



¿CÓMO GOBERNAR  
LOS PRODUCTOS DE DATOS?  
**HACIA UN GOBIERNO  
DE LOS DATOS FEDERADO**

---

# 01

## Introducción

El potencial **económico y social** de la utilización de datos es enorme, desde permitir la creación de nuevos productos y servicios basados en tecnologías innovadoras, pasando por aumentar la eficiencia de la producción, hasta dotarnos de instrumentos para combatir los problemas de la sociedad. Para poder alcanzar este potencial es necesario que haya datos disponibles, compartidos de manera confiable y fáciles de reutilizar técnicamente.

Conseguir un intercambio entre los productores y los consumidores de datos en un **entorno adecuadamente gestionado** es sin duda alguna uno de los resultados más relevantes de una iniciativa de gobierno de los datos, y a la vez uno de los más difíciles de alcanzar.

**La democratización de los datos** (tanto interna como externamente) es una necesidad imperiosa para ser capaces de utilizar los datos de manera holística en la toma de decisiones de negocio, pero hay muchos aspectos a tratar cuando se implanta un marco de gobierno de los datos donde la propiedad y la gestión de estos toman especial relevancia.

En el *paper*, "**Gobierno del 2.0: del gobierno del dato al gobierno de los productos de datos**", apuntamos que una de las características de la evolución de los modelos tradicionales de gobierno del dato es incorporar el **gobierno del ciclo de vida del dato**, y por lo tanto ampliar el "enfoque dato" al "enfoque del dato como producto y de los productos de datos".





Nuestra visión de un **gobierno del dato integral** se caracteriza por un enfoque que combina principios y políticas globales para los aspectos más relevantes, y una gestión descentralizada y federada por dominios de datos.

Esta aproximación a una estrategia de datos distribuida puede ir en contra de lo que se considera normal en otros aspectos de la vida. Siempre se nos ha enseñado que para poder correr tenemos que aprender a andar, para poder gestionar antes hemos tenido que operar. Sin embargo, en materia de datos creemos que, desde las áreas de negocio, lo primero que se debe asumir es la responsabilidad para después aprender a **generar valor a partir de los datos**.

El tratamiento de los datos como un producto, trasladando la propiedad a los equipos productores sigue la teoría de que la propiedad está mejor alineada con la creación de los datos, lo que aporta mayor **visibilidad y calidad** a estos. Este desplazamiento de las responsabilidades aborda las problemáticas derivadas de la explosión de los datos que han sobrepasado las capacidades de equipos de datos centralizados encargados de limpiar, armonizar e integrar los datos.

La delegación de responsabilidades sobre conjuntos de datos específicos significa que las organizaciones se dividen en dominios de datos cada una de los cuales asumen la responsabilidad de servir datos, mantener los metadatos asociados, mejorar la calidad de los datos, aplicar la gestión del ciclo de vida, realizar el control del código, etc.

