

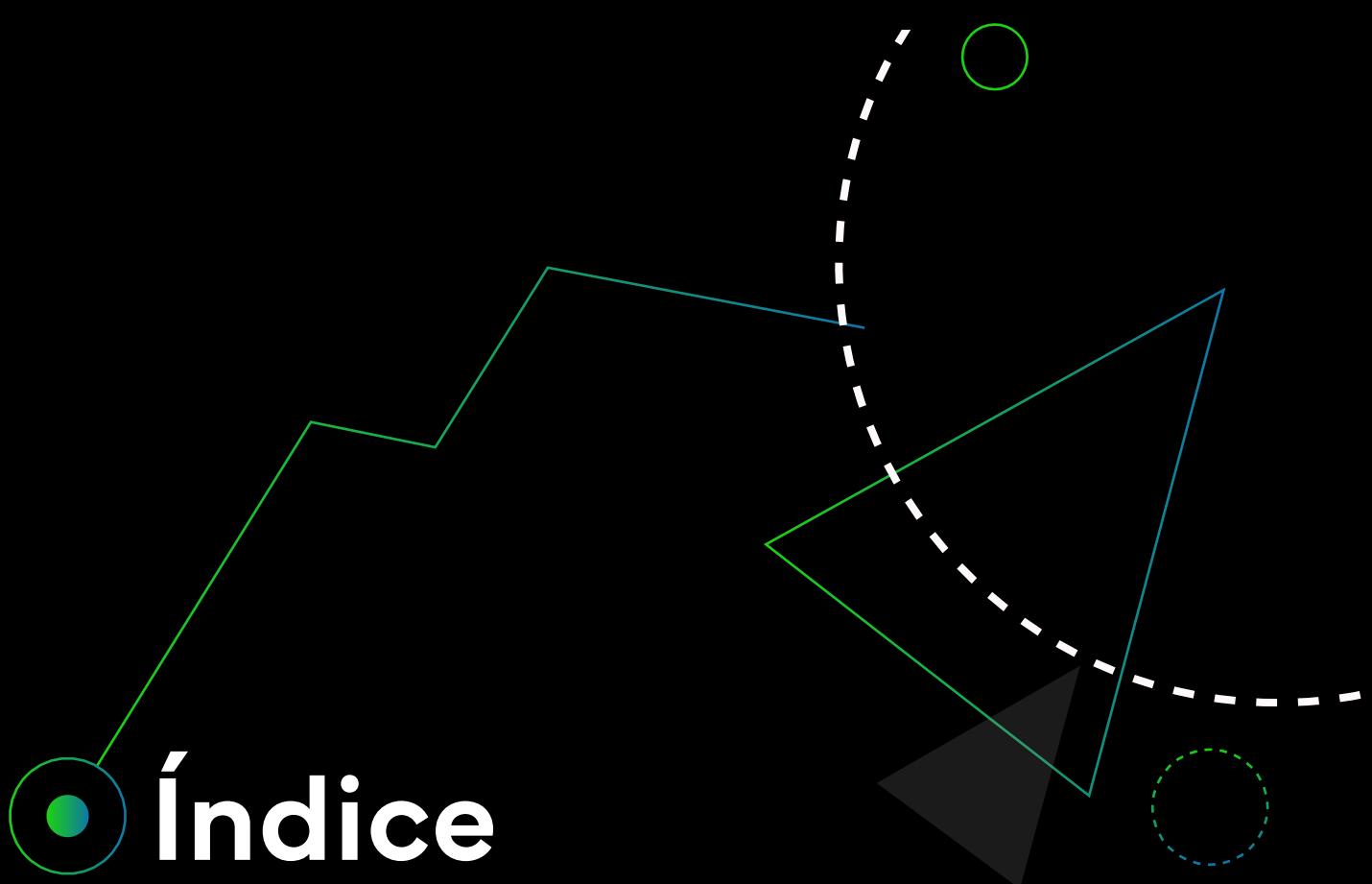
 PowerData



Guía para entender la gestión de datos

Todo lo que tu empresa tiene que saber
sobre la gestión de datos





Índice

Introducción	3
Eficiencia en la gestión de datos	5
Abundancia, síntesis y claridad	6
Práctica continua de Data Management	7
El gran desafío: la seguridad	8
Los 4 ejes del Data Management	9
Escenarios Cloud y multi-Cloud	10





