

Gobierno de datos. Un desafío clave para las organizaciones que pretendan destacarse en la era digital.

Área

ACTUALIZACION ACADEMICA

Autores

Ernesto Chinkes – echinkes@gmail.com

Alejandra Regueiro - maregueiro@yahoo.com

1. INTRODUCCIÓN

1.1. El valor estratégico de los datos en la era digital

La denominada cuarta revolución industrial (Schwab, 2017) o era digital, es un momento de la historia de la humanidad dónde se están produciendo importantes cambios económicos, organizacionales, políticos y sociales, que, de manera directa o indirecta están basados principalmente en las tecnologías digitales y el aprovechamiento de los datos.

El contexto en el que deben desenvolverse las organizaciones, en esta era, se presenta con una serie de características que lo evidencia como complejo y desafiante. Es un entorno cambiante, incierto, global, con disrupción fugaz de nuevos jugadores, y la irrupción de nuevos productos, servicios y modelos de negocios que en muchos casos convierten a los existentes en obsoletos en muy poco tiempo.

Muchos de esos grandes retos y desafíos, que tienen las organizaciones, son generados, impulsados o potenciados por el uso que distintos actores de su entorno hacen de los datos; pero, por otro lado, la capacidad que tendrán para mitigar los riesgos y amenazas latentes, o por el contrario, para aprovechar oportunidades únicas dependerá de la forma en que ellos mismos consigan gobernarlos.

El término "big data" ha pasado a ser un concepto muy difundido, y que, de alguna forma, aglutina la toma de conciencia colectiva de que existe una cantidad muy grande de datos que aumenta exponencialmente (sobre personas, objetos y hechos que las rodean), pero también sobre el que puede obtenerse valor (Chinkes et al., 2015).

La abundancia de datos existente se explica por las nuevas y diversas fuentes de datos que fueron surgiendo en las últimas dos décadas, y que han ido creciendo en forma incesante (Chandler, 2015). Por otro lado, la informatización de todos los procesos sobre la mayoría de las actividades que se realizan en la actualidad, implica una interacción

constante con sistemas web, apps, o equipamiento específico, que terminan registrando datos digitales.

Los siguientes factores permiten resumir los principales impulsores de este crecimiento (Chinkes, 2015): 1. El aumento de las transacciones y procesos cada vez más digitales, 2. La omnipresencia de las redes sociales, 3. La inclusión cada vez mayor de sensores incorporados en objetos conectados a Internet (IoT), 4. Una multimedia, cada vez más presente y realista (que favorece la virtualidad), 5. El aprovechamiento de datos externos mediante archivos y APIs, que potencial el uso de los datos por fuera de los límites de las organizaciones, y 6. la nube, la IA y los dispositivos móviles como dinamizadores claves que potencian el resto de los factores.

En esta economía, la exploración de dichos datos es uno de los activos más relevantes del que disponen las empresas; a la par con el capital y el trabajo (Kenneth, 2010); y abren la posibilidad de predecir y tomar decisiones más inteligentes (McAfee & Brynjolfsson, 2012).

Esas afirmaciones, que ya tienen más de una década, se han ido consolidando con la evolución de las tecnologías, y la evidencia del aumento en la cantidad de datos; pero además, la inteligencia artificial ha irrumpido y se viene afianzando con gran celeridad, incorporándose en diversas tareas que antes sólo podían realizar los seres humanos (Hilbert, 2016), pero por otro lado, esa inteligencia artificial está permitiendo el aprovechamiento de los grandes volúmenes de datos, ya que de otra forma sería imposible lograrlo.

Los datos, no sólo son el insumo para generar información que apoye a los decisores, sino también son los que habilitan la automatización de tareas, y la oportunidad de embeber inteligencia en productos, servicios y procesos, para personalizarlos, optimizarlos y/o mejorar la experiencia de los usuarios.

Puede decirse, entonces, que los datos han pasado a ser un activo extremadamente relevante y omnipresente, en toda organización que tenga la pretensión de destacarse; e inclusive algunas veces, si lo único que se pretende es subsistir.

1.2. Organizaciones basadas en datos.

Los directivos de las organizaciones tienen una visión cada vez más favorable hacia el aprovechamiento del big data y la inteligencia artificial, ya que pareciera ofrecerles la posibilidad de una mejora significativa en la toma de decisiones en áreas críticas de sus organizaciones (Davenport & Bean, 2018).

En esa misma línea, existe una postura que viene sumando adeptos entre autores de administración, sobre evolucionar hacia la que denominan organización basada (o dirigida) en los datos (data-driven organization). Según esta línea de pensamiento, estas organizaciones deberían procurar que la mayor parte de las decisiones que se tomen estén sustentadas en los datos. De alguna manera quieren minimizar las decisiones basadas en la intuición y poner de relevancia las decisiones que pueden sustentarse en información (como evidencia).

De una manera un poco más concreta, Berndtsson, M., Forsberg, D., Stein, D., & Svahn, T. (2018), plantean que una organización basada en datos implicaría: usar ampliamente la analítica avanzada (como es el caso de la analítica predictiva para estimar el futuro o la prescriptiva para generar recomendaciones o automatizarlas) a lo largo y ancho de gran parte de la organización. Postulan, que no es sólo la incorporación de tecnología, sino la implementación de un cambio cultural en la forma en que se toman las decisiones.

Nosotros consideramos que un enfoque purista de este tipo de organizaciones, lo único que logrará es diseñar organizaciones irreales, con personas ideales; que no existen. Donde se intente implementar dichas recomendaciones lo más probable es que generen resultados poco satisfactorios o inclusive desastrosos. Este enfoque, de alguna forma, tiene similitud con las escuelas clásicas de economía, en relación al hombre económico y la toma de decisiones totalmente racional. Hay muchas decisiones que no se toman basadas en datos, y que deberán seguir de igual forma, ya sea porque los factores importantes son otros o porque no es factible disponer de todos los datos que permitirían construir la información necesaria (inclusive considerando que nos encontramos en la era del big data).

