



Whitepaper

Gobierno de Datos e Inteligencia Artificial para obtener información confiable

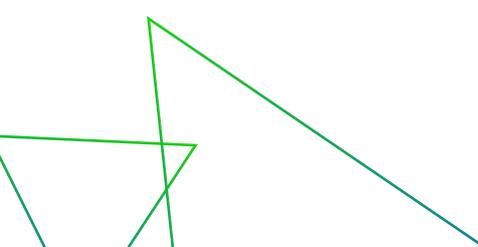
Maximice el valor del intercambio de datos de la
organización de manera imperativa





Tabla de contenido

El análisis requiere un uso ético y confiable de los datos y de AI/ML-----	3
Convergencia de datos y gobernanza de AI -----	4
La nube es el futuro para escalar la AI y el gobierno de datos-----	5
La automatización proporciona agilidad-----	6
Entregue valor, no restricciones-----	10
Conceptos clave y próximos pasos-----	13
Acerca de Ventana Research -----	20
Acerca de PowerData -----	21



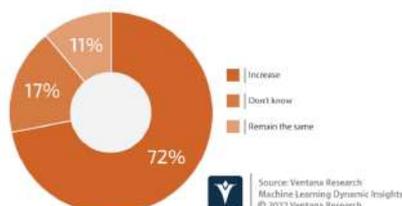
El análisis requiere un uso ético y confiable de los datos y de AI/ML

Los análisis son fundamentales para el funcionamiento eficiente y eficaz de la mayoría de las organizaciones modernas; con ese fin, los análisis confiables requieren datos confiables. Los datos inexactos y no confiables significan que las organizaciones no pueden confiar en los análisis producidos con esos datos. La confiabilidad requiere datos precisos y de alta calidad, pero también un uso apropiado de los datos que cumpla con las políticas de administración de riesgos y acceso de una organización. Los datos de alta calidad permiten tomar decisiones de alta calidad, pero con el volumen cada vez mayor de datos que fluyen dentro de las organizaciones, realizar un control de calidad puede ser extremadamente laborioso. Casi dos tercios (64%) de las organizaciones en nuestra investigación informan que garantizar la calidad de sus datos es una de las tareas en las que pasan más tiempo, solo superada por la preparación de datos para el análisis.

El análisis no solo necesita datos confiables, también la inteligencia artificial (AI). Los modelos de AI deben basarse en datos precisos para producir resultados precisos. Y con muchos modelos de AI que se utilizan en procesos automatizados, es importante invertir para garantizar los niveles de confianza y calidad en los datos que alimentan la IA. El uso ético de los datos y la AI también se ha convertido en una preocupación importante a medida que las organizaciones aumentan sus inversiones en IA y aprendizaje automático (AI/ML). Nuestra investigación muestra que casi las tres cuartas partes (72%) de las organizaciones planean aumentar su uso de AI/ML. Con el aumento del uso, también han aumentado las preocupaciones sobre el uso ético de AI/ML. Esto se debe a que los conjuntos de datos pueden contener sesgos que se manifiestan en los modelos desarrollados con esos conjuntos de datos. En consecuencia, las organizaciones están trabajando para detectar y prevenir sesgos antes de implementarlos en producción. Y los legisladores de Europa y América del Norte han introducido legislación que regularía el uso de AI/ML.

Para abordar estos problemas, las organizaciones están invirtiendo en la creación y expansión de programas de gobierno de datos que van más allá de los informes de cumplimiento, la protección de datos y la gestión de riesgos para centrarse también en brindar valor en forma de datos confiables para análisis, AI y más. Los programas de gobierno bien ejecutados hacen que los procesos de datos de las organizaciones no solo sean más sólidos sino también más efectivos. Capturan políticas comerciales, admiten flujos de trabajo y brindan colaboración entre los participantes. Los catálogos de datos son un componente central de los programas integrales de gobernanza, al igual que la calidad de los datos, la gestión de datos maestros, el intercambio de datos y la democratización. Cuando los datos son más accesibles, compartibles y mejor entendidos, las organizaciones pueden aprovecharlos al máximo para ejecutar sus procesos comerciales.

Planned Changes in Machine Learning Use
Most expect use will increase



Convergencia de datos y gobernanza de AI

El gobierno de datos es una de las principales preocupaciones de las organizaciones en la actualidad. Con regulaciones cada vez mayores, como el Reglamento general de protección de datos (GDPR) y la Ley de privacidad del consumidor de California (CCPA), las propuestas de la Unión Europea sobre la gestión de datos y la AI, y un enfoque cada vez mayor en los informes ambientales, sociales y de gobernanza (ESG), las organizaciones enfrentan más supervisión externa de sus prácticas de gobierno de datos. El riesgo de multas significativas asociadas con estas y otras regulaciones, junto con los requisitos de cumplimiento interno de las organizaciones, ha llamado más la atención sobre las prácticas de gobierno de datos e AI. Anticipamos que hasta 2024, la principal preocupación de más de las tres cuartas partes de los directores de datos será controlar la privacidad y la seguridad de los datos de su organización. Esto incluye preocupaciones incipientes pero crecientes con la AI. Aun así, aunque el interés ha aumentado, una organización puede estar perdiendo el bosque por los árboles en su estrategia de gobierno de datos e AI..

