

# Catalogación de datos para un buen gobierno de datos: cinco funcionalidades esenciales

Cómo impulsar los programas empresariales de gobierno de datos



# La oportunidad

En la actual economía basada en las innovaciones, su capacidad de aprovechar con eficacia los datos será lo que determine si realmente logra el éxito o el fracaso.

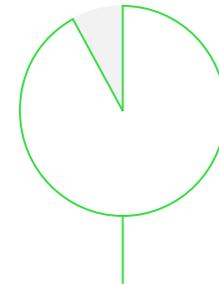
Podemos tomar como ejemplo el caso de un fabricante del sector automotriz que busque atraer al cliente con métodos nuevos, y por tanto, se plantee mejorar los resultados de sus análisis y comprenda que necesita un sistema de gobierno completo de todos sus datos empresariales. Una empresa global de servicios financieros, que busca lanzar nuevos servicios más rápidos, automatiza su programa de gobierno de datos para ampliar el acceso de los empleados a información clave.

Como muestran estos ejemplos, las empresas exitosas se dan cuenta de que los datos gobernados y confiables y como estos representan un activo comercial clave. La inteligencia artificial encontrara nuevas formas de aprovechar los datos, las organizaciones pueden mejorar los resultados: identificar nuevas oportunidades de productos, optimizar precios, reducir costos, riesgos y mejorar la experiencia del cliente, entre otras iniciativas de negocio.

Para cualquier empresa que busque una ventaja competitiva en los datos, el gobierno de datos empresariales se ha convertido en una prioridad estratégica. El gobierno de datos garantiza que los datos utilizados para tomar decisiones del negocio sean confiables. Esto le permite medir y monitorear la calidad de sus datos, para lograr un mayor éxito en todas sus actividades, desde los análisis hasta la experiencia de los clientes. Al mismo tiempo, contar con un programa de gobierno de datos empresariales fomenta el cumplimiento de las normativas, contribuye a la reducción de los riesgos y protege los datos sensibles y personales.

Un programa de gobierno de datos exitoso proporciona un gobierno integral de principio a fin de los datos empresariales críticos. Garantiza que los datos empresariales sean confiables, gobernados y protegidos permitiendo la colaboración entre su empresa y las partes interesadas de TI, de modo que los datos no sean solo responsabilidad de TI, sino algo que debe usarse y gobernarse en toda su organización.

Sin embargo, ninguno de estos beneficios es posible hasta que sepa qué datos tiene, cuál es su calidad, qué sistemas y procesos los utilizan. Aquí es donde un catálogo de datos empresariales marca la diferencia. Y no cualquier catálogo de datos, sino uno que permita una gestión ágil de los datos a escala, con la capacidad de proporcionar una amplia visibilidad de todos los activos de datos de toda la empresa, así como una comprensión profunda de esos activos.



El **92%** de los ejecutivos afirma que el gobierno de datos es sumamente importante para la adopción empresarial de Big Data.<sup>1</sup>



# No se puede gobernar lo que no se conoce o no se puede clasificar

Un programa de gobernanza de datos eficaz debe incluir dos aspectos fundamentales: descubrir y comprender datos relevantes. Sus iniciativas de gobierno no llegarán muy lejos si no incluyen estos dos procesos esenciales.

Sin embargo, el volumen y la variedad de datos que las empresas modernas deben gestionar plantean un reto importante. Los datos relevantes pueden residir localmente, en la cloud o en entornos de multicloud. Y si bien los datos a menudo existen en silos (aislados por límites como la función comercial, la geografía o la fuente), las organizaciones deben gestionar cada vez más los datos de manera integral, ya sea para aprovechar nuevas oportunidades o mitigar riesgos. Por ejemplo, las normativas centradas en los datos como el reglamento general de protección de datos de la UE (RGPD) y ley de privacidad del consumidor de California (CCPA) requieren una visión en toda la empresa de los datos personales del cliente.

Por lo tanto, el grado de éxito en el descubrimiento y el inventario sustenta el éxito de su programa de gobierno de datos. Veamos estos dos procesos más de cerca.

