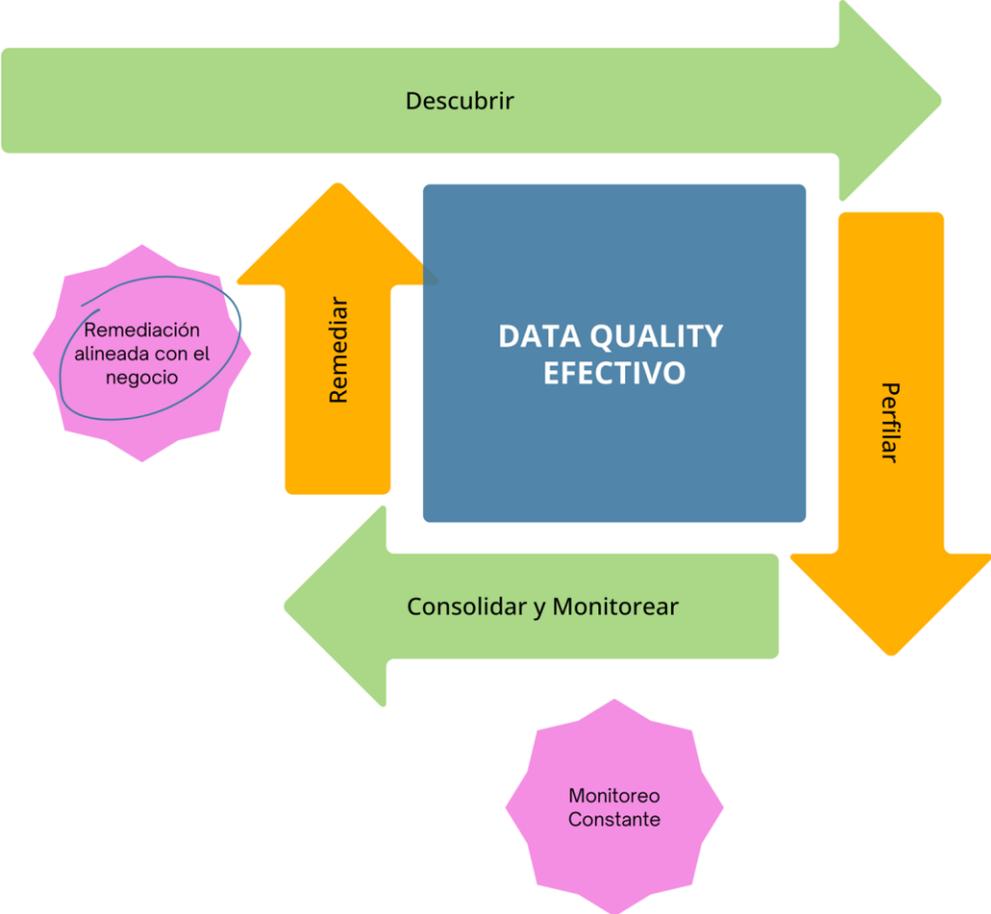
Identificar
Reglas de
Negocio

Evaluación
inicial de
Calidad de
datos

Implementar
una Operación
de Calidad de
Datos

Gobierno de Calidad de Datos

Operación de Calidad de Datos



Gobierno de Calidad de Datos

Reglas de Calidad de Datos

id Regla	id Termino De Negocio	id Dimensión De Calidad	Nombre De La Regla De Calidad	Descripción De La Regla	Medida	Métricas	Nivel de aceptación
RC_0001	TN_0015	DC_03	Compleitud para el campo de nombre central	El campo "Nombre de central" debe encontrarse completo en todos los registros	Se debe contar el total de registros que contienen algún valor no nulo y dividirlo por el total de registros y multiplicar el resultado por 100		>95% excelencia; >85% aceptable;
RC_0002	TN_0142	DC_05	Validez de la medición de los caudales	El valor de la medición de los caudales debe pertenecer al conjunto de los números decimales mayores o iguales a 0	Contar la cantidad de registros mayores a 0, dividirlo por el total de registros y multiplicar por 100		=100% excelencia; <100% no aceptable;

Gobierno de Datos

Módulo #14

Seguridad y Privacidad de Datos

Seguridad y Privacidad de Datos

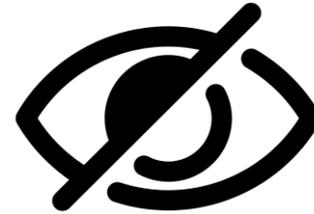
📌 ¿Por qué es importante la seguridad y privacidad de los datos?

- Los datos son un activo clave en las organizaciones.
- Un mal manejo puede llevar a fraudes, sanciones legales y pérdida de confianza.
- Regulaciones como **GDPR, HIPAA, CCPA** exigen protección de datos.



Seguridad y Privacidad de Datos

Diferencias entre Seguridad y Privacidad de Datos



Concepto	Seguridad de Datos	Privacidad de Datos
Definición	Protección contra accesos no autorizados, corrupción o pérdida.	Control sobre el uso y acceso a la información personal.
Objetivo	Evitar filtraciones, ataques y pérdida de datos.	Asegurar que los datos se usen según políticas y regulaciones.
Ejemplo	Cifrado de datos, autenticación de usuarios.	Consentimiento del usuario, derecho al olvido.

Seguridad y Privacidad de Datos



Principios Claves de Seguridad de Datos

- ◆ **Confidencialidad:** Solo las personas autorizadas deben acceder a los datos.
- ◆ **Integridad:** Los datos no deben ser alterados o dañados sin autorización.
- ◆ **Disponibilidad:** Los datos deben estar accesibles cuando sean necesarios.

💡 **Ejemplo:** Un banco debe garantizar que sus clientes puedan acceder a sus cuentas (disponibilidad), pero solo ellos y los empleados autorizados pueden ver la información (confidencialidad), y los registros de transacciones no deben ser alterados sin control (integridad).