Sin embargo, nuestra postura, no desmerece en absoluto el gran valor que tienen los datos en esta era, tal como fue planteado en la sección precedente. Los datos son un activo clave para las organizaciones (el más valioso, para muchas). Una institución que mejore su madurez con relación a su aprovechamiento, sin lugar a dudas podrá destacarse respecto de otras.

Berndtsson, M., Forsberg, D., Stein, D., & Svahn, T. (2018), plantean una serie de factores que una organización debe desarrollar, y monitorear, para lograr un buen aprovechamiento de los datos. Dichos factores son el compromiso y enfoque de la alta dirección, disponer de los datos adecuados, la incorporación de las herramientas informáticas, la organización de las áreas dentro de la organización que faciliten la captura, estructuración y aprovechamiento de los datos, y su nivel de incorporación en el proceso de decisión.

La alta dirección debe estar fuertemente involucrada, definiendo los lineamientos de la estrategia que se debe seguir. Si ello no pasa, es muy probable que se terminen implementando soluciones de nichos que faciliten la construcción de "silos" de información.

El segundo factor es el de disponer de los datos. La organización debe asegurarse que éstos cubran un amplio espectro de temas, que le permita abordar mediante su gestión, la mayoría de los desafíos de la organización. Para ello no puede quedarse solo con datos internos, sino que también debe incorporar datos externos; y no solo datos estructurados, sino también aprovechar muchos datos semi estructurados o no estructurados. Por otra parte, es fundamental asegurar la calidad de los mismos, y que éstos sean una fiel representación de la realidad.

El tercer factor es el de las herramientas. Más allá de disponer de herramientas, que permitan hacer análisis básicos y avanzados (descriptivos, predictivos o prescriptivos) también deben posibilitar que los datos se usen de manera masiva (a lo largo y ancho de toda la organización).

El cuarto factor, la organización, se refiere a la forma y el rol que debe tomar el área (o las áreas) de la organización para que puedan ayudar a que se evolucione en esta dirección. Se debe definir si es un área dentro de la de gerencia de tecnología, o un área específica; de quien depende la misma y cuál será su rol (si se dedicará a generar información y la reparte, o si su función es desarrollar la plataforma analítica, o, si se dedica a dar soporte a todas las áreas de negocio, dada que son ellas mismas las que van a generar la información).

Por último, también es necesario reflexionar y trabajar en la forma en que se toman las decisiones. Implica revisar la cultura institucional, para que se le dé relevancia a los hallazgos que pueden mostrar los datos, dónde sea valorada la experimentación con los datos para probar hipótesis, y se elabore una solución analítica que potencie la toma de decisiones comprendiendo en profundidad las posibilidades y límites de los datos para mejorarlas.

El gobierno de datos es una temática relativamente nueva, y como tal, su marco teórico muchas veces es ambiguo, e inclusive contradictorio. Este trabajo, por tanto, pretende ser un aporte tendiente a su clarificación. En la siguiente sección, se revisarán ciertas necesidades que tienen las organizaciones en la actualidad, así como, diversos factores que lo impulsan como un tema crítico y estratégico que debe ser encarado sin dilaciones. En la tercera sección, se pretende delimitar el alcance del gobierno de datos, revisando enfoques y diferenciándolo de otros conceptos relacionados. En la cuarta sección se

describen sus componentes y se sugieren una serie de lineamientos que deben considerarse para su implementación. Por último, en las conclusiones se busca presentar un resumen de las principales ideas tratadas.

2. NECESIDADES Y FACTORES.

2.1. La necesidad de un gobierno de datos

El gobierno de datos, como se profundizará luego, es el que permite dar el marco adecuado para que la organización pueda avanzar y madurar en los factores que se plantearon previamente, en forma sostenida y consistente. Incluye la definición y aplicación de políticas, procesos y responsabilidades asociados a los datos (Mahanti, 2018). Es decir, que define un marco de trabajo, a través de toda la organización, para manejar los datos como un activo estratégico de la institución, definiendo quien tienen los derechos y las responsabilidades en la toma de decisiones sobre este activo (Abraham et al., 2019).

Por lo tanto, su aporte más significativo es la posibilidad de desarrollar la **capacidad** institucional que habilita a la organización a realizar estrategias que de otra forma no serían posibles.

Brinda la capacidad de alinear los datos con los objetivos corporativos, y también habilita el marco necesario para abordar el trabajo colaborativo en torno a los datos, entre los distintos niveles y áreas de la institución. Por último, una buena gobernanza de datos sirve también para mejorar la calidad y seguridad de los datos, así como reduce inconsistencias y minimiza riesgos (Mahanti, 2018).

Para crear valor en las empresas es fundamental aprovechar los datos internos y externos que pueden tener a su disposición, y para ello la gobernanza debe ayudar para fijar un rumbo claro y un adecuado control para que la institución disponga de ellos con la mayor calidad posible. Pero inclusive, cuando se obtengan una cantidad significativa de datos, generalmente luego se usa una porción pequeña. Eso se debe a que la mayoría de los empleados disponen de un acceso y uso muy limitado. Es por ello que, la denominada democratización (como parte de los lineamientos del gobierno de datos) debe ser un tema prioritario para muchas instituciones (Lefebvre et al., 2021).

2.2 Factores que impulsan el gobierno de datos

La necesidad de implementar un programa de gobierno de datos está dada por la concurrencia de distintos factores que se realimentan entre sí. Se detallan a continuación los más relevantes.

Como se mencionó anteriormente, en las organizaciones cuya cultura está "basada en datos", este activo adquiere suma importancia, lo que deriva en la necesidad de que sea "gobernado" al igual que el resto de los activos estratégicos de la organización, para asegurar que su gestión esté alineada con su estrategia.