## Detección de datos

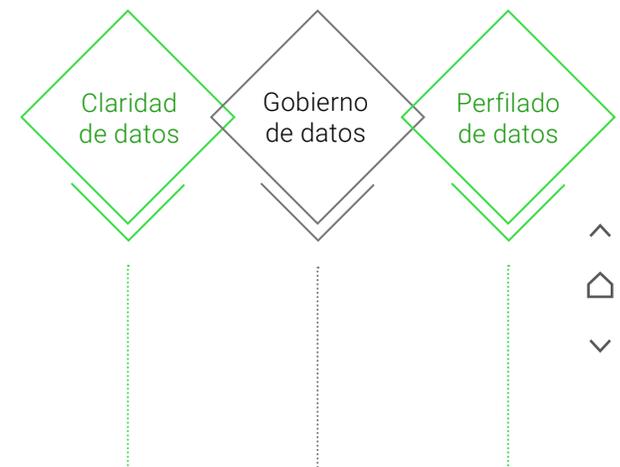
Cualquier iniciativa de gobierno de datos debe responder a unas sencillas preguntas:

- ¿Qué datos poseo?
- ¿Dónde se encuentran?
- ¿En qué procesos intervienen?
- ¿Cuál es su calidad?
- ¿Quién es el responsable de ellos?
- ¿De dónde proceden?
- ¿Con qué datos están relacionados? Etc.

Una solución de catálogo de datos proporciona respuestas. Le ofrece visibilidad de los datos sea cual sea su ubicación, para que pueda comprender qué datos incorporar a su programa de gobierno de datos empresarial.

Pero dado el gran volumen de datos con el que manejan las empresas hoy en día, un catálogo de datos debe ofrecer estas respuestas a escala. La única forma de hacerlo es aprovechando la inteligencia artificial (AI) y el aprendizaje automático (ML). Los esfuerzos manuales simplemente no son adecuados para la tarea.

Los tres principales impulsores de la inversión en un catálogo de datos: <sup>2</sup>



El catálogo de datos debe brindar visibilidad de los datos donde quiera que residan: en aplicaciones, sistemas de infraestructura, en las instalaciones o en el cloud. También debe poner los datos en contexto, para que los usuarios puedan ver de dónde provienen, quién los ha estado usando y cómo se han usado en el pasado. Este contexto ayuda a los usuarios a tomar decisiones informadas cuando utilizan datos para que, en última instancia, aporten un mayor valor al negocio.

### Comprensión de los datos

Los datos deben definirse y clasificarse para que se puedan buscar y utilizar en toda la empresa, y este paso va más allá de los datos en sí e involucra políticas, procesos e identificación de partes interesadas. Muchas organizaciones comienzan definiendo políticas y procesos y luego proceden al descubrimiento.

Aunque se puede utilizar la inteligencia artificial y el aprendizaje automatizado para obtener una gran ventaja en esta materia, es importante que los miembros de las empresas e IT involucrados colaboran en esta fase. Por ejemplo, al definir datos querrás tener un glosario de términos comerciales acordados. Este vocabulario proporciona un significado y un contexto claros que se vinculan con los metadatos técnicos subyacentes. Un glosario debe responder preguntas sobre términos comerciales como:

- ¿Cómo se define este término?
- ¿Quién es el propietario de este término?
- ¿Dónde se originan los datos relacionados con este término?
- Si hay cifras asociadas a este término, ¿cómo se calculan?

Los beneficios de tener un glosario incluyen mejorar la comunicación en toda la empresa, eliminar conflictos, reduce el tiempo para encontrar datos, reduce la cantidad de iteraciones necesarias para la entrega del proyecto y proporcionar una base sólida para las iniciativas de gobernanza de datos, cumplimiento normativo y administración de datos.

Llegados a este punto, es clave asociar términos, definiciones y políticas del glosario empresarial a activos de datos técnicos. Tendrá mayor éxito si puede realizar estas asociaciones automáticamente a escala aprovechando la inteligencia artificial y el aprendizaje automático.