Seguridad y Privacidad de Datos



Principios Claves de Privacidad de Datos

- ◆ **Minimización de Datos:** Solo recolectar los datos estrictamente necesarios.
 - ◆ **Consentimiento:** Los usuarios deben saber cómo se usarán sus datos y aprobarlo.
 - ◆ **Derecho al Olvido:** Los usuarios pueden solicitar que se eliminen sus datos.
 - ◆ **Transparencia:** Explicar claramente cómo se manejan los datos.
- 💡 **Ejemplo:** Un e-commerce que solicita el teléfono del usuario debe justificar por qué lo necesita y darle la opción de eliminar su cuenta si lo desea.



Seguridad y Privacidad de Datos

Mejores Prácticas para proteger los Datos Sensibles

- ✓ **Cifrado de Datos** → Usar cifrado en tránsito y en reposo.
 - ✓ **Gestión de Accesos** → Aplicar control de acceso basado en roles (RBAC).
 - ✓ **Autenticación Multifactor (MFA)** → Evitar accesos no autorizados.
 - ✓ **Monitoreo y Auditoría** → Revisar logs y detectar actividades sospechosas.
 - ✓ **Clasificación de Datos** → Identificar datos sensibles y aplicar protecciones adecuadas.
 - ✓ **Formación y Concienciación** → Capacitar a empleados sobre seguridad de datos.
- 💡 **Ejemplo:** Un hospital implementa cifrado de datos para registros médicos y usa autenticación multifactor para acceso del personal.



Seguridad y Privacidad de Datos

Brecha de Datos y sus Consecuencias

📌 **Ejemplo real:** En 2017, **Equifax** sufrió una filtración de datos que expuso información personal de 147 millones de personas.

📌 **Causas:** Falta de actualización de sistemas y gestión deficiente de accesos.

📌 **Consecuencias:**

- Multas de **\$700 millones de dólares.**
- Pérdida de confianza de clientes y caída en la reputación.
- Requerimiento de mejoras en su seguridad de datos.

💡 **Lección:** Una mala gestión de seguridad puede costar millones y afectar la credibilidad de la organización.

Seguridad y Privacidad de Datos

Resumen de Puntos Clave

- ✓ La **seguridad de datos** protege contra accesos no autorizados y alteraciones.
- ✓ La **privacidad de datos** asegura que los datos se usen correctamente.
- ✓ Existen múltiples amenazas, pero aplicando buenas prácticas podemos mitigarlas.
- ✓ Cumplir con regulaciones es fundamental para evitar sanciones y fortalecer la confianza.
- ✓ La seguridad de datos no es solo un **problema tecnológico**, sino también un **tema de cultura organizacional**.



Gobierno de Datos

Módulo #15

Data Lineage: Rastreando el ciclo de vida de los datos

Data Lineage: Rastreando el ciclo de vida de los datos

¿Qué es Data Lineage y por qué es importante?

- **Data Lineage** es el **rastreo del ciclo de vida de los datos** desde su origen hasta su consumo.
- Permite entender **cómo los datos fluyen a través de los sistemas**, facilitando auditorías, calidad de datos y cumplimiento normativo.
- Es clave en **Data Governance, Business Intelligence y Big Data**.



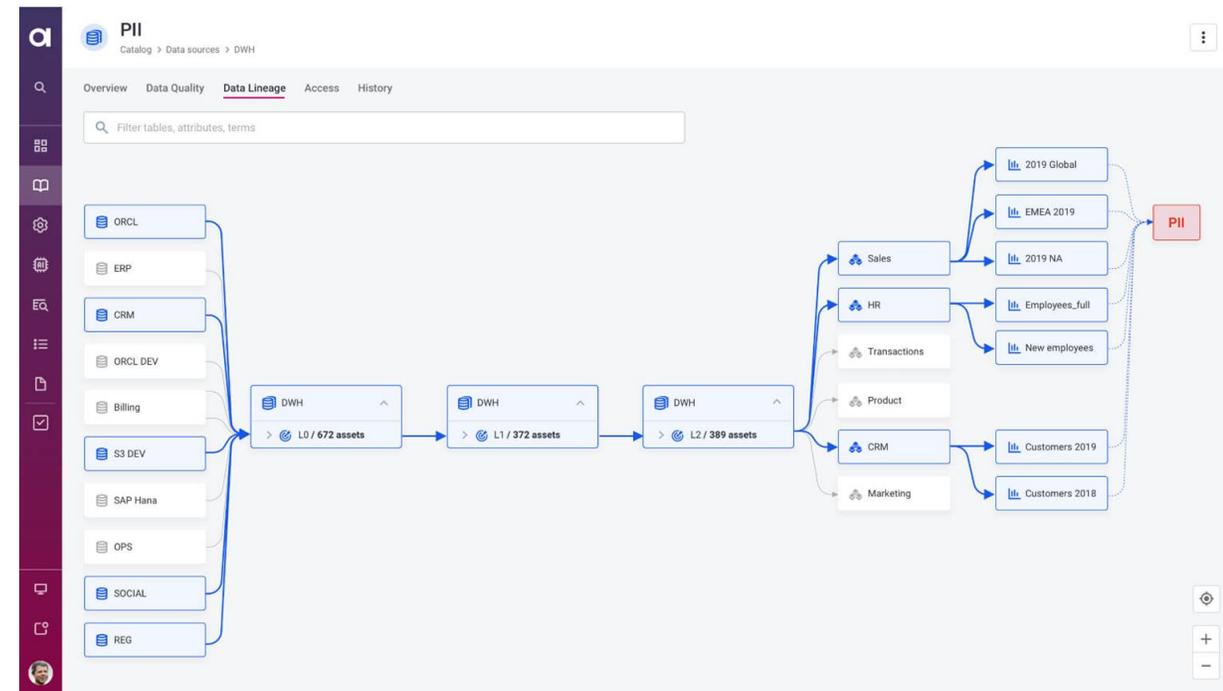
Data Lineage: Rastreamo el ciclo de vida de los datos

Definición

El **Data Lineage** es la trazabilidad del recorrido de los datos, incluyendo:

- ✓ **Origen:** De dónde provienen los datos (fuentes, sistemas, bases de datos).
- ✓ **Transformaciones:** Cómo son procesados y modificados a lo largo del tiempo.
- ✓ **Movimientos:** Cómo se transfieren entre sistemas y aplicaciones.
- ✓ **Destino:** Quién consume los datos y cómo se utilizan.