Introducción

Quien haya intentado alguna vez hacer malabares con pelotas, sabe que hacerlo con dos es muy fácil, que con tres es posible y que, con más, empieza a ser un desafío que requiere mucha preparación y entrenamiento. No sólo se trata de una habilidad manual sino, también, de una actividad que requiere de un gran esfuerzo de concentración. ¿Qué sucede cuando se agregan más elementos? La cuestión se complica; es habitual que no sólo se caiga la adicional, sino todas o casi todas las pelotas porque ante el primer problema, surge el caos y todo se desordena.

La alegoría vale para hablar de la situación que se presenta en la actualidad con la gestión de datos en las organizaciones (data management - DM). Manejar un volumen acotado de información era un juego de niños, pero la acelerada transformación digital que se transita en la actualidad, obliga a pasar a otro nivel.



El mercado proyectado de DM para 2031 superaría los 19 mil millones de dólares, con una tasa anual promedio de crecimiento mayor al 12%.

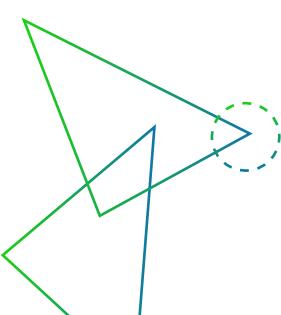
Fuente: Allied Market Research

Para toda persona de negocios estos números son muy llamativos. Cuando hablamos de gestión de datos, incluimos desde los motores básicos hasta las más sofisticadas soluciones analíticas, pasando por innumerables herramientas para la captación, almacenamiento, protección y aprovechamiento de los mismos, y todos los servicios asociados a ese ciclo de vida de la información.

Este crecimiento del negocio vinculado al data management se refleja en las implicancias que tenga para cada organización. Es decir, en las inversiones de tiempo y recursos necesarias para estructurar adecuadamente la información del negocio y para sacarle el máximo provecho a los datos que se generan en la operación cotidiana. Y de esta forma poder tomar decisiones con mayor sustento.

La gestión de datos tiene dos pilares -políticas y procedimientos- alineados con la estrategia global de la organización, y aquella específicamente trazada para la información.

- Las políticas incluyen criterios generales, definiciones que orientan las inversiones y los proyectos en los que TI tendrá que trabajar en articulación con las diferentes áreas de la estructura.