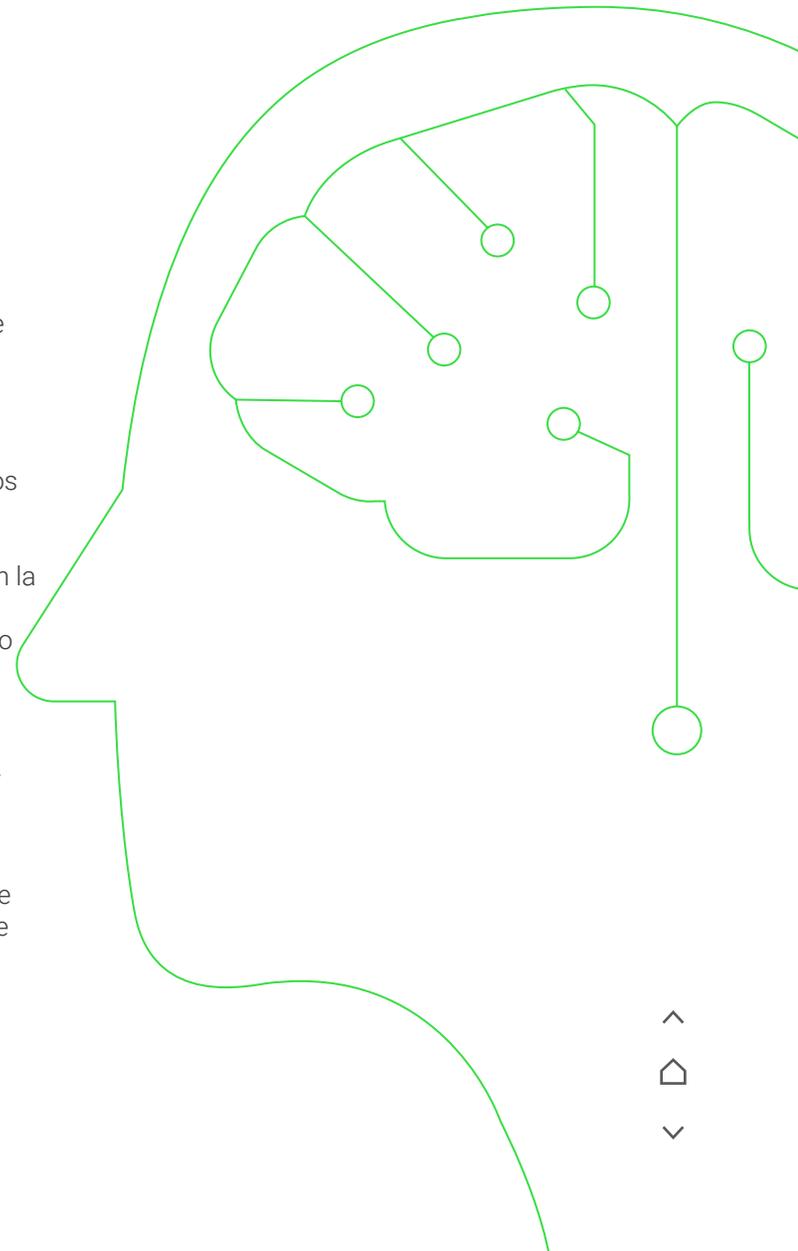
### El papel de la inteligencia artificial

El gobierno de datos debe diseñarse para reducir los procesos manuales que consumen mucho tiempo y que dificultan o imposibilitan el escalamiento con el aumento del volumen y la variedad de datos. También debería adaptarse a la expansión del número de usuarios finales que esperan un acceso fácil y bajo demanda a datos democratizados. Al aprovechar la inteligencia artificial para automatizar y optimizar procesos como el descubrimiento y la catalogación de datos, y al agregar un rico contexto empresarial a los datos, el gobierno de datos libera al personal talentoso para centrarse en análisis más profundos y procesos de mayor valor.