💡 **Ejemplo:** Un reporte financiero se genera a partir de múltiples bases de datos. El **Data Lineage** permite rastrear cómo se obtienen, transforman y combinan estos datos antes de llegar al informe final.

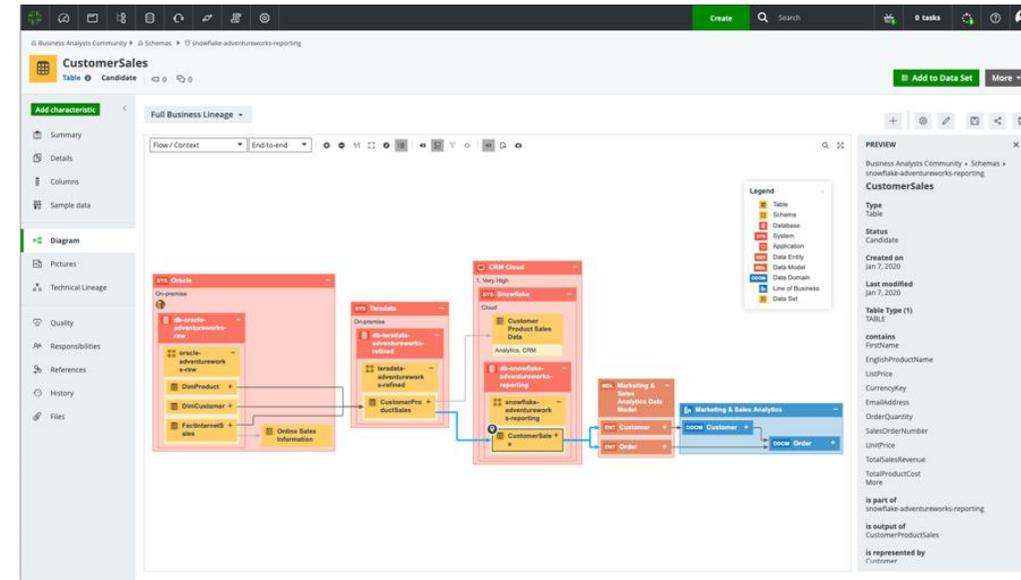


Data Lineage: Rastreando el ciclo de vida de los datos

Beneficios del Data Lineage

- ✓ **Mejor Comprensión de los Datos** → Permite visualizar cómo los datos fluyen en la organización.
- ✓ **Cumplimiento Normativo** → Ayuda a cumplir regulaciones como **GDPR**, **HIPAA**, **SOX**.
- ✓ **Optimización de Procesos** → Mejora la eficiencia en la integración de datos.
- ✓ **Mayor Calidad de Datos** → Permite detectar errores y evitar datos inconsistentes.
- ✓ **Facilita la Resolución de Problemas** → Reduce tiempos de respuesta ante incidencias.
- ✓ **Seguridad y Gobernanza** → Define quién puede acceder y modificar los datos.

💡 **Ejemplo:** Una empresa debe demostrar a una auditoría de GDPR cómo procesa y almacena datos personales. Con Data Lineage, puede mapear y documentar cada paso de los datos.