La estrategia de gobernanza de datos e AI es tan buena como su eslabón más débil. Desafortunadamente, tres cuartas partes (75%) de las organizaciones todavía confían en los procesos de datos basados en hojas de cálculo y casi seis de cada 10 (59%) confían en códigos y secuencias de comandos personalizados. Los datos también se transfieren a los procesos de análisis e inteligencia artificial que deben gobernarse con la misma disciplina que los procesos de datos. Por lo tanto, una estrategia holística de gobierno de datos considera el espectro completo de sus procesos de datos, análisis e inteligencia artificial.

La gobernanza integral de AI/ML requiere consideraciones adicionales más allá de la gobernanza de datos, como si se usaron los datos de capacitación apropiados para el modelo y determinar si los datos están documentados, son justos e imparciales, y están aprobados. Se necesita un monitoreo continuo del modelo para asegurarse de que continúe funcionando bien y para detectar la deriva de datos. El comportamiento, la salida y el linaje del modelo deben ser comprensibles para que la organización pueda colaborar en torno al uso del modelo y garantizar que se utilice correctamente. Esto es especialmente importante pero también un desafío continuo porque las organizaciones luchan por encontrar las habilidades de AI/ML que necesitan, con menos de una cuarta parte (23%) informando que tienen las habilidades necesarias. Comprender y catalogar los modelos de AI también ayuda a promover la reutilización de los análisis aprobados, impulsando una mayor eficiencia y aumentando el valor del programa de gobierno.



La nube es el futuro para escalar la AI y el gobierno de datos

La nube se está convirtiendo en el futuro de los datos y la AI. Un número creciente de organizaciones está utilizando la nube para sus procesos de datos y análisis: casi seis de cada 10 (59 %) usan la nube hoy y casi nueve de cada 10 (89 %) planean usar la nube en el futuro. La nube proporciona la escala necesaria para los procesos relacionados con datos y análisis y elimina muchas tareas administrativas y de mantenimiento, lo que libera al personal de TI para que se concentre en objetivos de mayor valor. La combinación de estos beneficios da como resultado un tiempo de valorización más rápido para los procesos relacionados con los datos.

La escala es fundamental para muchos procesos de AI/ML. Las correlaciones identificadas en los modelos AI/ML se determinan a nivel detallado; por ejemplo, si dos artículos se compraron juntos no se puede determinar a ningún nivel agregado. Además, la identificación de patrones en los datos a menudo requiere la presencia de muchas instancias, como en escenarios de detección de fraude donde debe haber suficientes casos de comportamientos fraudulentos para poder detectar los patrones. Por lo tanto, el desarrollo de modelos AI/ML requiere grandes cantidades de datos detallados y recursos informáticos. Además, ejecutar los modelos AI/ML en nuevos datos a medida que llegan requiere suficientes recursos informáticos para realizar los cálculos en tiempo real.

La realidad de “nube” significa multinube e híbrida. Por una variedad de razones, muchas organizaciones implementan sistemas a través de múltiples proveedores de nube y/o configuraciones locales. Nuestra investigación muestra que más de cuatro de cada 10 organizaciones (42%) operan sus procesos de datos y análisis a través de múltiples proveedores de nube. Y tres cuartas partes (75%) de los que operan en una nube también operan en un híbrido de local y nube. Por lo tanto, la mayoría de las organizaciones necesitan un sistema de gobierno y administración de datos que abarque todas estas configuraciones y brinde flexibilidad para las cargas de trabajo. Esta flexibilidad ayuda a evitar el bloqueo del proveedor a una sola plataforma y permite la rentabilidad en las diferentes plataformas.

“

La nube proporciona la escala necesaria para los procesos relacionados con datos y análisis y elimina muchas tareas administrativas y de mantenimiento, lo que libera al personal de TI para que se concentre

La automatización proporciona agilidad

Las operaciones de datos (DataOps) es la disciplina de la automatización de procesos de datos de forma flexible. Más de un tercio (37 %) de las organizaciones informan que su tecnología de preparación de datos no es flexible ni se adapta al cambio, lo que compromete tanto su agilidad como su capacidad para usar los datos de manera eficaz. DataOps ha surgido como un enfoque para abordar estos problemas y, al mismo tiempo, mantener la confiabilidad y el gobierno adecuados que requieren las organizaciones. DataOps toma su nombre de DevOps, que se enfoca en la entrega continua de aplicaciones de software frente a cambios constantes. Las organizaciones ahora están aplicando esos mismos conceptos a la entrega de datos.