Un catálogo de datos impulsado por la inteligencia artificial proporciona un motor de descubrimiento basado en aprendizaje automático para escanear y catalogar activos de datos en toda la empresa, ya sea en las instalaciones, el cloud y el Big Data en cualquier lugar. Busque capacidades de descubrimiento automático, enriquecimiento y conservación basadas en la inteligencia artificial que permitan:

- Identificar y clasificar automáticamente entidades en datos estructurados y no estructurados.
- Detectar automáticamente similitudes entre los conjuntos de datos y comprender las relaciones y dependencias.
- Asociar automáticamente términos y definiciones de negocio a conjuntos de datos físicos.

Si bien los catálogos de datos inteligentes tienen la capacidad de escanear y catalogar automáticamente activos en toda la empresa, no basta con proporcionar visibilidad de todos los datos. El catálogo también debe guiar a los usuarios hasta los datos más relevantes y confiables para sus necesidades empresariales. Mediante el aprovechamiento de la inteligencia artificial, un catálogo de datos inteligente puede proporcionar el contexto empresarial para los activos de datos a escala, impulsando el éxito de sus iniciativas de cumplimiento y gobernanza de datos.



# Cinco funcionalidades esenciales

**Para respaldar programas de gobierno de datos que ofrezcan datos confiables, gobernados y protegidos, asegúrese de que su solución de catálogo de datos brinde estas capacidades.**

Un catálogo de datos empresariales respalda la gobierno de datos, pero también proporciona una base para cualquier iniciativa que dependa de datos confiables, gobernados y protegidos. Así es como un catálogo de datos inteligente permite el éxito.

## 1. Conectividad de metadatos amplia y profunda

El éxito del gobierno de datos depende de la capacidad de descubrir datos mediante el escaneo y la catalogación de activos de datos en toda la empresa. La capacidad de trabajar sin problemas con datos de una amplia gama de fuentes (datos locales tradicionales, datos en el Cloud, Big Data, herramientas de inteligencia empresarial y otras aplicaciones, datos estructurados y no estructurados) es esencial. No se trata sólo de la amplitud de la conectividad. También es la capacidad de extraer diferentes tipos de metadatos que se pueden aprovechar para una comprensión más profunda de los datos y para automatizar procesos posteriores.

## 2. Análisis de linaje e impacto

Para gobernar los datos de manera eficaz, es necesario comprender dónde se encuentran, quién es su propietario, de dónde provienen y cómo se utilizan. Un catálogo de datos eficaz debe ser de principio a fin, de forma que los usuarios puedan comprender la procedencia de los datos críticos. Y al vincular ciertos datos (como, por ejemplo, números confidenciales de tarjetas de crédito) a términos y políticas

comerciales, proporciona orientación para realizar análisis de impacto relacionados con reglas internas y regulaciones externas.

## 3. Monitoreo de la calidad de los datos

Al fin y al cabo, los datos deben ser fiables y coherentes. Tener acceso a la calidad de los datos es imprescindible para el éxito de cualquier programa de gobierno de datos empresariales. Su solución de catálogo de datos debe proporcionar visibilidad de las estadísticas de perfiles de datos (como distribuciones de valor y otros patrones) y metadatos técnicos que faciliten una comprensión más profunda de la calidad de los datos. Al guiarlo hacia datos de mayor calidad, los resultados finales serán más adecuados para informes de cumplimiento y análisis avanzados, así como para brindar información significativa a los usuarios finales de toda la organización, hasta el nivel del director general y del director de TI.



#### 4. Escalabilidad impulsada por IA

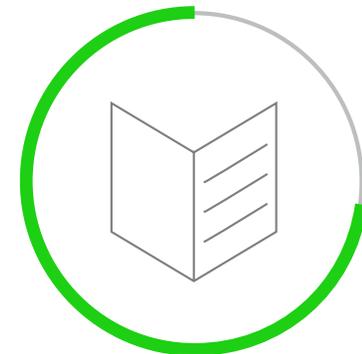
Lograr una escala empresarial no es práctico sin una plataforma impulsada por IA que pueda agilizar procesos, como la curación, el etiquetado y la clasificación automatizados, la detección automática de datos similares y la asociación automática de términos comerciales a activos de datos técnicos. Con miles de conjuntos de datos, cada uno de los cuales contiene cientos de columnas y miles de términos empresariales, la identificación y el etiquetado manuales serían una alternativa poco práctica. Si se necesitan semanas o meses para completar un proceso de descubrimiento de datos o generar un informe, los resultados pueden tener poco o ningún valor.

