

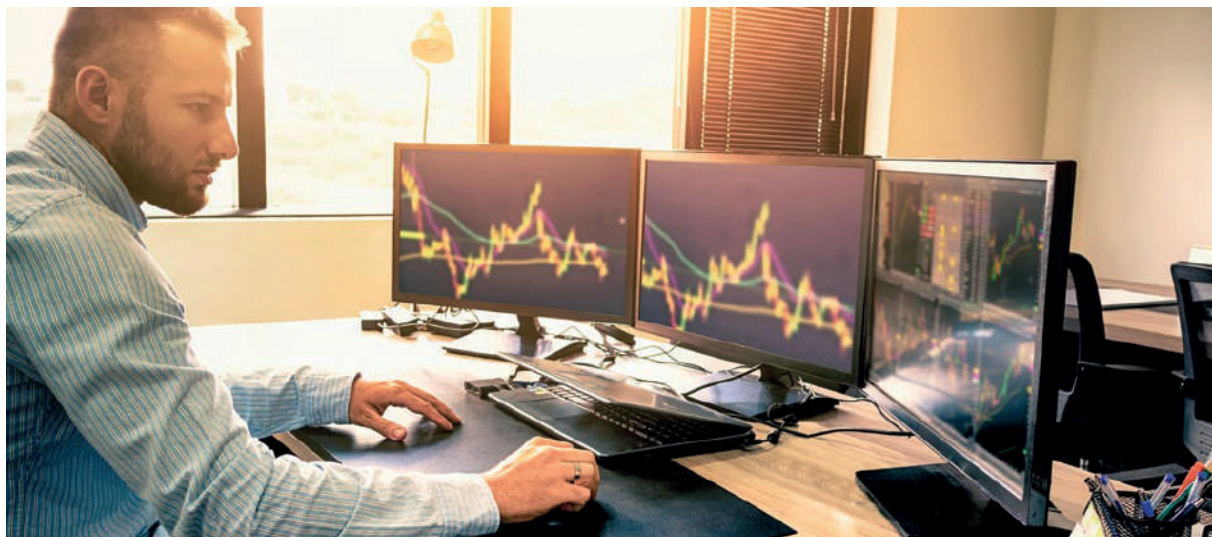
CÓMO CONSOLIDAR LAS APLICACIONES DE GESTIÓN DE RIESGOS Y CUMPLIMIENTO NORMATIVO MEDIANTE UNA PLATAFORMA DE DATOS UNIFICADA

Resumen ejecutivo

Las instituciones financieras han implementado continuamente, muchas aplicaciones distintas para la gestión de riesgos y el cumplimiento normativo en respuesta al aumento en los requerimientos regulatorios y la necesidad de menores riesgos operativos.

Sin embargo, últimamente muchas organizaciones han comenzado a unir estas aplicaciones separadas mediante una plataforma única, integral y transversal para toda la empresa. Hacerlo les permite obtener una visión más precisa y completa de sus datos, identificado los riesgos operacionales, de cumplimiento y las actividades sospechosas de forma más rápida y exhaustiva, para cumplir mejor con las nuevas regulaciones o sus modificaciones, además de simplificar la gestión y mantenimiento en general.

En este documento técnico se describen los factores que impulsan a consolidar las aplicaciones y los datos para la gestión de riesgo y el cumplimiento normativo. Luego, se discuten las fortalezas y limitaciones de diferentes tecnologías, incluyendo *data warehouse* tradicionales y nuevas tecnologías de big data. Después, se presenta una nueva tecnología: InterSystems IRIS Data Platform™, la que complementa los ambientes de gestión de datos ya existentes de las organizaciones para entregar una visión panorámica y unificada de todos los datos empresariales y una capa de procesamiento de ultra alto rendimiento y escalabilidad para realizar tareas de procesamiento complejas sobre datos distribuidos para alimentar aplicaciones y entregar respuestas rápidas y certeras a preguntas inesperadas de parte de reguladores y usuarios de negocios.