Data Lineage: Rastreando el ciclo de vida de los datos

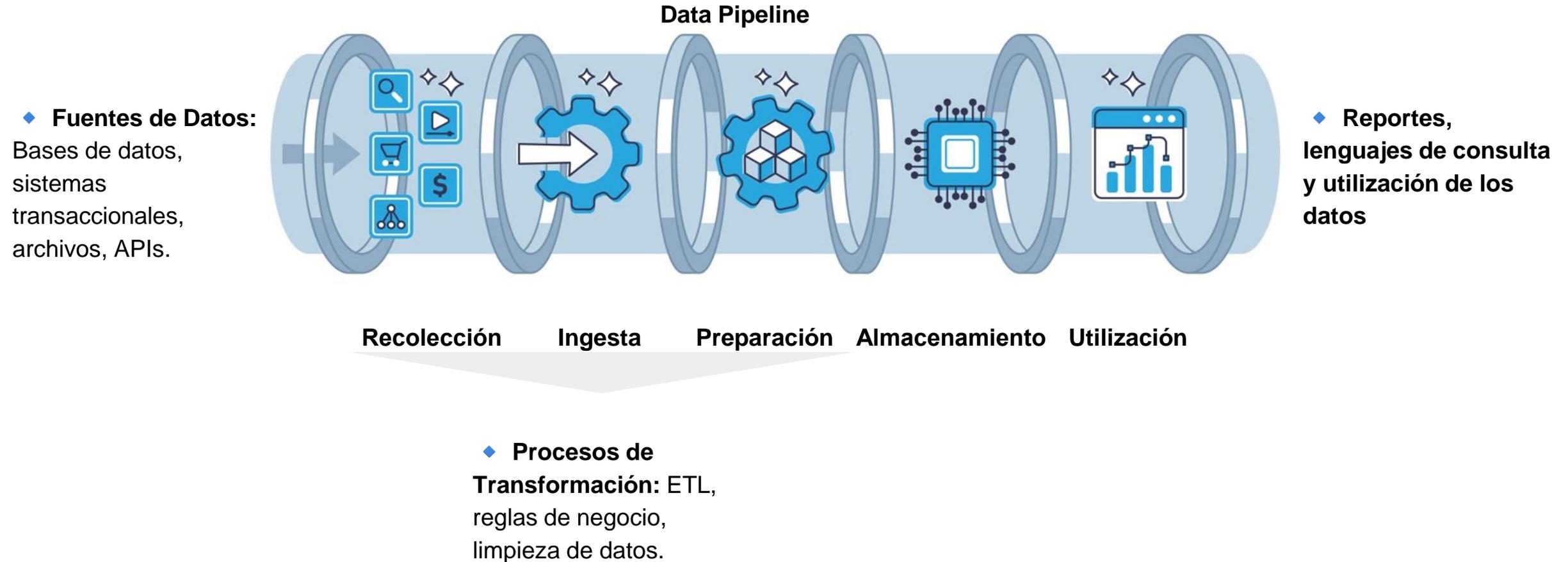
Tipos de Data Lineage

- ✦ **1 Lineage Físico** → Se centra en la trazabilidad técnica de los datos (ETL, bases de datos, APIs).
 - ✦ **2 Lineage Lógico** → Mapea el flujo de datos a nivel conceptual y de negocio.
 - ✦ **3 Lineage Activo** → Se actualiza en tiempo real con herramientas automatizadas.
 - ✦ **4 Lineage Pasivo** → Se obtiene a partir de documentación y análisis manual.
- 💡 **Ejemplo:** Un Data Lineage **físico** muestra cómo los datos pasan de un sistema de CRM a un Data Warehouse, mientras que un **lógico** describe el significado de cada campo y su relación con términos de negocio.



Data Lineage: Rastreando el ciclo de vida de los datos

Componentes del Data Lineage



Data Lineage: Rastreamento el ciclo de vida de los datos

Herramientas para Implementar el Data Lineage

📌 Soluciones Populares:

- ✅ **Informatica EDC** → Para grandes volúmenes de datos y gobernanza avanzada.
 - ✅ **Collibra Data Lineage** → Trazabilidad automatizada y cumplimiento normativo.
 - ✅ **Apache Atlas** → Open-source para ecosistemas Big Data.
 - ✅ **Microsoft Purview** → Solución en la nube para la trazabilidad de datos en Azure.
- 💡 **Ejemplo:** Una empresa de retail usa **Microsoft Purview** para mapear cómo los datos de clientes fluyen entre su CRM y plataformas de analítica.