Para agilizar el tedioso proceso de asociar metadatos técnicos con términos y políticas comerciales, una solución de catálogo de datos impulsada por IA puede asociar automáticamente términos y políticas comerciales específicos con los conjuntos de datos relevantes, una necesidad para lograr y mantener la escala empresarial. Los administradores de datos pueden centrarse en análisis de mayor valor de conjuntos de datos seleccionados por IA para ofrecer mucho más valor.

#### 5. Colaboración

Además de aprovechar el poder de la IA para la escalabilidad, un catálogo de datos debe aprovechar el conocimiento colectivo y la experiencia en la materia de todas las partes interesadas en contacto con los datos. Esto incluye arquitectos de TI, propietarios de datos, administradores de datos y consumidores de datos.

La colaboración entre todas estas partes interesadas es fundamental para un programa de gobierno de datos exitoso. Al proporcionar una visión holística de sus datos, junto con un rico contexto empresarial, el catálogo de datos debería permitir que TI respalde mejor las solicitudes y requisitos empresariales. Los usuarios empresariales deberían poder etiquetar, comentar, calificar, aprobar flujos de trabajo, etc. Con una estrecha integración entre las políticas y los datos bajo controles de gobernanza, un catálogo de datos inteligente permite que tanto las empresas como las partes interesadas de TI utilicen los datos de manera efectiva y responsable.



Se estima que el **73%** de las empresas que utilizan un catálogo de datos estarán más satisfechas con su capacidad para compartir datos en comparación con los usuarios que no utilizan catálogos.<sup>3</sup>

# La base de un gobierno de datos empresariales

## **Un catálogo de datos empresariales es esencial para controlar sus datos empresariales.**

A medida que crece la variedad y el volumen de datos empresariales, los métodos tradicionales para realizar un seguimiento de sus datos no pueden escalar. Un catálogo de datos empresariales inteligente impulsado por IA es la única solución para una visibilidad integral de los datos en toda la empresa.

Para descubrir el valor de sus activos de datos, es importante sentar las bases adecuadas. Su solución de catálogo de datos debe tener un conjunto sólido de capacidades que preparen sus programas de gobierno de datos para el éxito. Sólo entonces podrá descubrir activos de datos de forma rápida y sencilla, comprender qué es realmente útil y garantizar que los datos sean confiables.

La gobernanza de datos empresariales ya no es una función aislada: las empresas y la TI deben trabajar juntas para gobernar y aprovechar los datos. Al proporcionar una visibilidad integral de los datos junto con un marco que respalda la colaboración entre la empresa y la TI, un catálogo de datos empresariales se convierte en un requisito para programas eficaces de gobierno de datos y, en última instancia, para una transformación digital basada en datos.

# Fuentes de Informatica

Esta guía es una traducción al español de la versión original "[Data Cataloging for Data Governance: 5 Essential Capabilities](#)" publicado por Informatica.



**Informatica**  
PLATINUM PARTNER

PowerData, es una compañía multinacional de origen español con destacada presencia regional, como especialistas en gestión de datos en la nube, está preparada para ayudar a las organizaciones a acelerar su camino hacia la transformación digital y brindar la previsión necesaria para que sean más ágiles y aprovechen nuevas oportunidades de crecimiento.

PowerData e Informatica se convirtieron en socios desde el año 2000, desde entonces se ha estrechado la relación con más de 100 organizaciones de diversos sectores, categorías y nichos, quienes han depositado su confianza en las soluciones de Informatica, y en el expertise y Know-how de PowerData.

Te invitamos a explorar los proyectos donde aportamos valor con la gestión de datos.

[powerdata.es](https://powerdata.es)