Introducción

Durante las últimas dos décadas, las instituciones financieras han implementado muchas aplicaciones diferentes para soportar varias iniciativas de gestión de riesgo y cumplimiento normativo a través de múltiples líneas de negocio, regiones geográficas y entidades legales. Especialmente desde los eventos de 2001 y 2008, las nuevas regulaciones han generado una explosión en la implementación de soluciones específicas.

Las empresas de servicios financieros han creado un mosaico de aplicaciones desconectadas, cada una con su propio conjunto de reglas y datos almacenados en diferentes formatos y representaciones para cumplir con esos requerimientos de información intensificados. Como resultado, muchas organizaciones encuentran difícil obtener una visión completa y precisa a nivel transversal de sus datos para gestionar el riesgo, cumplir con las regulaciones y llevar a cabo la vigilancia

especialmente a medida que los volúmenes de datos crecen y se implementan nuevas regulaciones y cambian las existentes. Esto ha expuesto a las organizaciones a riesgos financieros y de reputación, además de multas y penalizaciones de los reguladores.

Muchas empresas están luchando por seguir el ritmo del volumen, variedad y veracidad de sus datos, a la vez que intentan analizar años e incluso décadas de datos empresariales para cumplir con la avalancha de nuevas y cambiantes regulaciones. Como si no fuera suficiente, las organizaciones también deben lidiar con la velocidad de los datos para soportar casos de uso en tiempo real. También deben cumplir con los reportes de liquidez intradía y otros requerimientos que requieren disminuir el tiempo entre transacciones, el conocimiento y la acción.

Iniciativas de big data y data warehouse

Algunas empresas han comenzado a evaluar e implementar Apache Hadoop, Apache Spark, y otras tecnologías de big data mientras trabajan en consolidar su gestión de riesgos y aplicaciones de Cumplimiento normativo para crear una plataforma única, consistente y de nivel empresarial. Estas tecnologías de big data pueden entregar algunos beneficios para las iniciativas de consolidación; por ejemplo, almacenamiento rentable de datos para satisfacer búsquedas simples y ciertos tipos de consultas. Por ejemplo:

- HDFS (Sistema de archivos distribuido Hadoop) y otros repositorios de datos pueden proporcionar almacenamiento escalable y rentable para grandes volúmenes de datos.
- MapReduce y otras herramientas como Impala y Hive pueden entregar respuestas a ciertos tipos de preguntas a partir de los datos.
- Spark, con su arquitectura en memoria, puede entregar altos niveles de rendimiento para ciertos tipos de consultas analíticas.

Sin embargo, para muchos escenarios, estos marcos de trabajo y herramientas de big data no son suficientes, llevando más bien a un “pantano de datos” en lugar de un *data lake*, donde se almacenan grandes cantidades de fuentes de datos dispares pero no se aprovechan fácilmente y sigue siendo difícil o imposible encontrar respuestas certeras a preguntas complejas o inesperadas.

Por ejemplo, Hadoop y Spark pueden distribuir una búsqueda o una tarea de procesamiento simple de manera eficiente a través de múltiples nodos. Sin embargo, suelen tardar horas, o no completar tareas más complejas que requieran unir datos desde diferentes tablas o nodos. Esto es especialmente importante a medida que las organizaciones buscan hacer nuevos usos de la información distribuida. Por ejemplo, intentar correlacionar múltiples actividades y productos asociados con la cuenta del cliente que están distribuidos en tablas múltiples en diferentes nodos usando MapReduce o Spark suele agotar el tiempo de espera, sin llegar a una respuesta. Hadoop presenta límites adicionales para iniciativas de riesgos consolidados y de cumplimiento, por ejemplo, que involucren seguridad, gobernanza de datos y procedencia.

Las organizaciones financieras están descubriendo que estos marcos de trabajo y tecnologías de big data, si bien dan algunas ventajas para almacenar grandes conjuntos de datos y para ciertas clases de herramientas de procesamiento, no son suficientes.