Informatica



Collibra



Apache Atlas



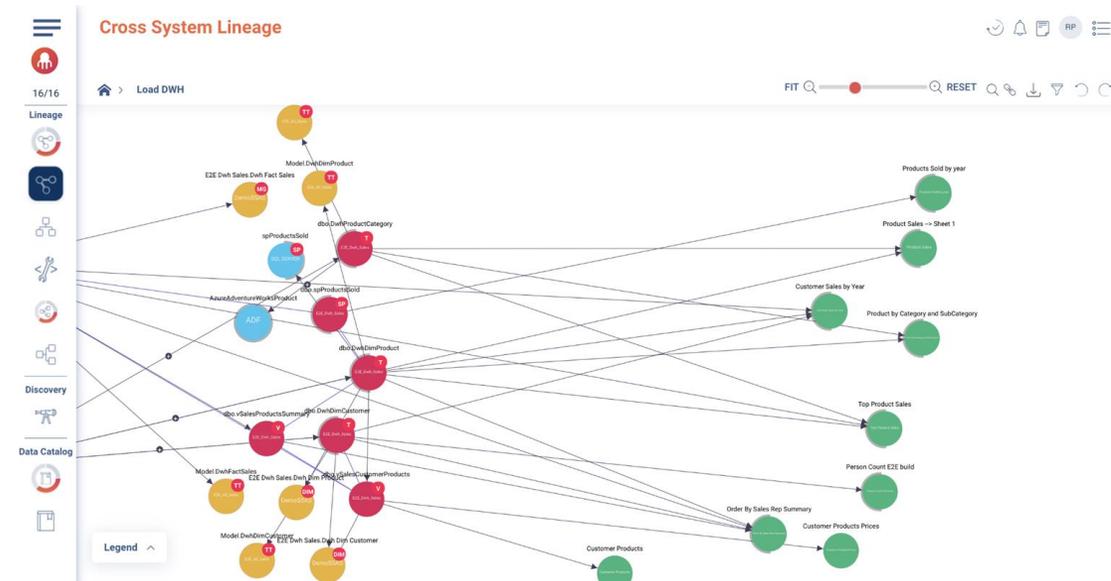
Azure Purview

Data Lineage: Rastreando el ciclo de vida de los datos

Caso de Uso: Data Lineage en una Organización

📌 Ejemplo: Data Lineage en un Banco

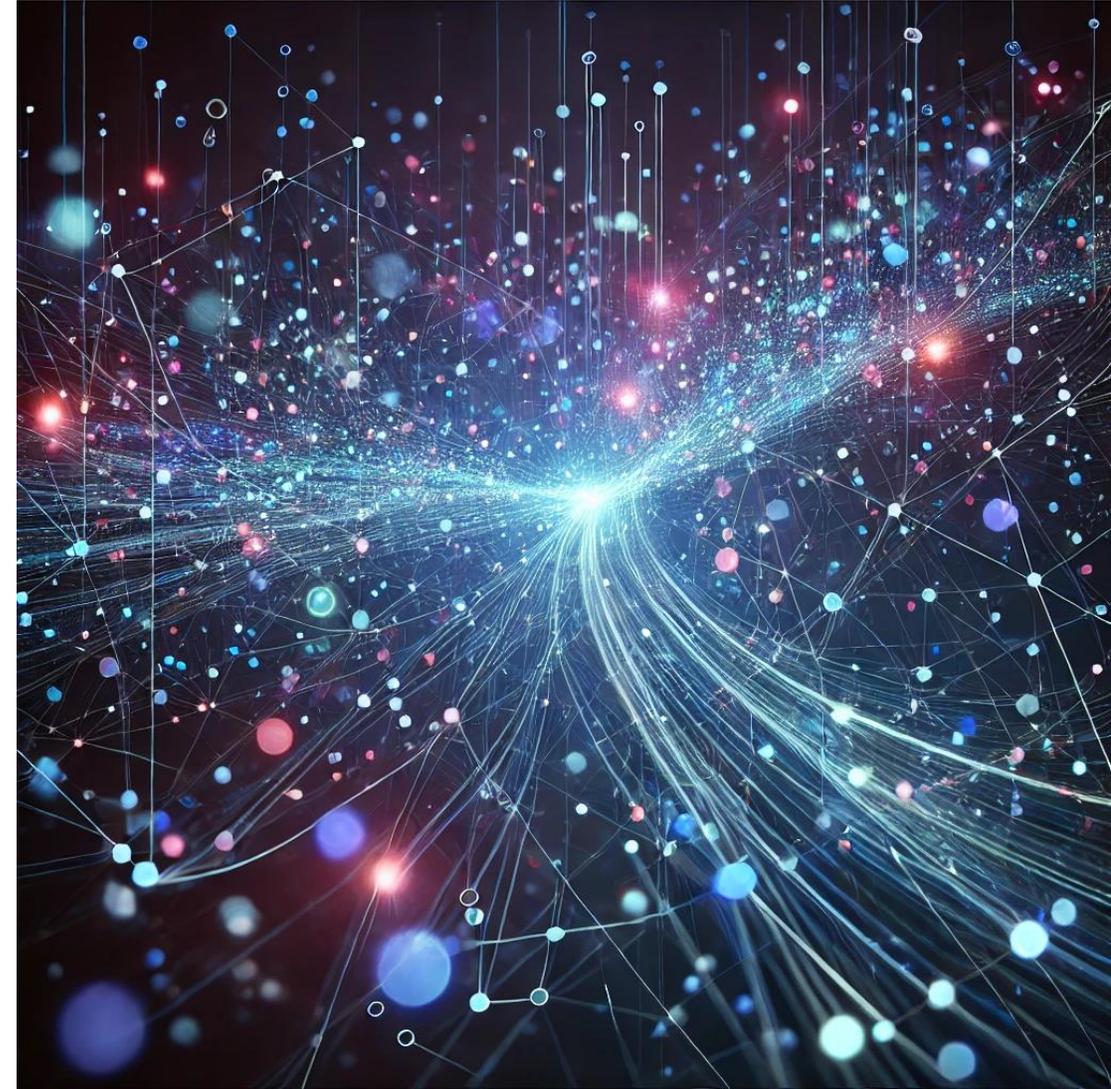
- ✅ Un banco usa **Data Lineage** para rastrear el origen de los datos usados en el cálculo del riesgo crediticio.
- ✅ Si un error afecta la evaluación de un cliente, pueden revisar todo el flujo de datos hasta encontrar la causa.
- ✅ Esto garantiza que los procesos sean **transparentes, auditables y confiables**.



Data Lineage: Rastreando el ciclo de vida de los datos

Resumen de Puntos Clave

- ✓ El **Data Lineage** permite rastrear el ciclo de vida de los datos desde su origen hasta su consumo.
- ✓ Facilita **cumplimiento normativo, calidad de datos y gobernanza**.
- ✓ Hay diferentes **tipos de Data Lineage** (físico, lógico, activo, pasivo).
- ✓ Se pueden usar **herramientas especializadas** para automatizar el seguimiento.
- ✓ Implementarlo con éxito requiere una **estrategia clara y cultura de datos**.



Gobierno de Datos

Módulo #16

Gobierno de Datos en Arquitecturas Modernas

Gobierno de Datos en Arquitecturas Modernas

Aplicación en Data Mesh, Data Fabric y sistemas distribuidos

✦ ¿Por qué es clave el Gobierno de Datos en arquitecturas modernas?

- Las empresas adoptan modelos **descentralizados** como **Data Mesh** y **Data Fabric** para escalar sus datos.
- En estas arquitecturas, la **gestión de datos se vuelve más compleja**, lo que exige un gobierno de datos sólido.
- Un Gobierno de Datos bien implementado garantiza **calidad, seguridad, interoperabilidad y cumplimiento normativo**.



Gobierno de Datos en Arquitecturas Modernas

Definición

Definición

El **Gobierno de Datos en arquitecturas modernas** es el conjunto de **políticas, procesos y tecnologías** que garantizan el control de los datos en entornos distribuidos.

Objetivos Claves:

- **Asegurar la calidad y confiabilidad de los datos** en entornos descentralizados.
- **Garantizar la seguridad y cumplimiento normativo** (GDPR, HIPAA, CCPA).
- **Facilitar la interoperabilidad** entre sistemas y dominios de datos.
- **Definir ownership y accountability** en arquitecturas distribuidas.

Gobierno de Datos en Arquitecturas Modernas

Comparación de Arquitecturas: Data Mesh vs Data Fabric

Característica	Data Mesh	Data Fabric
Enfoque	Descentralizado	Híbrido (centralizado + descentralizado)
Gestión de Datos	Datos gobernados por dominios	Automatización de integración y acceso
Gobierno de Datos	Autonomía por dominio con reglas comunes	Control centralizado con visibilidad global
Automatización	Baja (requiere procesos manuales)	Alta (metadatos y AI para optimización)
Casos de Uso	Grandes empresas con equipos distribuidos	Empresas que buscan integración fluida de datos

💡 **Ejemplo:** Una empresa de retail global usa **Data Mesh** para que cada región gestione sus propios datos, mientras que una aseguradora usa **Data Fabric** para conectar datos de múltiples fuentes de manera centralizada.

Gobierno de Datos en Arquitecturas Modernas

Retos del Gobierno de Datos en Data Mesh y Data Fabric

- ⚠ **Descentralización y falta de control** → Difícil mantener coherencia en los datos.
 - ⚠ **Calidad de Datos variable** → Diferentes equipos pueden aplicar estándares inconsistentes.
 - ⚠ **Seguridad y Cumplimiento** → Dificultad para aplicar reglas uniformes de acceso y protección.
 - ⚠ **Interoperabilidad** → Desafíos en la integración y compatibilidad de datos entre sistemas.
- 💡 **Solución: Establecer principios de gobierno claros** en cada modelo.