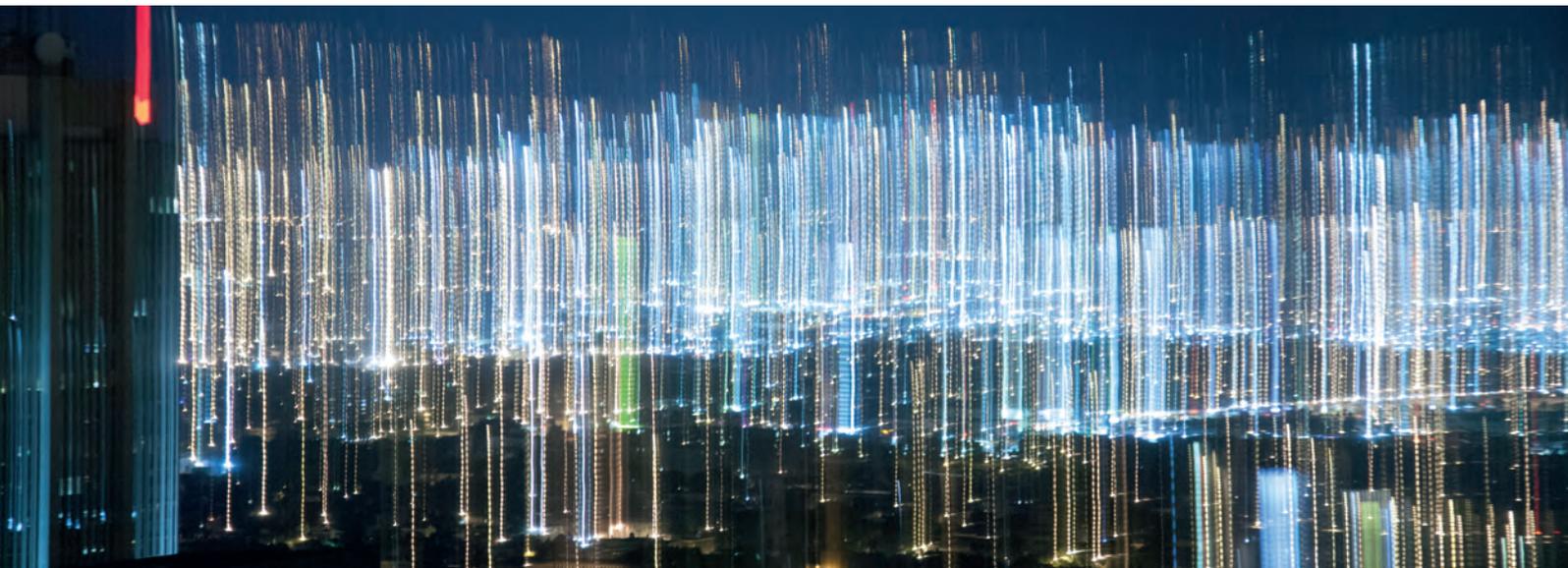
Encontrar el **equilibrio** entre el grado de centralización y de distribución de infraestructura y gobierno de los datos es hoy una cuestión de primer orden en cualquier organización. Para la decisión es importante tener en cuenta criterios como: dimensión y complejidad del negocio, volumen, aspectos regulatorios, tipos de datos y grado de madurez y talento en la gestión de datos de las diferentes unidades de negocio. La existencia o no de estas capacidades distribuidas, el tiempo y el que pueda resultar crearlas son criterios fundamentales de decisión para tener en cuenta.

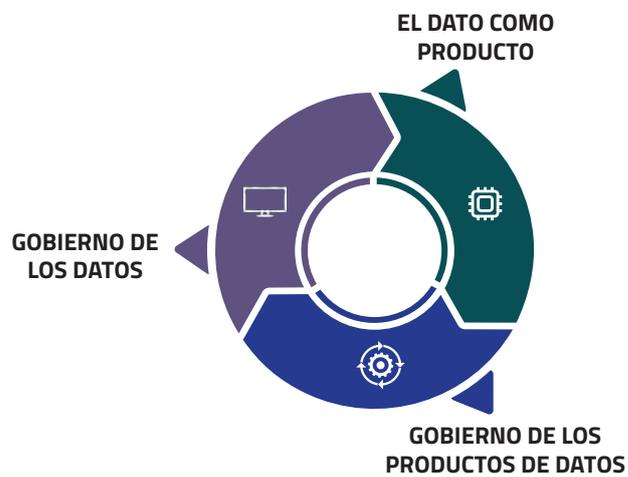
# 02

## El rol de la Oficina Central (*Data Management Office*) en un entorno federado

En este entorno federado de generación de valor a partir de los productos de datos, el rol de la **oficina del dato** es de una unidad central que actúa como *control tower* y que proporciona a las diferentes áreas de negocio y dominios de datos:

- Unos **principios globales de Gobierno del Dato** que incluyen un marco de trabajo común de acceso al dato, metadato, linaje, y seguridad.
- Los principios que definen un producto de datos y las **reglas para crearlos**. Por ejemplo, los diferentes dominios federados son considerados como productos de datos y tienen responsabilidades sobre el contenido, calidad e interoperabilidad de los datos que se gestionan. También son considerados productos las diferentes soluciones informacionales, analíticas, Webs, Apps que encapsulan tres componentes: código, datos y metadatos e infraestructura.
- La **descripción y responsabilidades** de los diferentes participantes que tienen algún rol en el Gobierno del Dato distribuido: *Owner* de negocio del dominio, *Owner* a nivel IT, responsabilidad del equipo DevOps, etc.
  - ◇ Para poder considerar en una organización el dato como un producto, los datos de los dominios y también los productos de datos de esos dominios deben tener un **Product Owner**. Un responsable de que los datos sean detectables, es decir accesibles y que puedan ser direccionados hacia los consumidores. También que sean autodescriptivos, es decir, que entendamos su significado de forma que tenga sentido para el negocio.
  - ◇ Es también relevante la administración de quién puede acceder y publicarlos y guiar en el consumo de datos con metadatos, como casos de uso y consultas de ejemplo y por supuesto monitorizar el autoconsumo para evaluar el retorno de la inversión.





Por otra parte, debemos tener en cuenta la creación de **mecanismos de compartición de datos** (*data sharing agreements*, o DSA) que establecen los principios de **responsabilidad, calidad y seguridad** entre productor/consumidor. Los DSA establecen:

- Normas de calidad y tolerancias.
- Identificación de la responsabilidad de los datos y los roles relacionados.
- Uso aceptable de los datos: Se indica para qué fines aceptamos el uso, tratamiento y explotación de los datos.
- Protocolos de infracción.

El uso de los DSA permite:

- La trazabilidad de la cesión y la adecuada gestión de identidades.
- Aplicar mecanismos de cifrado/anonimizado.
- La reusabilidad de los datos.





# 03

## Infraestructura o plataforma de datos en un entorno federado

Disponer de una infraestructura o plataforma de datos **dotada de aceleradores técnicos** que sean utilizados por toda la organización es básico para que los equipos de producto ganen en eficiencia y productividad. Por consiguiente, es básico disponer de un presupuesto suficiente y gestionado de manera óptima para la construcción de la plataforma. Normalmente los diferentes componentes se despliegan en paralelo a la construcción de los productos evitando al máximo la generación de “deuda técnica”.

Además de los componentes de uso común por los diferentes negocios, existen arquitecturas de referencia para los productos que normalmente funcionan como arquitecturas preferenciales, pero no son obligatorias.

Esta infraestructura o plataforma de datos *as a service* es diseñada y mantenida por un equipo central de ingenieros y arquitectos de datos, y los **componentes de uso común** son:

- Herramienta de Gobierno del Dato (*multi-cloud/híbrida*).
- Marketplace de datos y productos de datos generados por la organización y puestos a disposición de las diferentes unidades organizativas.
- *Knowledge store*: Documentación de las soluciones; código, arquitectura...
- Catálogo de activos IT reaprovechable para los productos de datos.
- Seguridad e identificación de identidades.
- Zona *Help Desk* para la gestión de incidencias y mantenimientos.

Cada organización tiene su propio ecosistema de datos y sus propios retos. No todas las organizaciones están preparadas para la descentralización a gran escala. Sin embargo, podemos encontrar un equilibrio eligiendo una topología de referencia que mejor se ajuste a nuestras necesidades.

# 04

## Los dominios de datos

La definición de los dominios de datos en una organización federada es un aspecto clave y aconsejamos:

- Establecer **dominios de datos con fronteras claras** y cada una de ellas con un equipo *DataOps* asignado. Un dominio de datos puede ser por ejemplo un área funcional o negocio de la organización. Ha de tener sentido para la organización o de lo contrario es complejo establecer un *ownership* claro, cada dominio debe tener un *Product Owner*.
- Cada dominio es responsable de **depurar y establecer los controles de calidad de sus datos**. Es responsable *end-to-end* de tener datos dispuestos para su consumo y entrega a los productos de datos que lo necesiten sean de su dominio o de otros dominios. El tratar los datos como un producto hace referencia precisamente a tener claro los estándares que los datos deben cumplir y aplicarlos por los *Product Owner* de cada dominio.
- Es importante en cada dominio **separar los datos de los productos de datos** de esta forma los datos son reutilizables por varios productos. De lo contrario, si los datos ya incorporan unas determinadas transformaciones o lógicas de negocio, impedimos que sean reutilizables para otros usos.
- Todos los dominios comparten los mismos **estándares de interoperabilidad** que les permiten compartir datos. En ocasiones cuando los datos son compartidos por varios dominios hay que establecer esquemas. Hay varias soluciones como establecer dominios cliente-proveedor de datos.