Algunas empresas están intentando utilizar tecnologías de *data warehouse* tradicional para estas iniciativas de consolidación. Pero los *data warehouses* requieren que los datos sean pre-procesados y estructurados, lo que puede eliminar información valiosa, requerir almacenamiento y recursos de procesamiento costosos, tener un escalamiento vertical (en lugar de arquitecturas de escalabilidad horizontal) y limitar la cantidad de clientes simultáneos. Y es difícil o imposible estructurar los datos por adelantado para responder a cualquier posible consulta puntual que puedan hacer las entidades de regulación.

InterSystems IRIS Data Platform para la gestión de riesgos y el cumplimiento normativo

InterSystems IRIS Data Platform entrega capacidades fundamentales para ayudar a las compañías a consolidar sus aplicaciones de gestión de riesgos y cumplimiento normativo en una sola plataforma.

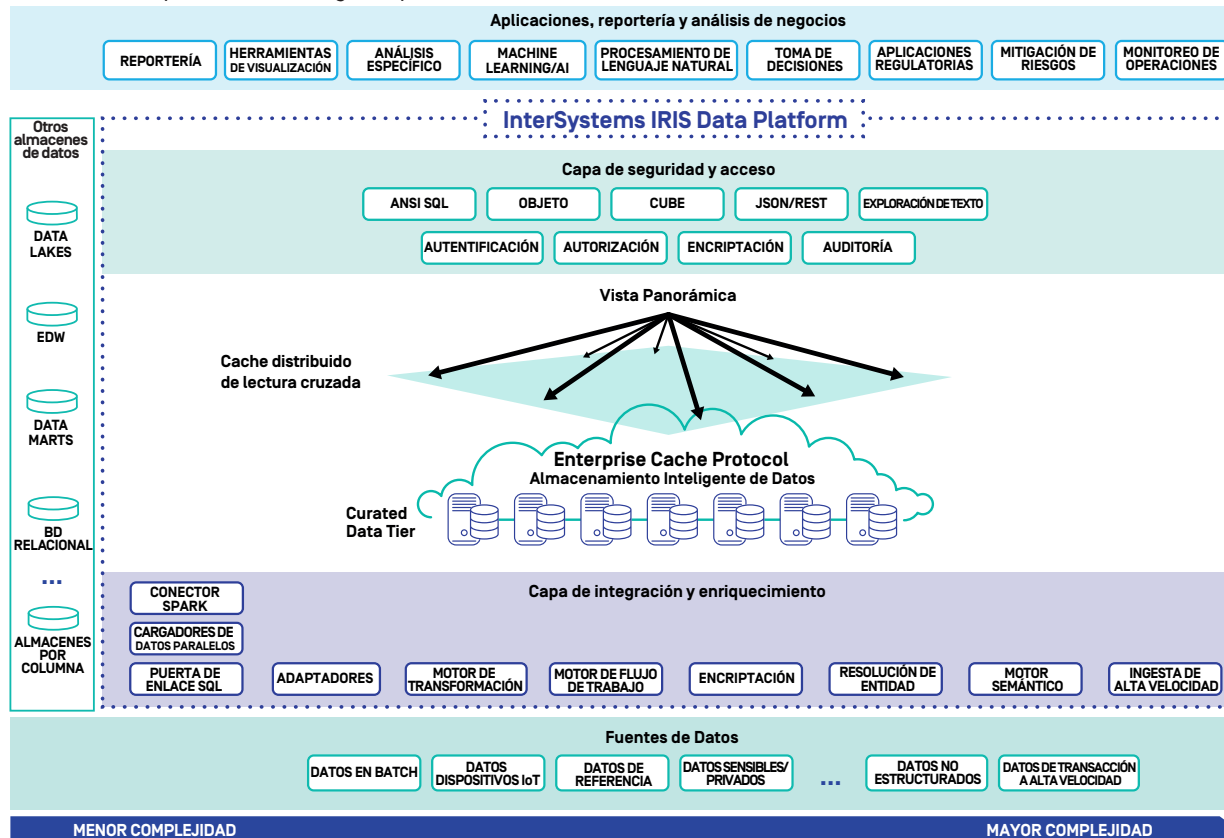
Se integra fácilmente a la infraestructura de gestión de datos ya existente de una organización, incluyendo las aplicaciones heredadas, *data warehouses*, tecnologías de big data y *data lakes*. Permite a las organizaciones obtener una vista panorámica y unificada de todos los datos de múltiples fuentes en toda la organización, mediante una capa de almacenamiento en caché distribuida y en tiempo real, que entrega acceso seguro a datos desde diferentes fuentes.