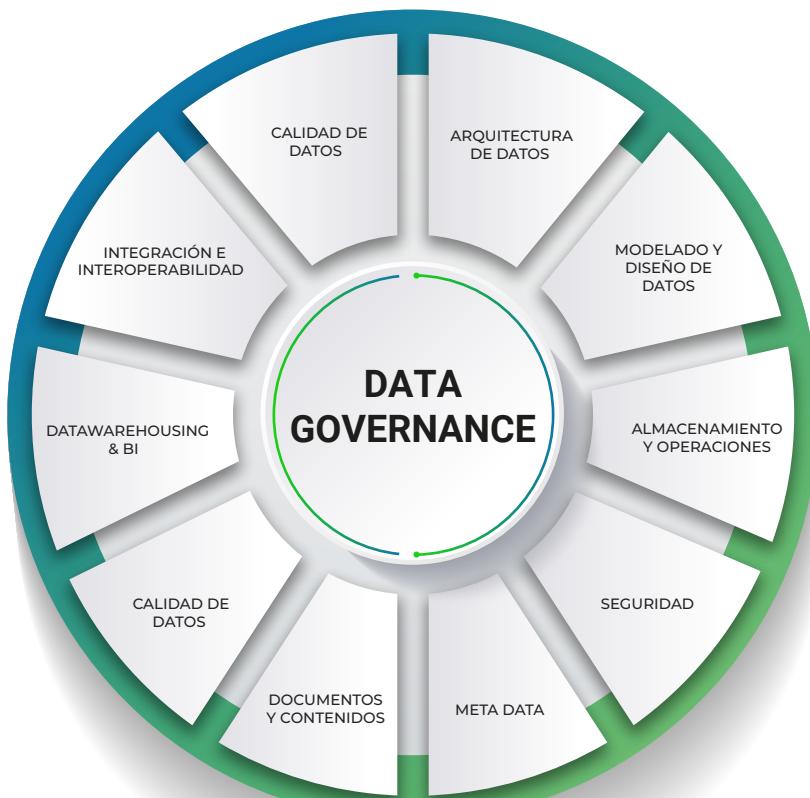


- Los procedimientos son el conjunto de rutinas que habrá que seguir y respetar en cada proceso de la actividad, tanto por el equipo de tecnología -especialmente de desarrollo- como así también todas aquellas personas que, de alguna manera, tienen datos entre sus manos. La idea es disponer de parámetros y guías de práctica para cuidar la exactitud, accesibilidad y seguridad de los datos.

Desde una perspectiva de análisis de negocios, la gestión de datos ayuda a organizar y procesar información. Desde el punto de vista de la ciencia de datos, se utiliza el data management para analizar grandes volúmenes de información. Todo esto ayuda a los tomadores de decisión, incluso, a saber qué información necesitan tener, cuál debe ser la política de permisos de acceso y uso, cómo se debe construir el plan de disaster recovery, etc.

Un gráfico de la comunidad de data management de Dinamarca, expone muy sintéticamente todos los elementos que hacen a una solución integral de gestión de datos.

Fuente: DAMA.dk





Eficiencia en la gestión de datos

Los datos son un activo de las organizaciones; por lo tanto, es importante que sean visibles y que estén al alcance. Así como los activos contables se computan en el balance, y eso valoriza a la organización -además de ser un requisito obligatorio de las normas vigentes-, lo mismo sucede con la información: tenerla ordenada, a disposición, segura y utilizarla de manera inteligente, tiene un impacto positivo en el negocio. Por cierto, año tras año se multiplican las reglas gubernamentales que tienen que ver con el manejo de datos, por lo que esto también es un estímulo para llevar adelante un buen data management.

Adicionalmente, que el personal sepa qué información tiene a su disposición, dónde está y cómo debe utilizarla, contribuye a elevar la productividad y a empoderar a los equipos con información valiosa para tomar acción.

A la hora de tomar decisiones, con una buena práctica de gestión de datos se tiene información precisa y actualizada, lo que permite identificar tendencias y patrones no necesariamente evidentes. Además, permite monitorear indicadores clave predeterminados

(key performance indicatorsKPI), para evaluar progresos y áreas de mejora. Indudablemente, esto mejora ese proceso decisorio y redundar en beneficios de negocios tangibles.

El historial de las decisiones tomadas y el contexto de cada una, con el respaldo de la información correspondiente, permite analizar la situación con mayor perspectiva, ver la big picture, contemplando ciclos más extensos del negocio con una mirada más estratégica.





Abundancia, síntesis y claridad

Es común que existan diferentes caminos de llegada de información a la organización; por un lado están las diferentes áreas (marketing, ventas, servicios, atención al cliente) y, por otro, los diferentes canales (sistemas online, redes sociales, puntos de venta, interacción telefónica, correo electrónico, etc.). Cuando todos los procesos son digitales, el volumen de información crece exponencialmente y hay un riesgo importante de registrarla con redundancias que derivan en inconsistencias.