Gobierno de Datos en Arquitecturas Modernas

Principios de Gobierno de Datos en Arquitecturas Modernas

- ✓ **Data as a Product** → En **Data Mesh**, cada dominio es responsable de la calidad de sus datos.
- ✓ **Interoperabilidad** → En **Data Fabric**, se deben definir estándares para la integración automática.
- ✓ **Metadatos Activos** → Uso de catálogos de datos, linaje y trazabilidad para gestión eficiente.
- ✓ **Acceso Seguro y Federado** → Implementar **Zero Trust Security** y políticas de privacidad granular.
- ✓ **Monitoreo Continuo** → Aplicar herramientas de observabilidad para medir calidad y uso de datos.

💡 **Ejemplo:** Un banco que usa **Data Fabric** debe aplicar gobierno para garantizar que todas sus fuentes de datos cumplan con los mismos estándares de calidad y seguridad.

Gobierno de Datos en Arquitecturas Modernas

Herramientas para Gobierno de Datos en Arquitecturas Modernas

Data Mesh:

- ✓ **Collibra** → Gobernanza descentralizada con workflows por dominio.
- ✓ **Alation** → Catálogo de datos con trazabilidad para equipos autónomos.

Data Fabric:

- ✓ **Microsoft Purview** → Gobierno de datos en arquitecturas híbridas.
- ✓ **Informatica Axon** → Automatización de calidad y cumplimiento normativo.

Sistemas Distribuidos:

- ✓ **Apache Atlas** → Open-source para linaje y catalogación de datos.
- ✓ **AWS Lake Formation** → Gobernanza en entornos de lago de datos.

 **Ejemplo:** Una empresa de manufactura usa **Collibra** para descentralizar su gobierno en Data Mesh, mientras que una fintech usa **Microsoft Purview** para gobernar su Data Fabric en la nube.

Gobierno de Datos en Arquitecturas Modernas

Caso de estudio: Implementación de Gobierno de Datos en Data Mesh

✦ **Contexto:** Una empresa global de e-commerce con equipos de datos en diferentes países.

◆ **Problema:**

- Los equipos manejaban sus datos sin estándares comunes.
- Había problemas de calidad e inconsistencias en los reportes.

◆ **Solución:**

- Se implementó **un marco de gobierno descentralizado** basado en Data Mesh.
- Cada dominio definió reglas claras de calidad y acceso a sus datos.
- Se usó **Collibra** para gestionar políticas de gobierno de manera federada.

◆ **Resultados:**

- ✓ Reducción de errores en datos en un **30%**.
- ✓ Mejor eficiencia en la integración de datos entre regiones.
- ✓ Mayor cumplimiento normativo en la gestión de datos personales.

Gobierno de Datos en Arquitecturas Modernas

Conclusión

Resumen de Puntos Clave:

- ✓ **Data Mesh y Data Fabric requieren enfoques diferentes de gobierno de datos.**
- ✓ **Es clave definir ownership, calidad, seguridad e interoperabilidad.**
- ✓ **Las herramientas adecuadas permiten automatizar y escalar el gobierno de datos.**
- ✓ **El gobierno debe adaptarse a la descentralización sin perder control.**

Gobierno de Datos

Módulo #17

Gobierno de Datos para la Inteligencia Artificial y Analítica Avanzada

Gobierno de Datos para la Inteligencia Artificial y Analítica Avanzada

Introducción

¿Por qué es clave el Gobierno de Datos para la IA y Analítica Avanzada?

- Los modelos de **Machine Learning (ML)** y **Analítica Avanzada** dependen de **datos confiables y bien gobernados**.
- Datos sin gobernanza pueden generar **sesgos, resultados erróneos y riesgos de seguridad**.
- Un marco sólido de **Gobierno de Datos** garantiza **calidad, seguridad y cumplimiento normativo** en los modelos predictivos.

Gobierno de Datos para la Inteligencia Artificial y Analítica Avanzada

¿Cómo impacta el Gobierno de Datos en la IA y Analítica Avanzada?

✦ **El Gobierno de Datos impacta en:**

- ✓ **Calidad de Datos:** Los modelos de IA solo pueden ser tan buenos como los datos que los entrenan.
- ✓ **Transparencia y Explicabilidad:** Facilita la trazabilidad de las fuentes de datos y sus transformaciones.
- ✓ **Cumplimiento Normativo:** Permite asegurar que los datos usados en modelos cumplan regulaciones como **GDPR, HIPAA, CCPA**.
- ✓ **Interoperabilidad:** Facilita la integración de datos de diferentes fuentes para mejorar modelos predictivos.
- ✓ **Gestión de Sesgos:** Evita que modelos aprendan patrones discriminatorios o erróneos.

💡 **Ejemplo:** Un banco usa IA para evaluar riesgos crediticios. Sin gobierno de datos, el modelo podría estar entrenado con datos sesgados, afectando decisiones financieras.

Gobierno de Datos para la Inteligencia Artificial y Analítica Avanzada

Desafíos en la Gestión de Datos para IA y Analítica

- ⚠ **Falta de calidad y consistencia** → Datos sucios o incompletos generan modelos poco confiables.
- ⚠ **Sesgo en los datos** → Datos históricos con prejuicios pueden generar discriminación en modelos predictivos.
- ⚠ **Falta de trazabilidad y linaje** → Sin un rastro claro de cómo se procesaron los datos, es difícil confiar en los resultados.
- ⚠ **Cumplimiento normativo** → Modelos que usan datos personales deben cumplir con regulaciones.
- ⚠ **Seguridad y privacidad** → Uso indebido de datos puede derivar en problemas legales y pérdida de confianza.