Por otro lado, el gran aumento de la volumetría, su variedad (datos estructurados, semi estructurados y no estructurados), y la posibilidad de abaratar costos y sumar eficiencia mediante el procesamiento en la nube, generaron arquitecturas de datos cada vez más complejas y heterogéneas; y por consiguiente más difíciles de gestionar.

Los datos provienen de diversas fuentes y tienen formatos cada vez más variados, lo cual también aumenta la complejidad de su integración. Los grandes esfuerzos en la integración de datos que las organizaciones tienen que llevar a cabo, derivan en un detrimento importante de la calidad de los datos (Cheong & Chang, 2007).

Estas arquitecturas, tanto de datos como de aplicaciones (más complejas y heterogéneas) también generan la aparición de nuevos roles y perfiles asociados a la gestión de los datos (data scientist, data product manager, data owner, entre otros). Es decir, que existen más participantes en la organización con incumbencias en las decisiones que deben tomarse a lo largo del ciclo de vida de los datos.

En este contexto, donde crece la complejidad de la gestión de los datos, aumenta la vulnerabilidad y se ven afectadas las tres características fundamentales de la información; integridad, disponibilidad y confidencialidad (Regueiro Pérez, 2022), y por lo tanto es necesario también un marco institucional que asegure la gestión de la calidad de los mismos.

En ocasiones, la gestión de los datos trasciende las fronteras organizacionales, y éstas deben compartir datos con terceros, ya sea por cuestiones legales u oportunidades comerciales, lo cual suma nuevos desafíos respecto de la seguridad, privacidad y calidad.

El avance y difusión de la tecnología en materia de procesamiento masivo de datos y el uso de la inteligencia artificial, disparó numerosas discusiones respecto de la privacidad, lo que llevó a la actualización en materia legislativa de la protección de datos personales a nivel mundial. Este marco normativo sobre el tratamiento de los datos personales, tanto a nivel nacional como internacional, también es un condicionante para la gestión de datos, ya que las organizaciones deben velar por la seguridad de la información y la confidencialidad de las personas cuyos datos gestionan.

Por los factores antes descritos, como el aumento de la volumetría y variedad de los datos, la complejización de las arquitecturas de datos, la actualización de la legislación sobre la

privacidad de las personas, la dificultad de gestionar su integración y calidad, la voluntad de instaurar una cultura basada en datos en las organizaciones; su monetización y su relevancia como activo estratégico; gobernar la gestión de los datos es una tarea cada vez más compleja, y más necesaria, y que también se torna en una cuestión de "compliance" para las organizaciones.

3. CARACTERIZACIÓN DEL GOBIERNO DE DATOS.

3.1 Relación con el Gobierno corporativo y el Gobierno de TI

Weill & Ross, en su obra *"IT Governance: How Top Performers Manage IT Decision Rights for Superior Results"*, toman la definición propuesta por la Organización para la Cooperación y el Desarrollo Económicos, en la publicación de 1999; "Principios de la OCDE para la gobernanza empresarial", por la cual, "el gobierno corporativo proporciona la estructura para determinar los objetivos organizacionales y monitorear el desempeño para asegurar que se logren los objetivos"; y definen al gobierno de TI como "la especificación de los derechos de decisión y el marco de responsabilidad para fomentar el comportamiento deseable en el uso de TI", es decir, los derechos de decisión y responsabilidad que determinan quién toma las decisiones sobre los activos de TI (2004).

El gobierno de TI es una parte integral del gobierno corporativo, [...] y aborda la definición e implementación de procesos, estructuras y mecanismos relacionales en la organización, que permiten, tanto a la gente de negocios como a la de TI, ejecutar sus responsabilidades [...] y la creación de valor comercial [...] (De Haes, Joshi, Huygh, & Jansen, 2017)

Podemos afirmar entonces, que mientras el gobierno corporativo tiene un amplio alcance sobre todos los activos de la organización, el gobierno de TI se enfoca sólo en el conjunto de activos relacionados con la información y las tecnologías que la soportan.

Lo mismo ocurre con el gobierno de datos; éste se enfoca sólo en los activos de datos y su arquitectura. La norma ISO/IEC 38505-1: 2017, define al gobierno de datos como un subconjunto o subdominio del gobierno de IT, el cual, a su vez, es un subconjunto o subdominio del gobierno corporativo (ISO, 2021).

Dentro de la literatura relacionada, también aparece el término "Gobierno de la Información", el cual presenta una sutil diferencia respecto del gobierno de datos. "El gobierno de datos es parte de un gobierno de TI más amplio, que también forma parte de una gobernanza de la información aún más amplia." (Smallwood, 2014).

Según este autor, el gobierno de datos se diferencia del gobierno de la información, ya que éste último se centra en los controles a la información que los sistemas de información generan, mientras que el gobierno de datos se enfoca en la materia prima de ésta, los datos y su calidad (2014). En definitiva, la diferencia entre el gobierno de información y de datos es la que existe entre dichos conceptos (información y datos). Si bien parece importante tener claridad entre ambos, a la hora de incursionar en su gobierno y gestión, no se considera apropiado diseñar esquemas de gobierno por separado, ya que se debe poner especial cuidado en no duplicar esfuerzos ni recursos a la hora de incorporar sus mecanismos. Es decir que el gobierno de datos debe incluir a ambos, datos e información, como activos dentro de su alcance.

Como se puede observar, en todos los casos el "gobierno" busca definir qué y quién puede tomar decisiones respecto de su alcance o área de incumbencia; sea que se refiere a todos los activos (gobierno corporativo), a los activos de TI (gobierno de TI), o los activos de datos e información (gobierno de datos).

Cada uno de estos marcos de gobierno tiene a su vez, dominios específicos sobre los cuales tiene que establecer los derechos de decisión y responsabilidad que le competen; y son implementados mediante mecanismos tales como políticas, procesos, procedimientos, estándares y comités, entre otros.