Para completar el proceso, las organizaciones deben extender la automatización a AnalyticOps, incluido MLOps. Las canalizaciones de datos automatizadas y repetibles son fundamentales para los procesos de análisis, al igual que lo son para DataOps. A medida que cambian las fuentes y los objetivos, es necesario modificar las canalizaciones para adaptarse a los cambios; mediante el uso de metadatos, muchas de las modificaciones necesarias pueden automatizarse, eliminando así la necesidad de intervención manual. Esperamos que para 2024, un tercio de las organizaciones adopten un enfoque de operaciones analíticas similar e integrado con sus procesos de operaciones de datos para mejorar la capacidad de respuesta y la agilidad. Las organizaciones también deben crear MLOps automatizados y repetibles para respaldar los esfuerzos de AI/ML.

La implementación de esta automatización proporciona la agilidad que necesitan las organizaciones en la actualidad. Cuando se hace correctamente, la integración de los enfoques DataOps, AnalyticOps y MLOps con un buen marco de gobierno ayuda a que los datos sean más accesibles y los procesos de datos más receptivos. Como resultado, la organización puede responder más rápidamente a los cambios en los requisitos comerciales.



Entregue valor, no restricciones

El gobierno de datos debe habilitar a la organización, no deshabilitarla. Con demasiada frecuencia, las organizaciones centran sus esfuerzos de gobierno de datos en el cumplimiento, la evaluación de riesgos y los controles que limitan la exposición y el uso de los datos. Si bien estos son objetivos válidos, existe una mayor oportunidad de permitir que la organización logre sus objetivos comerciales. Cuando los programas de gobierno se diseñan desde esta perspectiva, pueden ser muy empoderadores y dar como resultado que las organizaciones tengan acceso oportuno a datos bien entendidos y análisis confiables.

Nuestra investigación muestra algunos de los beneficios de los programas de buen gobierno. Los programas de buen gobierno incluyen un marco de políticas y administración, incluidos los flujos de trabajo apropiados para administrar e implementar las políticas. También incluyen la capacidad de vincular políticas y proporcionar el contexto empresarial a los datos. Los catálogos de datos ayudan a proporcionar ese contexto: dónde se originaron los datos, cuándo se actualizaron por última vez y cómo pueden y deben usarse. Los catálogos de datos también mejoran la gobernanza y reducen el riesgo de errores al indicar qué fuentes están aprobadas y certificadas. Quizás el beneficio más fundamental de un catálogo de datos es que facilita la búsqueda y el hallazgo de la información necesaria para un análisis. Nuestra investigación encuentra que los usuarios del catálogo de datos están significativamente más satisfechos (68%) con la capacidad de su organización para analizar big data que aquellos que no usan catálogos de datos (39%).

El buen gobierno de datos implica reunir a las partes interesadas comerciales y técnicas para colaborar y acordar políticas de datos, definiciones, procesos de flujo de trabajo y procedimientos. Un buen gobierno significa mejores resultados para la organización. Cuando las partes interesadas entienden y confían en los datos, es más probable que los usen y se beneficien del valor que pueden proporcionar. Los datos se comparten en toda la organización con medidas de seguridad apropiadas, pero flexibles, lo que permite a los usuarios comerciales y técnicos acceder a los datos en forma de autoservicio. Y, si los datos son consistentes en toda la organización (debido a la gestión de datos maestros y los esfuerzos de calidad de los datos), habrá más alineación y se perderá menos tiempo corrigiendo las inconsistencias. Cuando los modelos AI/ML se entiendan y gobiernen, se implementarán correctamente y se mantendrán para lograr la máxima eficacia. Además, los esfuerzos de gobierno garantizarán la protección de datos, la gestión de riesgos y el cumplimiento de los requisitos normativos e internos, minimizando los riesgos de multas y daños a la reputación de la organización.





Conceptos clave y próximos pasos

- La gobernanza de datos, análisis y AI/ML está convergiendo para abordar las brechas presentes en los programas de gobernanza menos integrales.
- Automatice los procesos de datos, incluido el gobierno, mediante el diseño de programas que abarquen DataOps, AnalyticOps y MLOps.
- Reconocer que el gobierno de AI/ML crea requisitos adicionales más allá de las capacidades estándar de gobierno de datos.
- Explorar tecnologías y procesos de gobierno confiables que crean una garantía de confianza de extremo a extremo.
- Con la gobernanza para brindar un uso confiable de los datos y la AI, las organizaciones podrán maximizar el valor de los datos que administran

Acerca de Ventana Research

Ventana Research es la firma de servicios de consultoría e investigación de tecnología empresarial de referencia más autorizada y respetada. Brinda información y orientación experta sobre tecnologías convencionales y disruptivas a través de un conjunto único de ofertas basadas en investigación que incluyen evaluaciones comparativas de investigación y evaluación de tecnología, talleres educativos y nuestros servicios de investigación y asesoramiento, Ventana On-Demand.

Este Whitepaper es una traducción al español de la versión original [Data and AI Governance for Trusted Insights and Data Sharing](#) 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