💡 **Solución: Implementar un marco de Gobierno de Datos adaptado a IA.**

Gobierno de Datos para la Inteligencia Artificial y Analítica Avanzada

Principios de Gobierno de Datos para IA y Analítica

- 📌 **1 Calidad de Datos** → Implementar reglas de validación, limpieza y estandarización.
- 📌 **2 Data Lineage y Transparencia** → Trazabilidad completa de los datos desde su origen hasta su uso en modelos.
- 📌 **3 Seguridad y Privacidad** → Aplicar controles de acceso y anonimización de datos sensibles.
- 📌 **4 Cumplimiento Regulatorio** → Asegurar que los datos usados en modelos respeten normativas globales.
- 📌 **5 Gestión de Sesgos y Equidad** → Monitorear y mitigar posibles sesgos en los datos.

💡 **Ejemplo:** Una empresa de salud debe asegurarse de que sus modelos de predicción de enfermedades cumplan con **HIPAA** y no discriminen a ciertos grupos de pacientes.

Gobierno de Datos para la Inteligencia Artificial y Analítica Avanzada

Estrategias para Implementar Gobierno de Datos en IA

- ✓ **Establecer estándares de calidad y limpieza de datos** antes de usarlos en modelos.
- ✓ **Implementar trazabilidad automática de datos (Data Lineage)** para mejorar explicabilidad.
- ✓ **Definir controles de acceso basados en roles** para proteger datos sensibles.
- ✓ **Usar técnicas de anonimización y enmascaramiento de datos** para cumplimiento normativo.
- ✓ **Crear políticas de auditoría y monitoreo** para revisar la equidad y sesgos de modelos.

💡 **Ejemplo:** Un retailer usa **Data Lineage** para asegurar que los datos de ventas usados en modelos de predicción de demanda provengan de fuentes confiables.

Gobierno de Datos para la Inteligencia Artificial y Analítica Avanzada

Caso de Estudio: Implementación de un Gobierno de Datos en IA

📌 **Contexto:** Un hospital usa IA para predecir el riesgo de enfermedades en pacientes.

◆ **Problema:**

- Falta de trazabilidad en los datos usados para entrenar modelos.
- Preocupación sobre privacidad y cumplimiento normativo (HIPAA).

◆ **Solución:**

- Se implementó **Microsoft Purview** para rastrear fuentes de datos y aplicar control de accesos.
- Se usó **Privitar** para anonimizar datos y entrenar modelos sin exponer información personal.
- Se establecieron **controles de calidad y monitoreo** para mejorar confiabilidad del modelo.

◆ **Resultados:**

- ✓ Modelos de IA más confiables y con menor riesgo de sesgo.
- ✓ Cumplimiento total de regulaciones de privacidad.
- ✓ Mayor aceptación de modelos predictivos por parte del equipo médico.

Gobierno de Datos para la Inteligencia Artificial y Analítica Avanzada

Conclusiones

Resumen de Puntos Clave:

- ✓ La calidad, transparencia y seguridad de datos son clave para IA confiable.
- ✓ **Data Lineage y catálogos de datos** ayudan a mejorar la trazabilidad.
- ✓ Herramientas de **seguridad y anonimización** protegen los datos sensibles.
- ✓ El cumplimiento normativo es esencial para evitar problemas legales.
- ✓ Una estrategia de gobierno sólida mejora la adopción y confianza en la IA.

Gobierno de Datos

Módulo #18

Medición del Éxito de Gobierno de Datos

Medición del Éxito de Gobierno de Datos

Introducción



¿Por qué medir el éxito del Gobierno de Datos?

- Gobierno de Datos no es solo sobre políticas y procesos, sino sobre generar **valor medible**.
- Sin métricas claras, es difícil demostrar el impacto en el negocio.
- Medir el éxito permite **ajustar estrategias, optimizar procesos y mejorar la toma de**

decisiones.

Medición del Éxito de Gobierno de Datos

¿Qué son los KPIs y Métricas en Gobierno de Datos?

📌 Definición:

✅ **KPIs (Key Performance Indicators):** Indicadores clave de rendimiento que miden el éxito del Gobierno de Datos.

✅ **Métricas:** Datos específicos que ayudan a evaluar el desempeño de los procesos de gobierno.

💡 Ejemplo:

📊 Un KPI puede ser el **porcentaje de datos con calidad validada**.

📌 Una métrica asociada puede ser el **número de registros duplicados eliminados mensualmente**.