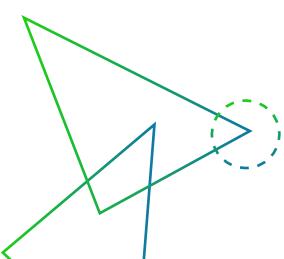
Para prevenir esto, una de las cuestiones principales es asignar roles y responsabilidades en lo concerniente a la gestión de los datos, lo cual va de la mano de la definición -a priori- de procesos y procedimientos, junto a los controles respectivos. Todo esto constituye un marco de trabajo para la gestión de datos.

Uno de los procesos esenciales tiene que ver con una limpieza (cleaning up) regular de las bases de datos: eliminar duplicados, estandarizar tipos de datos y formatos, completar y clasificar metadata, de forma tal que haya una sola versión de cada registro, indexada adecuadamente. La estandarización de parámetros de metadata no sólo ayuda a mantener limpio y ordenado el universo de información disponible, sino también agiliza y acelera el acceso a los datos de manera oportuna y precisa.

Otro elemento significativo tiene que ver con la definición de reglas de validación y el monitoreo de la salud de los datos para prevenir errores o inconsistencias, resguardando la calidad de los datos.

Todos los procesos de limpieza, control de calidad y, de ser necesario, transformación de datos, pasan a otro nivel cuando se logra implementar la automatización, de la mano del machine learning e, idealmente, la inteligencia artificial.

Finalmente, el data management es un antídoto contra la pérdida de información. Si hay abundancia y desorden, es común perder cosas; incluso, suele ocurrir que se descubre el problema demasiado tarde. La otra cara de este riesgo está formada por la posibilidad de que sucedan incidentes imprevistos. En ambos casos, las prácticas sistemáticas de copias de seguridad (backup), las definiciones para los procedimientos de restauración (restore) o, integrados ambos elementos en un plan de respuesta ante desastres (disaster recovery), son la protección necesaria, para lo cual hay un sinnúmero de herramientas a disposición.





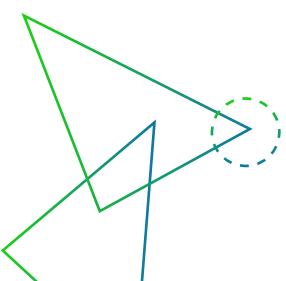
Práctica continua de Data Management

Existen factores internos y externos para implementar una práctica continua de data management. Entre los últimos, crecen año a año los requerimientos formales de parte de las autoridades a partir de las políticas de protección de datos personales, la confidencialidad de la información y otros derechos de cuarta generación. Las normas de compliance, deben cumplirse en tiempo y forma, para evitar sanciones.

Otro aspecto a considerar tiene que ver con la imagen de la organización; si, por alguna vía, se difundiera que fueron vulnerados los registros (robados o modificados) de manera espúrea, la reputación se vería afectada y eso también impacta en el negocio.

Entre las razones internas -que, obviamente, impactan por fuera también- se destaca la mejora en la eficiencia que se logra cuando hay un buen sistema integral de gestión de datos.

Reducir el tiempo empleado en encontrar la información necesaria en cada momento, que sea la adecuada y esté accesible para quien la necesita, mejoran la satisfacción de clientes y proveedores, la experiencia del personal y la celeridad de respuesta en requerimientos tanto externos como internos. Ser una organización impulsada por los datos (data driven) también ayuda a reducir costos y ganar en agilidad. Cuando un gerente o el directorio necesitan información urgente para tomar una decisión, el tiempo es oro y la calidad de los datos es esencial.





El gran desafío: la seguridad



Durante el tercer trimestre de 2022, los usuarios de Internet de todo el mundo sufrieron aproximadamente 15 millones de violaciones de datos, un 167% más que en el trimestre anterior.

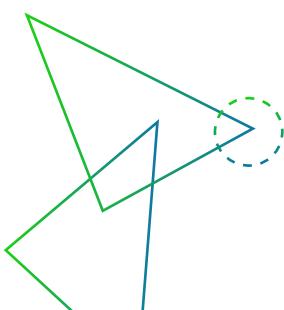
Fuente: Statista.