Los dominios de decisión del gobierno de TI son: los principios de TI, arquitectura de TI, infraestructura de TI, requerimientos de aplicaciones del negocio, inversión en TI y priorización (Weill & Ross). Khatri & Brown, (2010), parten de estos dominios de TI y realizan un paralelo con los datos, y a partir de su análisis, establecen cinco dominios para el gobierno de datos (2010). Estos son: principios de datos, calidad de datos, metadatos, acceso a datos y ciclo de vida de los datos, los cuales abordaremos más adelante.

3.2 Gestión de datos y Gobierno de Datos

De acuerdo con Weill y Ross (2004) la diferencia fundamental entre el gobierno y la gestión (en sentido amplio) es que diseñar y analizar el gobierno requiere abstraerse de las decisiones del día a día de la organización, y concentrarse en establecer la mejor manera de tomar esas decisiones: definir cuáles deben ser tomadas, por quién, y cómo se tomarán y se monitorearán dichas decisiones.

Según Gartner,

La gestión de datos consta de las prácticas, las técnicas de arquitectura de datos, y las herramientas para el acceso y la distribución de datos consistentes a todas las áreas de la organización, para cumplir con los requisitos de consumo de datos de todas las aplicaciones y procesos de negocios.

Mientras que, define al gobierno de datos como “la especificación de los derechos de decisión y un marco de responsabilidad para garantizar el comportamiento adecuado en la valoración, creación, consumo y control de los datos y su análisis” (2023).

Por lo tanto, parece importante hacer foco en la diferencia mencionada, porque ambos términos no son sinónimos ni pueden utilizarse indistintamente, a pesar de que están intrínsecamente relacionados. Una distinción importante para señalar es que el gobierno de datos lo que especifica es quién puede (o debe) tomar determinadas decisiones, y no qué decisión puntualmente se toma.

“El gobierno de datos es importante porque define políticas y procedimientos para asegurar una gestión de datos proactiva y efectiva” (Cheong & Chang, 2007). Es decir, el objetivo del gobierno de datos es establecer los lineamientos necesarios para que las decisiones (estratégicas, tácticas u operativas) sean coherentes y sinérgicas entre sí, y definir quién o quiénes pueden autorizarlas, ejecutarlas o simplemente, deben ser informados o consultados acerca de dicha decisión. En la operatoria diaria se toman numerosas decisiones respecto de la gestión de datos, la importancia de la existencia del gobierno de datos es garantizar que esas decisiones sean tomadas por quienes tienen la potestad de hacerlo, y lo hagan siguiendo los lineamientos institucionales.

Estas decisiones involucran a todas las áreas que deban estar relacionadas con la gestión de datos; sean del negocio, relacionadas con la tecnología que le da soporte, o de apoyo como seguridad informática, áreas legales, etc.

3.3 Enfoques del Gobierno de Datos

Fadler, M., Lefebvre, H., & Legner, llevaron a cabo un estudio donde analizaron la implementación de los mecanismos de gobierno de datos en distintas empresas, y a partir de éste, identificaron tres arquetipos bien definidos: el arquetipo I, caracterizado por poner el foco en la “calidad de los datos maestros”, el arquetipo II, cuya principal característica es la de “habilitar la gestión de datos en toda la empresa”, y el III, que se centra en coordinar los esfuerzos y recursos de la organización con el objetivo de monetizar los datos (2021)

En su análisis, pudieron distinguir características muy similares en cada uno de los grupos. El primero, se caracteriza por focalizarse en la resolución de incidentes de calidad, centrándose en los datos maestros; mediante la asignación de tareas más bien de carácter operativo y con responsabilidades centralizadas. El segundo grupo, correspondiente al arquetipo II, desarrolla más los mecanismos relacionales, ya que su foco principal está en lograr una gestión de datos que satisfaga los requerimientos de negocio, lo cual requiere una estrecha colaboración entre las áreas funcionales y de TI. En cuanto al tercer grupo,

dado que, en éste, los datos son reconocidos como un activo impulsor de nuevos negocios, las organizaciones crean un nuevo puesto: el de director de datos (“*Chief Data Officer*”), para velar por el cumplimiento de los objetivos de la gestión de datos y su monetización. En general, estas últimas tienen mayor experiencia en la gestión de datos y su calidad. (Fadler, Lefebvre, & Legner, 2021)

En base a esto, podría decirse que los tres focos principales hacia los cuales pueden orientarse los esfuerzos del gobierno de datos son la calidad de datos, la democratización de los datos y su monetización. A su vez, también puede concluirse que la gestión de la calidad de datos sería un primer objetivo deseable, al momento de implementar un esquema de gobierno de datos, ya que los otros enfoques también dependen de ella. Por último, se considera que, para tener el impacto planteado en la sección 1 es necesario abordar el tercer enfoque, aunque será relevante tener resueltos previamente aspectos de los otros dos.

4. LA IMPLEMENTACIÓN DEL GOBIERNO DE DATOS

4.1 Componentes principales

Como se mencionó anteriormente, el gobierno de datos al igual que el gobierno de TI, implica definir los aspectos sobre los que deben tomarse decisiones (los “dominios”), quiénes tienen la responsabilidad de decidir, y cómo deberían hacerlo para lograr los mejores resultados. Dentro de las áreas de decisión sobre la gestión de datos, se encuentran los siguientes dominios (Khatri & Brown, 2010):