Por ejemplo, Apache MapReduce y Spark distribuyen eficientemente una búsqueda o una tarea de procesamiento simple sobre múltiples nodos. Sin embargo, suelen tardar horas, o simplemente no logran completar tareas más complejas que requieran la unión de datos distribuidos.

InterSystems IRIS, entrega un rendimiento ultra alto, una capa de procesamiento de escalamiento horizontal para realizar tareas de procesamiento complejas en *batch* (por lote) y en tiempo real, sobre grandes conjuntos de datos distribuidos, incluyendo la capacidad de realizar uniones complejas de tablas múltiples en datos distribuidos en *shards* sin requerir *co-sharding*¹, replicación de datos

ni transmisiones en la red. Esto hace posible *entregar respuestas rápidas a consultas que de otro modo podrían tardar horas o agotar el tiempo de respuesta sin completarse y al mismo tiempo se reducen los costos operacionales.*

InterSystems IRIS Data Platform permite a las organizaciones obtener una visión unificada a través de todos los datos de su empresa, además de obtener inteligencia más precisa y oportuna para sus negocios, asegurar un mejor cumplimiento con las regulaciones y responder con mayor rapidez a las preguntas de los reguladores financieros y analistas de cumplimiento, además de entregar las funcionalidades necesarias de acceso basado en roles, seguridad, encriptación, gobernanza y de procedencia. Es tecnología horizontalmente escalable que soporta el escalamiento horizontal, arquitecturas de *shards* para gestionar y analizar conjuntos de datos muy grandes usando nodos de almacenamiento y procesamiento distribuidos de bajo costo.



InterSystems realiza tareas de procesamiento complejas de forma rápida y eficiente sobre conjuntos de datos distribuidos sin requerir *co-sharding* o duplicación de datos para realizar uniones de tablas múltiples, por ejemplo, sin la necesidad de transmisiones por red. El resultado es un rendimiento y confiabilidad considerablemente mucho más altos, con un menor costo operacional, entregando respuestas rápidas a tareas que de otro modo podrían tardar horas o agotar el tiempo de respuesta sin completarse.

Figura 1: InterSystems IRIS Data Platform para riesgos consolidados y entre cumplimiento: Arquitectura de referencia

¹ Los datos *co-sharding* se refieren a datos particionados en una clave común.

InterSystems IRIS: Una base de datos HTAP multimodelo

El núcleo de InterSystems IRIS Data Platform es una base de datos multimodelo, de alcance empresarial y de procesamiento analítico transaccional híbrido (HTAP, por sus siglas en inglés), diseñada para trabajar con grandes cantidades de datos heterogéneos. Ingiera, almacena e indexa grandes volúmenes de datos transaccionales a velocidades de ingesta muy altas para responder a casos de uso analíticos en tiempo real.

Proporciona la flexibilidad de almacenar datos de fuentes dispares en el formato más apropiado. Los datos se almacenan una vez y se pueden describir en múltiples representaciones como SQL, objetos, conjuntos multidimensionales, pares valor/clave, documentos, etc. Esto elimina la necesidad de duplicar datos o proporcionar mapeos entre diferentes representaciones (por ejemplo, un mapeo de objeto a relacional) para un mayor rendimiento y eficiencia. Soporta de manera natural arquitecturas de *sharding*, escaladas horizontalmente y distribuidas, proporcionando una plataforma de datos rentable para trabajar con grandes conjuntos de datos usando recursos básicos. Entrega fuertes medidas de seguridad de nivel empresarial, integración con medidas de seguridad Kerberos y LDAP, control de acceso basado en roles y encriptación para datos en tránsito y en reposo.

