



Puntos destacados

- Interfaz gráfica de usuario fácil de usar con conectores de datos “arrastrar y soltar”
 - Desarrollo acelerado a través de funciones de transformación integradas
 - Limpieza de datos integral y personalizable
 - Sondeo y clasificación de datos automatizados para una mejor gobernanza
-

IBM BigInsights BigIntegrate y BigInsights BigQuality

Autonomía de información para su ecosistema de Big Data

El análisis de Big Data ha surgido como un factor impulsor clave para obtener valor y otras ventajas organizacionales a partir del rápido crecimiento de los volúmenes de datos dentro de las empresas. Aunque muchas organizaciones expertas ya utilizan Hadoop por sus inigualables capacidades analíticas y su costo conveniente, Hadoop en sí no fue diseñada para la integración de datos. La integración de datos exitosa tiene sus propios requisitos específicos, como el soporte para la gobernanza, la administración de metadatos, la calidad de los datos y los distintos estilos flexibles para su entrega.

Sin embargo, muchas capacidades esenciales de integración de datos en Hadoop no están completamente desarrolladas o simplemente no existen, y deben ser abordadas en los proyectos vigentes. Más importante aún, muchas organizaciones se están dando cuenta de que Hadoop por sí sola no es lo suficientemente robusta para una integración de datos exitosa, y las capacidades y el esfuerzo que requiere desarrollar una solución integral son insostenibles.

Para ayudar a las organizaciones a aprovechar por completo el alcance y el compromiso de Hadoop, IBM ha desarrollado una solución poderosa y rentable que admite proyectos críticos e iniciativas analíticas claves dentro de todo el ecosistema de datos. Desarrolladas mediante IBM® InfoSphere® Information Server, IBM BigInsights® BigIntegrate y BigInsights BigQuality proporcionan de punta a punta las capacidades de gobernanza e integración de información que permiten a las organizaciones lo siguiente:

- Comprender los datos
- Limpiar, supervisar, transformar y entregar datos
- Colaborar en toda la organización y reducir la brecha entre negocios y TI

Al aprovechar las capacidades flexibles de integración que ofrecen IBM BigInsights BigIntegrate y BigInsights BigQuality, las organizaciones pueden garantizar que la información que impulsa las iniciativas comerciales y estratégicas (desde sistemas de registro hasta sistemas de visión) sea confiable, consistente y administrada en tiempo real.





IBM BigInsights: análisis en grandes conjuntos de datos

BigInsights ayuda a las organizaciones a buscar patrones al observar todos los datos, no solo las muestras. Para lograr análisis en conjuntos de datos muy grandes, BigInsights proporciona una gama completa de algoritmos de aprendizaje automático y de gran escalabilidad (anteriormente denominada SystemML en IBM Research) que son programables mediante el conocido lenguaje R. Para desarrollar un modelo estadístico exitoso es fundamental un conjunto de datos lo suficientemente grande. La existencia de mejores modelos permite contar con capacidades predictivas más precisas y, en última instancia, genera mejores resultados. Siendo parte de la familia de productos de BigInsights, BigInsights BigIntegrate y BigInsights BigQuality ofrecen los medios para ensamblar, transformar, limpiar y poner los datos a disposición para el análisis hacia abajo.

BigInsights BigIntegrate y BigInsights BigQuality: integración de datos nativos

BigInsights BigIntegrate y BigInsights BigQuality permiten a las organizaciones integrar y transformar cualquier dato con Hadoop, al aprovechar tanto las fuentes de datos nuevas como las existentes para las iniciativas de Big Data.

Una interfaz gráfica fácil de usar ayuda a las organizaciones a transformar rápidamente la información de toda la empresa. La paleta repleta de funciones incluye conectores a una variedad muy amplia de fuentes de datos, que incluye todas las principales bases de datos tradicionales en todas las plataformas (distribuidas, IBM z/OS®, etc.), tipos de archivos, aplicaciones empresariales (como Oracle, Salesforce.com y SAP) y Hadoop. Los programadores ponen a disposición estas fuentes de datos a través de funciones simples de “arrastrar y soltar” y ajustan los parámetros de configuración para maximizar el rendimiento y la flexibilidad.

BigInsights BigIntegrate y BigInsights BigQuality ayudan a mejorar el rendimiento y son altamente escalables. La conectividad nativa con las fuentes de datos más utilizadas está disponible a través de interfaces específicas, y cientos de funciones de transformación integradas ayudan a acelerar el desarrollo. Como resultado, muchas tareas de codificación se pueden completar en una fracción del tiempo empleado en la codificación manual. En un estudio, a un programador le llevó dos días terminar una tarea que anteriormente implicaba escribir durante 30 días casi 2000 líneas de código sin la documentación adecuada, sin reutilización y con desafíos de mantenimiento importantes.

Las últimas características de Hadoop utilizan el mismo entorno de diseño gráfico sencillo que los clientes de IBM ya han utilizado para desarrollar millones de aplicaciones de integración. Dentro de este entorno, las organizaciones pueden desarrollar nuevas aplicaciones de uso intensivo de la información basadas en Hadoop sin necesidad de volver a capacitar a su equipo de desarrollo sobre los nuevos lenguajes emergentes que requieren codificación manual y carecen de soporte para la gobernanza.