Dominios de decisión del Gobierno de Datos		Decisiones del dominio
Principios de Datos	declarar/clarificar el rol de los datos como activos	<ul style="list-style-type: none"> ¿Cuáles son los usos de los datos para el negocio? ¿Cuáles son los mecanismos para comunicar los usos comerciales de los datos de forma continua? ¿Cuáles son los comportamientos deseables para emplear datos como activos? ¿Cómo se identifican las oportunidades para compartir y reutilizar los datos? ¿Cómo influye el entorno regulatorio en los usos comerciales de los datos?
Gestión de la calidad de datos	establecer los requisitos del uso intencional de los datos	<ul style="list-style-type: none"> ¿Cuáles son los estándares para la calidad de los datos (precisión, oportunidad, integridad y credibilidad)? ¿Cuál es el programa para establecer y comunicar la calidad de los datos? ¿Cómo se evaluará la calidad de datos y el programa asociado?
Gestión de los metadatos	establecimiento de la semántica o contenido de datos, de modo que sea interpretable por los usuarios	<ul style="list-style-type: none"> ¿Cuál es el programa para documentar la semántica de los datos? ¿Cómo se definirán y modelarán consistentemente los datos para que sean interpretables? ¿Cuál es el plan para mantener actualizados los diferentes tipos de metadatos?
Acceso a los datos	especificar los requerimientos de acceso a los datos	<ul style="list-style-type: none"> ¿Cuál es el valor comercial de los datos? ¿Cómo se llevará a cabo la evaluación de riesgos de forma continua? ¿Cómo se integrarán los resultados de la evaluación con los esfuerzos generales de monitoreo del cumplimiento? ¿Cuáles son los estándares y procedimientos de acceso a datos? ¿Qué es el programa de seguimiento periódico y auditoría de cumplimiento? ¿Cómo se difunde y concientiza sobre seguridad? ¿Cuál es el programa de copia de seguridad y recuperación?
Ciclo de vida de los datos	determinar la definición, producción, retención y retiro de los datos	<ul style="list-style-type: none"> ¿Cómo se realiza el inventario de los datos? ¿Cuál es el programa para la definición, producción, retención y retiro de datos para diferentes tipos de datos? ¿Cómo afectan los problemas de cumplimiento relacionados con la legislación a la retención y el archivo de datos?

Figura 1. Traducción propia. Fuente: Khatri, V., & Brown, C. (2010). *Designing data governance*.

A los ejemplos dados, creemos pertinente considerar también los siguientes, a saber:

Principios de los datos: ¿cómo y con qué frecuencia se revisarán, actualizarán y comunicarán?

Gestión de la calidad de datos: ¿cómo se resolverán los incidentes de calidad y qué medidas se implementarán para evitar su propagación?, ¿en qué puntos y/o repositorios se ubicarán los controles de calidad para mejorar su eficiencia?, ¿cómo se dará certidumbre a los usuarios sobre la calidad de los datos? y ¿cuál es el porcentaje de calidad esperable para cada dominio?

Gestión de los metadatos: ¿cómo logra que los distintos actores de la organización puedan tener una interpretación común de los datos y de la información? ¿cómo se logra una visión compartida de la realidad a través de la representación de los datos?

Acceso a los datos: ¿quién puede solicitar la incorporación de nuevas fuentes y cómo? ¿qué niveles de acceso a los datos se pueden otorgar de acuerdo con los niveles de seguridad y confidencialidad de la información?

Ciclo de vida de los datos: ¿de qué forma se pueden registrar los datos para un mejor aprovechamiento futuro? ¿cómo se actualizan los datos y sus fuentes?

Por otro lado, hay decisiones que están encuadradas dentro del gobierno de TI, pero, sin embargo, dado que tienen un alto impacto en los objetivos propuestos por la estrategia de datos o requieren una *expertise* específica, podrían constituirse dominios de gobierno de datos adicionales a los cinco propuestos.

Uno de ellos es la arquitectura de datos. Como se detalló en forma precedente, presenta una alta complejidad en entornos de big data y analytics, dado que se manejan grandes volúmenes y variedad de datos, y se requieren enormes esfuerzos y recursos para su integración. Por otro lado, las decisiones que se tomen tendrán un alto impacto para la organización a largo plazo.

Otro caso, es el de las decisiones relacionadas con la priorización de proyectos de datos. Si bien son parte de un dominio del gobierno de TI, y los proyectos que se elijan seguramente serán un subconjunto de la cartera total, cuando están asociadas directamente con la estrategia de datos, es una cuestión del gobierno de datos asegurarse que dicho subconjunto sea decidido mediante los mecanismos estructurales y de proceso que los mismos necesitan.

Ambos dominios de decisión, arquitectura de datos y priorización de proyectos de datos, podrían también ser englobados dentro de las decisiones relacionadas con la gestión del ciclo de vida de los datos (según la clasificación mencionada previamente), pero dado que tienen peso específico y alto impacto, sería deseable que se encuentren bien diferenciadas.

4.2. Mecanismos

Una vez identificadas dentro de la organización, cuáles son las áreas de decisión (dominios) que considera pertinentes, es necesario definir los mecanismos que permitan implementar el esquema de derechos de decisión y responsabilidades.

Para ello, pueden diseñarse distintos tipos de mecanismos: estructurales, procedimentales y relacionales. Dentro de los mecanismos estructurales, se encuentran: la estructura organizacional, los roles y responsabilidades y la coordinación y alineamiento. Los mecanismos de procesos se dividen en: modelos de procesos, monitoreo y evaluación, y modelos de desarrollo; y dentro de los relacionales: visión compartida, la colaboración y la transferencia de conocimiento (Baijens, Helms, & Velstra, 2020).

4.1.1 Mecanismos estructurales

a) Estructura organizacional (quiénes – de qué área)

En primer lugar es necesario definir cuáles serán las áreas de la estructura organizacional que posibiliten el desarrollo de la organización, para alcanzar el mayor grado de madurez posible en cuanto al aprovechamiento de los datos. Esto requiere definir áreas, sus funciones, y su nivel de dependencia jerárquica.