Al mismo tiempo que definimos los diferentes dominios, un gobierno del dato federado ha de asegurar que la distribución de responsabilidades y activos entre elementos centrales y dominios es la que consigue maximizar el valor que obtenemos con los datos. Es importante tener en cuenta que:

- Aunque hablemos de un modelo de gobierno y arquitectura del dato federada es imprescindible tener una **perspectiva global y elementos comunes** para que el modelo sea eficiente y los activos que genera sean reutilizables.
- **Cada dominio utiliza aceleradores comunes de la plataforma global**, pero puede necesitar su propio set de tecnología específico.
- **Los dominios operan sus propias aplicaciones y productos** y se adhieren a los principios globales.
- Aunque existan varios dominios debe existir un catálogo de datos común a los dominios. Es decir, existe un **central data management hub** que mantiene un catálogo común y hace una monitorización global tanto operativa como económica. Además, este *hub* provee servicios para automatización, *data discovery* y comparte mejores prácticas para reutilizar los activos que se van generando en todos los dominios.

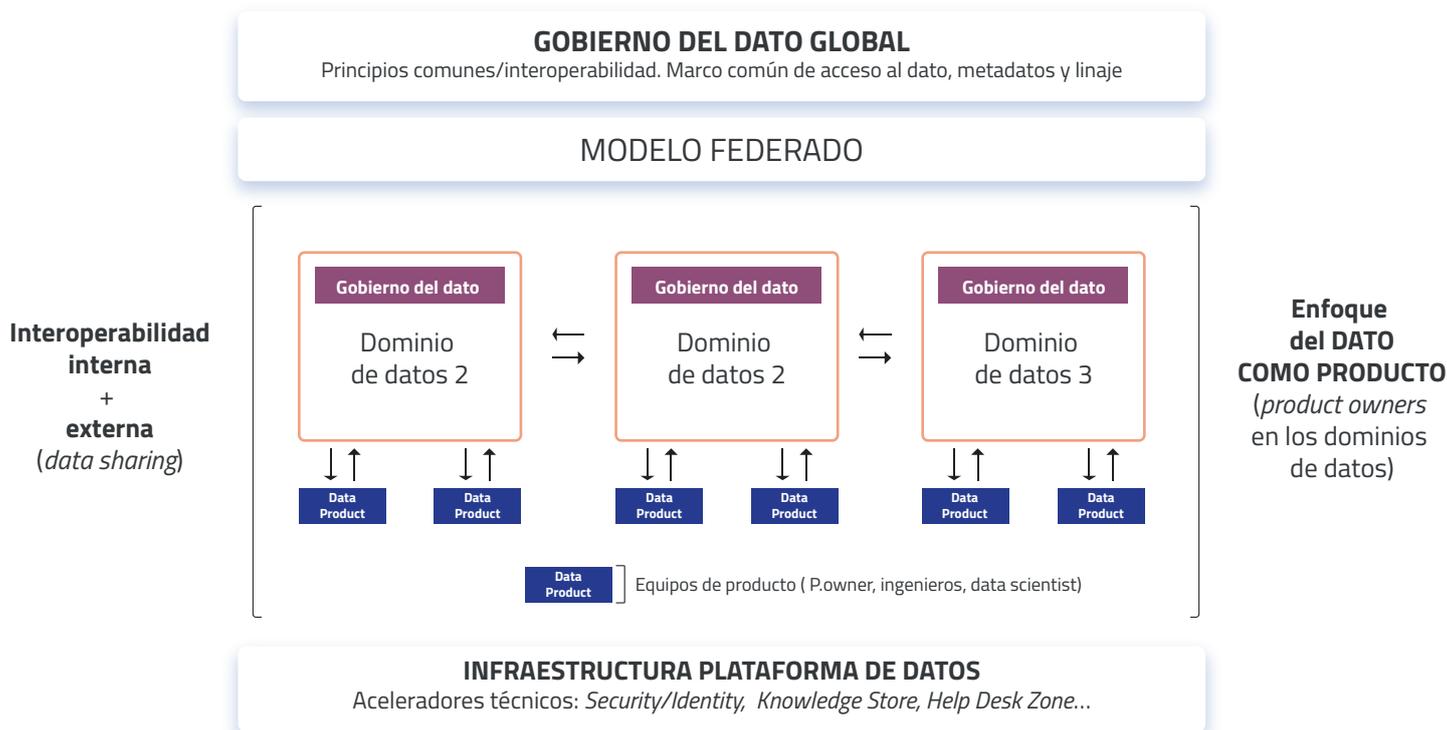


## La transición hacia la aportación de valor mediante los datos

Las organizaciones intentan utilizar los datos para impulsar **experiencias diferenciales para los clientes**, crear plataformas y servicios, generar conocimiento y redefinir sus operaciones gracias a procesos de trabajo inteligentes, mediante el uso de tecnologías como la Inteligencia Artificial.

**Los datos son el activo intangible más importante de las organizaciones**, pero suelen estar distribuidos en múltiples aplicaciones transaccionales, entornos analíticos, y entornos *Cloud* dentro y fuera de la organización, dificultando el acceso y el autoconsumo lo que provoca la aparición de silos de información y la duplicidad de trabajos.

**La evolución del gobierno de los datos ampliando su alcance**, pasando de gestionar los datos en bruto a gestionar los datos como un producto y los productos generados a partir de ellos, mejora significativamente el impacto que tienen estas iniciativas en términos de valor aportado a negocio.



En el tránsito hacia el Gobierno del dato 2.0 nuestras **recomendaciones incluyen:**