Colaboración estrecha con herramientas de diseño totalmente integradas

Una experiencia de diseño integrada y un marco de metadatos común ayudan a garantizar la colaboración cruzada, ya que todos los grupos trabajan a partir de los mismos metadatos consistentes. A través de la interfaz gráfica, los usuarios finales pueden automatizar el proceso de especificación de diseño y enriquecer los metadatos con una visión empresarial vinculada e integral de los activos técnicos. Las herramientas de diseño integradas permiten a los programadores desarrollar trabajos que combinen los procesos de integración de datos y de calidad de datos en un único artefacto de trabajos para alcanzar el más alto nivel de eficacia. Con un enfoque de diseño común, las organizaciones también tienen la flexibilidad para agregar garantía de calidad al proceso, donde resulte más conveniente.



Además, BigInsights BigIntegrate y BigInsights BigQuality ofrecen características de gobernanza e integración líderes en la industria. Los trabajos de transformación e integración de datos se pueden ejecutar directamente dentro de un clúster de Hadoop, así las características poderosas de conectividad de datos, transformación, limpieza, optimización y entrega de datos están disponibles de inmediato para ejecutarse dentro de la plataforma de Hadoop. Este enfoque ayuda a resolver algunos de los desafíos de datos más complejos, dado que minimiza el movimiento de datos y acelera el tiempo para proporcionar valor empresarial.

BigInsights BigQuality: limpieza de datos nativos para casos de uso de Hadoop

BigInsights BigQuality permite a las organizaciones limpiar los datos y supervisar su calidad de manera continua, al garantizar que los datos se conviertan en información confiable. Los datos confiables son esenciales para obtener valor empresarial de los recursos de datos existentes, de manera que las organizaciones puedan mejorar el proceso de toma de decisiones y agilizar las operaciones comerciales.

La función de limpieza de datos integral y personalizable por lotes y en tiempo real automatiza la investigación de datos de origen y la clasificación de datos. Esta automatización permite que el equipo de administradores de datos gestione los activos de datos de manera eficaz y responda ante los objetivos comerciales con datos confiables. Las funciones de automatización también ayudan a procesar y a administrar el aumento repentino de datos que ahora se envían a Hadoop, que pueden escalar si es necesario para seguir el ritmo del volumen.

Los volúmenes de datos de rápido crecimiento también significan que, a medida que la comunidad de proveedores y consumidores de datos se expande, también lo hace la incertidumbre sobre la sensibilidad de los datos y sobre cómo cumplir con los requisitos regulatorios exigidos. BigInsights BigQuality ayuda a sondear las diversas fuentes de datos, como Hadoop, y garantiza la ubicación y el uso correctos de los datos de acuerdo con las políticas predefinidas. Además, BigInsights BigQuality ayuda a estandarizar y a correlacionar los registros de acuerdo con reglas empresariales personalizables. Al establecer métricas de calidad de datos alineadas con los objetivos empresariales, las organizaciones pueden descubrir rápidamente los problemas de calidad y establecer un plan de reparación. Estas capacidades son claves para mantener un programa de gobernanza de datos eficaz.

¿Por qué IBM?

Las nuevas capacidades de BigInsights BigIntegrate y BigInsights BigQuality para Apache Hadoop expanden el portafolio de soluciones de Big Data de IBM Analytics, que abarca software, servicios, investigación y hardware. El portafolio combina las herramientas tradicionales de almacenamiento de datos con las técnicas y tecnologías actuales de Big Data (como Hadoop, stream computing, exploración de datos, análisis avanzados, integración empresarial y computación cognitiva de IBM Watson™) para crear una solución integrada que pueda incorporar el Big Data y los análisis en cualquier organización.

Cómo obtener más información

Para obtener más información acerca de BigInsights BigIntegrate y BigInsights BigQuality y otras soluciones de plataformas de integración de datos, comuníquese con su representante de ventas de IBM o haga clic aquí: ibm.com/software/data/information-integration-governance



© Copyright IBM Corporation 2015

IBM Analytics
Route 100
Somers, NY 10589

Producido en los Estados Unidos de América
Septiembre de 2015

IBM, el logotipo de IBM, ibm.com, BigInsights, IBM Watson, InfoSphere y z/OS son marcas comerciales de International Business Machines Corp., registradas en diversas jurisdicciones a nivel mundial. Otros nombres de productos y servicios podrían ser marcas comerciales de IBM o de otras compañías. Hay una lista actualizada de las marcas comerciales de IBM disponible en la web en “Copyright and trademark information” en ibm.com/legal/copytrade.shtml

Este documento está actualizado a la fecha inicial de su publicación y puede ser modificado por IBM en cualquier momento. No todas las ofertas están disponibles en todos los países donde opera IBM.

LA INFORMACIÓN DE ESTE DOCUMENTO SE PROPORCIONA “TAL CUAL” SIN GARANTÍAS DE NINGÚN TIPO, YA SEAN EXPRESAS O IMPLÍCITAS, INCLUYENDO CUALQUIER GARANTÍA DE COMERCIALIZACIÓN O DE IDONEIDAD PARA UN PROPÓSITO ESPECÍFICO Y CUALQUIER GARANTÍA O CONDICIÓN DE NO VIOLACIÓN. Los productos de IBM están garantizados según los términos y condiciones de los acuerdos bajo los cuales se brindan.

El cliente es responsable de garantizar el cumplimiento de las leyes y regulaciones aplicables. IBM no proporciona asesoramiento legal ni declara o garantiza que sus servicios o productos asegurarán que el cliente cumpla con cualquier ley o regulación.



Por favor, recicle