Para llevar a cabo las actividades y funciones relativas al gobierno de datos, podría optarse por crear un área específica, o bien podrían crearse subdivisiones dentro de una o varias áreas relacionadas con la gestión de datos. Esta área y/o equipo de trabajo serían los encargados de impulsar el programa de gobierno de datos, y dentro de sus competencias principales estará la definición de los (macro) procesos y estándares de calidad de datos, la convocatoria a la realización de los comités, la designación, coordinación y el acompañamiento de los custodios, gestores y responsables de los datos, entre otras.

Con relación a la gestión de datos, la estructura organizativa puede implementarse de tres formas diferentes. Se puede optar por un modelo centralizado, donde todas las funciones se colocan en una misma área o unidad; por un modelo descentralizado donde las funciones de análisis se distribuyen en diferentes unidades o áreas, o un modelo híbrido (que combina ambas opciones). (Grossman & Siegel, 2014). Cada una de estas opciones presenta ventajas y desventajas.

b) Roles y responsabilidades (quiénes)

La gestión de datos requiere de la existencia y definición de roles específicos, con conocimientos técnicos, del negocio, o ambos, a saber: científicos de datos, arquitectos de datos, analistas de datos, analistas funcionales, desarrolladores, testers, líderes de

proyectos, y analistas de gobierno de datos, entre otros. También surgen dos nuevos roles, el de "custodio de datos" y el "propietario de datos". Hay responsabilidades que están claras en entornos de aplicaciones transaccionales, pero que dejan de estarlo debido a las particularidades de los entornos analíticos (Grossman & Siegel, 2014; Fadler, Lefebvre, & Legner, 2021). En el contexto del gobierno de datos es necesario establecer y delimitar claramente estas responsabilidades.

Custodia de los datos

Según Cheong & Chang, los activos de datos deben ser administrados por el custodio de datos. Es responsable también de su calidad. Su función consiste en resolver los problemas planteados por los usuarios, en respaldar los planes de gestión de datos, de limpieza de datos, de garantizar que los datos sean adecuados para su propósito, y de convertir los planes estratégicos en planes tácticos, además de la gestión de cambios y la gestión de partes interesadas. Una de sus principales funciones es asignar los accesos aprobados por los propietarios de los datos, y dar seguimiento a la aplicación de los estándares, políticas y procedimientos, que fueron establecidos por el gobierno de datos. En general, se espera que quienes ejerzan estos roles cuenten con conocimientos técnicos, y estas funciones pueden estar asignadas dentro de la órbita del área de sistemas o afines.

Propiedad de los datos

Según Fadler y Legner, existen diferencias entre el propietario de los datos (de la fuente), que es responsable de éstos como entrada (o materia prima), y el propietario del producto de datos que es responsable de garantizar el valor del resultado obtenido a partir de éstos. En un entorno de almacenamiento de datos tradicional, el otorgamiento de accesos y los requisitos de calidad son conocidos y pueden establecerse desde un principio. Sin embargo, en el entorno del análisis avanzado es el propietario de los datos quién debe autorizar el acceso a los datos y debe reaccionar continuamente a los requisitos de datos cambiantes, ya que el propósito de los datos se desconoce intencionalmente para permitir la exploración y experimentación (Fadler y Legner, 2020). En general, los propietarios de los datos pertenecen al negocio, y trabajan en colaboración con TI, pero no necesariamente poseen conocimientos técnicos. Los propietarios de datos se definen por dominios del negocio, es decir, no hay un "sólo propietario" de datos en la organización, sino que cada dominio del negocio tendrá uno. Entre sus responsabilidades se encuentra definir cuáles son los orígenes mandatorios de los datos de los dominios que tienen a su cargo, y determinar los riesgos y el nivel de confidencialidad de la información de su dominio.

c) Comités

La creación de comités con la participación de los roles interesados es otro de los mecanismos estructurales, que le permiten a la organización tomar decisiones con una mirada más amplia y representativa (Grossman y Siegel, 2014). Si bien cada organización planteará los comités que considere necesarios de acuerdo con la complejidad de su gestión y estructura, se considera importante destacar dos tipos de comités (como mínimo):

- un comité estratégico que incluya a la alta dirección, a representantes de la organización de TI y diversas partes interesadas del negocio, con la función de priorizar los proyectos; y obtener los recursos necesarios (Grossman y Siegel, 2014). Este comité tendrá a su cargo, la toma de decisiones a nivel estratégico relacionadas con el gobierno de datos (por ej. definición y aprobación de políticas, etc).
- un comité de operación y calidad de datos que garantice que los datos y metadatos de la organización sean precisos, completos y coherentes. Su carácter es más bien operativo, y busca resolver cuestiones puntuales asociadas a la calidad y a la gestión de los datos (Grossman y Siegel, 2014).

Si bien estos son los comités sugeridos en la literatura del tema, cada organización deberá pensar cuáles serían los apropiados de acuerdo con su tamaño, necesidades, cultura, entre otros factores, y definir su periodicidad y participantes.

Nos parece importante destacar que, si bien los comités son necesarios como mecanismo de coordinación para determinados aspectos, las decisiones tratadas de esta forma suelen insumir mayor tiempo. Por lo tanto, es deseable plantear en los comités sólo las decisiones que sean estrictamente necesarias, para no demorar las mismas en forma innecesaria.

4.1.2. Mecanismos de Procesos – (Cómo)

Un objetivo importante del gobierno de datos, además de determinar quiénes serán los encargados de tomar las decisiones en cuanto a los activos de datos, es también indicar cómo, o bajo qué criterios, deberían ser tomadas estas decisiones. “Los procesos de alineación son técnicas para asegurar una participación amplia y eficaz en decisiones de gobernanza y su implementación” (Weill & Ross, 2004). “Los mecanismos de procedimiento complementan las estructuras organizativas y los roles al definir cómo se toman decisiones.” (Fadler, Lefebvre, & Legner, 2021)

Dentro de los procesos propios del gobierno de datos se puede señalar aquellos relacionados con los dominios de decisión mencionados anteriormente en la figura 1. Por ejemplo, una situación que se presente en la organización podría ser la incorporación de una nueva fuente de datos a los repositorios de explotación (datawarehouse o data lakes); este caso corresponde al dominio de gestión del ciclo de vida de los datos, y es el gobierno de datos quién establecerá cómo serán los procesos o procedimientos para solicitar el alta de la nueva fuente, la asignación de permisos de acceso, la asignación de los responsables (custodios, usuarios, dueños, etc) y demás; Luego el área de gestión de datos, será quien incorpore físicamente dicha fuente al repositorio correspondiente. Otra situación podría ser qué estándares de datos utilizar en determinados dominios, el área de gobierno definirá quienes tienen la potestad de definir los estándares elegidos, y el área designada indicará cuáles serán dichos estándares.