Los problemas para las organizaciones han sido -y son- múltiples y variados, e implican graves riesgos y costos.

El prisma de la protección de datos tiene muchas caras. Por un lado, un adecuado control de acceso a la información; el identity and access management (IAM) es una de las claves para minimizar los riesgos, en base a definir muy claramente quién puede acceder a qué sistemas o bases de datos, en qué momento y de qué manera, partiendo de la filosofía del zero trust (confianza cero) y uso del doble factor de autenticación.

Conocer las amenazas que están circulando y adoptar herramientas para protegerse específicamente de ellas, es imprescindible. Luego, aparecen otros elementos que no pueden faltar:

- Plan de backup y restore, no solo por ciberataques sino también por fallas de equipamiento o desastres fortuitos (incendios, terremotos, inundaciones, etc.). Además, es una herramienta clave para el almacenamiento por largos períodos de tiempo para el cumplimiento de las normativas.
- Sistemas de encriptación, para reducir la posibilidad de mal uso, difusión o acceso no autorizado a información sensible, tanto de parte de cibercriminales como de personal descontento.

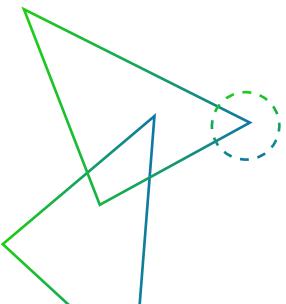




Los 4 ejes del Data Management

La gestión de datos es una estrategia global que puede estructurarse en cuatro ejes:

- **Data mining:** para extraer datos crudos de las bases de datos para extraer de ellos información de valor que sea de utilidad para tomar decisiones.
- **Data warehousing:** para almacenar grandes cantidades de datos estructurados, no estructurados y semi estructurados.
- **Data integration:** como el proceso de combinación de datos maestros que puedan provenir de diferentes fuentes, posibilitando una vista simplificada y coherente. Hoy en día también se aborda esta cuestión bajo el paradigma de los data lakes.
- **Data analysis:** que es el proceso de organización, manejo y entendimiento de lo que los datos significan. Esto ayuda a encontrar tendencias, patrones y relaciones en los datos principales.





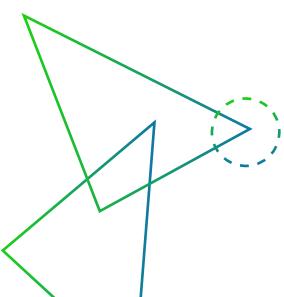
Escenarios Cloud y multi-Cloud

Todo lo señalado en las líneas anteriores, aplica a cualquier escenario: bases de datos y sistemas on premise, en un entorno de nube pública o privada, en arquitecturas mixtas o, incluso, distribuidas en diversas plataformas de nube.

A medida que las organizaciones migran su operación a la nube, ya sea en forma parcial o total, se popularizan prácticas específicas del cloud-based data management, un modelo que aporta escalabilidad, flexibilidad y mayor efectividad en términos de costos. La mencionada estrategia de data lakes es una de las que crecen con mayor impulso.

Hoy en día existen muchas alternativas: Microsoft Azure, Google Cloud Platform, Boomi Master Data Hub, Amazon Web Services, Talend y Panoply, son algunas de las más conocidas. No obstante, cada caso requiere de un análisis particular, para entender qué solución es la más adecuada según el estadío en el que se encuentre cada organización.

En resumen, entender cómo funciona -y cómo hacer que funcione- la gestión de datos, y cómo pueden las organizaciones poner en marcha las políticas y procedimientos más efectivos para obtener el máximo beneficio de la información disponible, implementando las mejores prácticas conocidas, es la clave para liberar el potencial de la información como un factor clave en el camino al éxito.





PowerData, es una compañía multinacional de origen español con gran presencia regional, está enfocada en todo lo relacionado con la Gestión y Gobierno de Datos, tiene una trayectoria de más de 20 años impulsado una cultura Data-Driven en las empresas de la mano de sus aliados tecnológicos.

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