- Dedicar esfuerzo y tiempo a entender las capacidades y dedicación interna que necesitaremos tanto a nivel de negocio (para cada dominio) como a nivel de las áreas de datos e IT para establecer un nuevo Gobierno 2.0.

- No improvisar en cuanto a qué componentes de la plataforma de datos son centrales y cuales se dejan a la gestión de cada dominio de datos. La dispersión tecnológica actual no ayuda, pero a nivel conceptual ha de estar muy claro.
- Implantar en la arquitectura y en las herramientas de Gobierno del Dato los principios de Gobierno, interoperabilidad y seguridad. Los tiempos de generar normativa y papel que nadie conoce o consulta ya han pasado.
- A pesar de hablar de modelos federados o distribuidos continúa siendo de la máxima relevancia la función del *Central Data Management Hub* puesto que tiene una función de Gobierno y monitorización global. Ha de tener una dotación de recursos humanos y económicos suficiente entre otras razones porque gestionará la adopción global del modelo.

Por último, resaltar que **la extensión del gobierno a los productos de datos ayuda a dar respuesta a imperativos estratégicos de las organizaciones:**

- Ser capaces de gobernar los modelos analíticos de explotación de los datos estableciendo métricas de rendimiento, creando catálogos, etc.
- Medir la aportación real de las iniciativas mediante la valoración económica de los datos y de los productos de datos, y la generación de modelos de monetización interna y externa.
- Impulsar la adopción de la nueva cultura, los nuevos conocimientos y procesos y políticas incorporando aspectos de gestión del cambio y *data literacy*.
- Tener en cuenta La ética de los datos referida a las buenas prácticas en torno a cómo se recogen, comparten y utilizan los datos. Estas prácticas son especialmente relevantes cuando las actividades de datos tienen el potencial de impactar negativamente en las personas y la sociedad.
- Apoyar en el cumplimiento normativo de regulatorios de protección de datos personales como GDPR de la Unión Europea, y otros que están por venir en el corto plazo y que impactarán en las estrategias de datos de las organizaciones: *Data Governance Act, Data Act, y el Artificial Intelligence Act.*

## ¿Hablamos?

¿Tú también crees en esta evolución del Gobierno del Dato?

**¡Nos encantaría hablar contigo!**

Contacta con nosotros en [info@bluetab.net](mailto:info@bluetab.net).