4.1.3. Mecanismos Relacionales

Los mecanismos enumerados previamente están estrechamente relacionados, ya que se necesita establecer qué áreas tienen que participar en las decisiones, de qué dominios, con qué alcance, cuáles son los roles requeridos, y con qué procesos llevarlos a cabo; sin embargo, la interrelación entre todos estos mecanismos requiere de un gran esfuerzo entre las partes interesadas. Para ellos es deseable contar con instrumentos que favorezcan la relación y colaboración entre los involucrados. Dentro de los mecanismos de gobernanza relacional, se identifican tres ejes principales: las percepciones compartidas, la colaboración y la transferencia de conocimientos y experiencia (Baijens, Helms y Velstra, 2020). Para las organizaciones es importante prestar especial atención a la gestión del conocimiento, y contar con herramientas informáticas que permitan el resguardo de al menos, parte del conocimiento explícito generado durante los procesos de gestión de datos (Regueiro Pérez, 2022). Estos mecanismos tienen como objetivo almacenar y difundir el conocimiento explícito de los expertos, por ejemplo, mediante la utilización de herramientas de trabajo colaborativas, tales como intranets, archivos compartidos, videoconferencias, herramientas de gestión de proyectos, de gestión de tickets, etc. De esta manera, los interesados pueden consultar y acceder a definiciones, documentos, tableros, informes, videos, cursos de capacitación y demás recursos compartidos, relacionados con la gestión de los datos, la gestión de la calidad de los datos y el gobierno de datos. Por ejemplo, lo acordado en las reuniones de los comités queda registrado en minutas accesibles a los participantes y a los interesados, quienes pueden consultarlas para darles seguimiento. En la intranet pueden publicarse todos los procesos de gestión de datos y las políticas y estándares que la organización utiliza para su tratamiento, entre otras acciones.

4.2. Lineamientos para su implementación

Los mecanismos descritos previamente, como componentes del gobierno de datos, no pueden definirse en forma aislada, sino que deben ser parte de decisiones que los integren de forma sinérgica para lograr los objetivos buscados.

Basados en los mecanismos mencionados, a continuación, se sugieren una serie de acciones, que deben trabajarse en forma coordinada a través de un programa de gobierno de datos.

1. Dominios.

En primer lugar es necesario definir **cuáles son los temas (dominios)**, relacionados con los datos, sobre los cuales deberán tomarse decisiones en la organización. Desde ya que deben ser temáticas relevantes, sobre las cuales tengan importancia dejar claridad sobre quiénes y cómo deben tomarse las mismas.

2. Responsabilidades y procesos

Se deben definir qué **personas** y **grupos** participarán en la toma de decisiones relacionadas con los dominios identificados, y cuál es el **tipo de participación** que tendrán y bajo qué mecanismos de **procesos** deberán hacerlo. Esto varía según el tipo de decisión y las características de cada organización. Se debe buscar que las decisiones las asuman las personas que estén mejor preparadas para tomarlas, pero que también tengan la autoridad necesaria para que luego se conviertan en acción, y que, además puedan realizarse con la agilidad que ese tipo de decisión requiera. Una forma de lograr decisiones consensuadas y con una mirada más amplia es el uso de comités, pero sin embargo deben usarse con cuidado, ya que implica un proceso decisorio más lento.

3. Estructura organizativa para apoyar el gobierno de datos

Será necesario diseñar e implementar una estructura organizativa (sea mediante la generación de un área nueva o la reestructuración de existentes) que asuma responsabilidades con relación al programa de gobierno de datos. Las responsabilidades, poder y dimensiones que se le asignen deben posibilitar la implementación del programa de gobierno (es un tema transversal a toda la organización).

4. Visión compartida en toda la organización

Deben considerarse diversos mecanismos relacionales (como la formación, comunicación y concientización del personal de las distintas áreas sobre el valor de los datos) que aseguren que los datos se trabajan en la organización como un recurso valioso y

compartido para la gran mayoría de los empleados de la misma, así como lograr que las decisiones que se tomen en relación a los datos sean aceptadas por todos.

5. Estrategia de datos

Se deben definir políticas para la gestión de datos. Es decir, que las máximas autoridades deben definir un conjunto de macro decisiones, que le permitan a la organización tener un rumbo en relación las decisiones relacionadas con los datos que se deberán tomar en el futuro. Luego, también es relevante que se realice el seguimiento y evaluación de su cumplimiento, mediante la creación de métricas específicas que permitan su monitoreo. Para ello deben definirse como parte del gobierno; quienes (mecanismos estructurales) y cómo (mecanismos de proceso) se define, implementa y monitorea dicha estrategia.

6. Gestión de una cartera de proyectos

Es clave disponer de una cartera de proyectos relacionada con los datos (posiblemente un subconjunto de la cartera de proyectos de TI), que permita la implementación de la estrategia de datos a través de iniciativas específicas. Para que las mismas estén alineadas y sean factibles, es importante asegurarse, mediante mecanismos estructurales y de procesos del gobierno, cómo se identifican, priorizan, asignan recursos y monitorea la cartera.