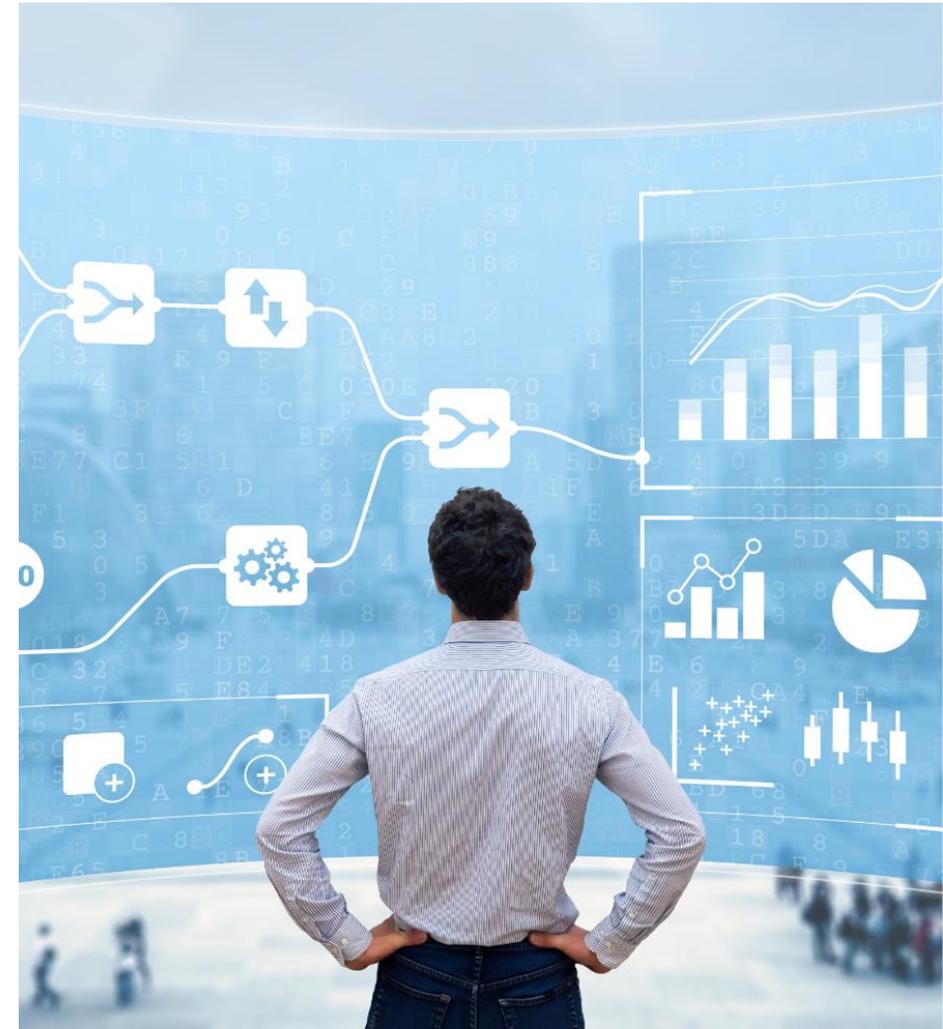
Medición del Éxito de Gobierno de Datos

Categorías de KPIs en Gobierno de Datos

📌 Los KPIs en Gobierno de Datos se pueden agrupar en:

- ✅ **1 Calidad de Datos:** Evalúa la exactitud, completitud y consistencia de los datos.
- ✅ **2 Seguridad y Cumplimiento:** Mide el cumplimiento de normativas y la protección de datos.
- ✅ **3 Accesibilidad y Uso:** Analiza la disponibilidad y usabilidad de los datos.
- ✅ **4 Eficiencia Operacional:** Evalúa el impacto del Gobierno de Datos en procesos y costos.
- ✅ **5 Adopción y Cultura de Datos:** Mide la aceptación del gobierno dentro de la organización.

💡 **Ejemplo:** En una empresa de seguros, un KPI de calidad de datos podría ser el porcentaje de pólizas sin errores en la base de datos.



Medición del Éxito de Gobierno de Datos

KPIs Claves para Calidad de Datos

📌 Indicadores de Calidad de Datos:

- ✅ **Precisión:** % de datos que reflejan información real.
- ✅ **Completitud:** % de datos sin valores nulos o faltantes.
- ✅ **Consistencia:** % de datos alineados entre sistemas.
- ✅ **Unicidad:** Número de registros duplicados eliminados.
- ✅ **Actualización:** % de datos que cumplen con los tiempos de actualización definidos.

💡 Ejemplo:

📊 *“El 95% de los datos de clientes deben estar completos y sin duplicados en la base de datos CRM.”*



Medición del Éxito de Gobierno de Datos

KPIs Claves para Seguridad y Cumplimiento

📌 Indicadores de Seguridad y Cumplimiento:

- ✅ **Incidentes de seguridad de datos:** Número de violaciones de seguridad reportadas.
- ✅ **Accesos no autorizados detectados:** % de accesos bloqueados por políticas de seguridad.
- ✅ **Cumplimiento de regulaciones:** Nivel de cumplimiento con normativas como **GDPR**, **HIPAA**, **CCPA**.
- ✅ **Tiempo de respuesta ante auditorías:** Días promedio en responder a auditorías de datos.

💡 Ejemplo:

📊 “Reducir los incidentes de acceso no autorizado en un 30% en el próximo año.”



Medición del Éxito de Gobierno de Datos

KPIs Claves para Eficiencia Operacional

📌 Indicadores de Eficiencia Operacional:

- ✅ Tiempo de procesamiento de datos en pipelines ETL.
- ✅ Costo de almacenamiento y procesamiento de datos.
- ✅ Número de procesos optimizados por mejoras en Gobierno de Datos.
- ✅ Impacto financiero de la mejora en calidad de datos (reducción de costos por errores).

💡 Ejemplo:

📊 “Reducir en un 20% los costos operativos asociados a corrección de errores en datos.”



Medición del Éxito de Gobierno de Datos

KPIs Claves para Adopción y Cultura de Datos

📌 Indicadores de Adopción de Gobierno de Datos:

- ✅ Número de empleados capacitados en gobierno de datos.
- ✅ Porcentaje de cumplimiento en la documentación de definiciones de datos.
- ✅ Nivel de satisfacción de los usuarios con la calidad de datos.
- ✅ Número de iniciativas de datos alineadas con el marco de Gobierno de Datos.

💡 Ejemplo:

📊 “Capacitar al 80% del equipo de analítica en el uso del catálogo de datos en los próximos 6 meses.”



Medición del Éxito de Gobierno de Datos

Caso de Estudio: Medición del Éxito en una Empresa

📌 Contexto:

Una empresa de retail implementa Gobierno de Datos y necesita medir su impacto.

◆ KPIs seleccionados:

- ✓ Reducción del 30% en registros duplicados en CRM.
- ✓ Disminución del tiempo de búsqueda de datos de 6 horas a 2 horas.
- ✓ Cumplimiento del 100% de auditorías regulatorias sin sanciones.

◆ Resultados:

- ✓ Mayor confianza en los datos para decisiones estratégicas.
- ✓ Ahorro de costos en limpieza de datos.
- ✓ Cumplimiento normativo sin multas.

Medición del Éxito de Gobierno de Datos

Conclusión

📌 Resumen de Puntos Clave:

- ✅ Medir el éxito del Gobierno de Datos es clave para demostrar su impacto.
- ✅ Existen diferentes **KPIs y métricas** según calidad, seguridad, accesibilidad, eficiencia y adopción.
- ✅ La elección de los KPIs debe alinearse con los **objetivos del negocio**.
- ✅ La mejora en Gobierno de Datos se traduce en **mayor confianza, eficiencia y cumplimiento normativo**.



Gobierno de Datos

Módulo #18

Medición del Éxito de Gobierno de Datos