5. CONCLUSIONES

En el presente trabajo se ha revisado la relevancia que tienen los datos para las organizaciones en una sociedad que es cada vez más digital. Se ha visto cómo, un contexto donde los datos pasan a ser uno de los activos más relevantes, moldea las denominadas organizaciones basadas en datos, y cómo todo ello se ve potenciado por una serie de factores que hacen imprescindible un gobierno sobre los mismos.

En la tercera sección se ha definido el alcance de este concepto, ya que era uno de los objetivos del artículo. Se ha explicado que el gobierno de datos es un subconjunto del gobierno de tecnologías de la información, el que a su vez lo es del gobierno corporativo; así como su relación con el gobierno de la información. Por otra parte, se lo ha diferenciado del concepto de gestión de datos. El primero trata sobre qué, quiénes y cómo se toman las decisiones; y el segundo son las decisiones en sí. Por último, también se ha dejado claro que existen distintos arquetipos. En particular se ha mencionado el que se focaliza en la calidad de los datos (más operativo), el que lo hace en los datos como apoyo a las áreas del negocio (táctico) y los que pretenden que los datos habiliten oportunidades de negocio (estratégico).

Por último, en la sección cuatro, se ha profundizado sobre cómo puede abordarse la implementación del gobierno de datos. Para ello, se revisaron los dominios y los mecanismos estructurales, de procesos y relacionales, para luego detallar una serie de iniciativas sugeridas que debieran ser encaradas en una organización para implementar un gobierno de datos.

Como se dijo, el artículo ha intentado clarificar ciertos conceptos, pero también dar una guía inicial para quienes lo deben abordar en sus organizaciones. Por último, se espera que esta publicación sirva como base para que este tema siga evolucionando.

6. REFERENCIAS

- Baijens, J., Helms, R., & Velstra, T. (2020). Towards a Framework for Data Analytics Governance Mechanisms. *Proceedings of the 28th European Conference on Information Systems (ECIS), An Online AIS Conference*. 81. Marrakech: ECIS 2020 Research Papers. Obtenido de https://aisel.aisnet.org/ecis2020_rp/81
- Berndtsson, M., Forsberg, D., Stein, D., & Svahn, T. (2018). Becoming a Data-Driven Organisation. (E. 2. (AISEL)., Ed.) 43. Retrieved from https://aisel.aisnet.org/ecis2018_rip/43/
- Chandler, D. (2015). A world without causation: Big data and the coming of age of posthumanism. *Millennium-Journal of International Studies*, 43(3), 833-851.
- Cheong, L., & Chang, V. (2007). The Need for Data Governance: A Case Study.
- Chinkes, E., Fernandez Blanco, M. L., & Coronel, L. (2015, septiembre 19). *BIG DATA: El Dato en un Rol Estratégico, un Desafío para las Soluciones de Gestion de Datos*. Jornada Académica del Departamento Pedagógico de Sistemas 2015, Buenos Aires, Facultad de Ciencias Economicas de la Universidad de Buenos Aires.
- Davenport, T. H., & Bean, R. (2018, febrero 15). Big Companies Are Embracing Analytics, But Most Still Don't Have a Data-Driven Culture. *Harvard Business Review*. <https://hbr.org/2018/02/big-companies-are-embracing-analytics-but-most-still-dont-have-a-data-driven-culture>
- De Haes, S., Joshi, A., Huygh, T. Jansen, S. (2017). Exploring How Corporate Governance Codes Address IT Governance. *ISACA Journal*, 4.
- Fadler, M., & Legner, C. (2020). Who owns the data in the enterprise? Rethinking data ownership in times of big data analytics. En A. f. (AISEL) (Ed.), *European Conference on Information Systems*. 207. Marrakech: ECIS 2020 Proceedings Research Papers. Obtenido de https://aisel.aisnet.org/ecis2020_rp/207/
- Fadler, M., Lefebvre, H., & Legner, C. (2021). Data governance: from master data quality to data monetization. *ECIS 2021 Research Papers*. 155.
- Fadler, M., Lefebvre, H., & Legner, C. (2021). Data governance: from master data quality to data monetization. En A. f. (AISEL) (Ed.), *Twenty-Ninth European Conference on Information Systems*. 155. Marrakech: ECIS 2021 Research Papers. Obtenido de https://aisel.aisnet.org/ecis2021_rp/155
- Gartner. (04 de 2023). <https://www.gartner.com>. Obtenido de <https://www.gartner.com/en/information-technology/glossary>
- Grossman, R., & Siegel, K. (2014). Organizational Models for Big Data and Analytics. (O. D. Community, Ed.) *Journal of Organization Design*, 20-25. doi:10.7146/jod.9799
- Hilbert, M. (2016). Big Data for Development: A Review of Promises and Challenges. *Development Policy Review*, 34(1), 135-174. <https://doi.org/10.1111/dpr.12142>

- Kenneth, C. (2010). Data, data everywhere. *The Economist*, 13.
- McAfee, A., & Brynjolfsson, E. (2012, octubre 1). Big Data: The Management Revolution. *Harvard Business Review*, October 2012. <https://hbr.org/2012/10/big-data-the-management-revolution>
- Regueiro Pérez, M. A. (junio de 2022). Gobernanza de datos: aplicación en una empresa de telecomunicaciones de la Argentina. *Trabajo Final de Posgrado. Universidad de Buenos Aires*. Obtenido de http://bibliotecadigital.econ.uba.ar/download/tpos/1502-2309_RegueiroPerezMA.pdf
- Smallwood, R. (2014). *Information Governance. Concepts, Strategies and Best Practices*. Wiley.
- Schwab, K. (2016). *La cuarta revolución industrial*. Debate. https://books.google.com/books/about/La_cuarta_revoluci%C3%B3n_industrial.html?hl=es&id=BRonDQAAQBAJ
- Weill, P., & Ross, J. (2004). IT Governance: How Top Performers Manage IT Decision Rights for Superior Results.