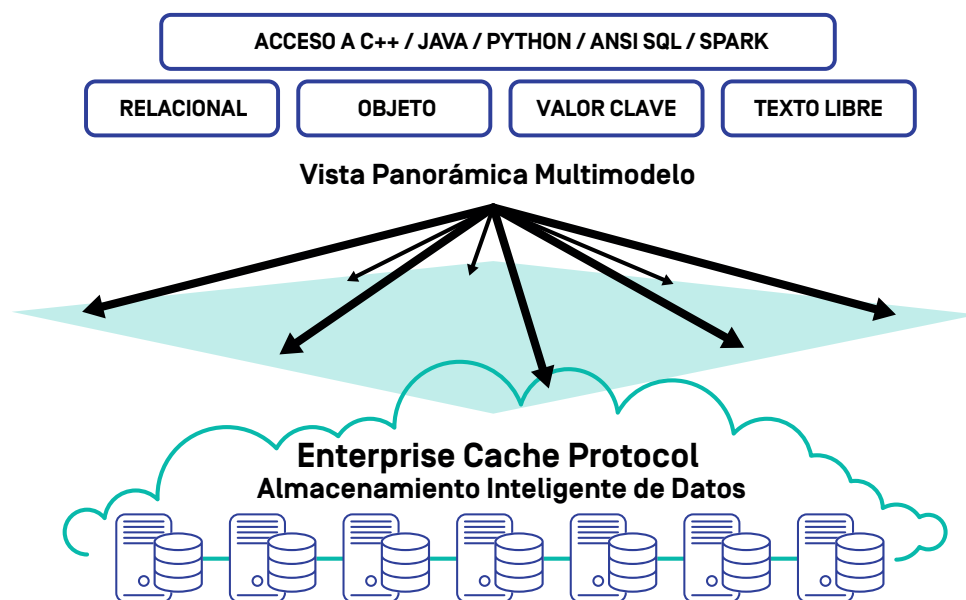


Figura 2: Base de datos transaccional-analítica híbrida multimodelo InterSystems IRIS

Enterprise Cache Protocol (ECP)

InterSystems proporciona potentes funcionalidades para un procesamiento de carga de trabajo múltiple confiable, distribuida y de alto rendimiento a muy alta escala. Esto se debe en gran parte a una tecnología única, Enterprise Cache Protocol (ECP), implementado, optimizado y reforzado en miles de ambientes de producción de misión-crítica.

ECP es una función integral de InterSystems IRIS. Entrega respuestas rápidas y confiables a consultas de conjuntos de datos distribuidos sin importar como estén organizados. ECP soporta de manera nativa las arquitecturas distribuidas y fragmentadas. Las uniones complejas se procesan de forma local, en lugar de difundirse a través de la red, eliminando las latencias e intervalos de espera típicamente asociados a las uniones de difusión, lo que aumenta el rendimiento y reduce los costos operacionales. ECP permite a los reguladores y analistas de cumplimiento obtener respuestas rápidas y certeras a consultas no planeadas sin pre-procesamientos ni replicaciones de datos que son costosas y consumen tiempo.

El uso de ECP es transparente y no requiere cambios en la aplicación ni técnicas especializadas. Las aplicaciones y las tareas de procesamiento simplemente tratan todos los datos como si fuesen locales. Los beneficios de rendimiento y escalabilidad son enormes, permitiendo a las organizaciones obtener respuestas, correlacionar información e identificar patrones en conjuntos de datos distribuidos, sin *co-sharding*, aumentando el rendimiento y reduciendo los costos operativos.

Integración con Apache Spark

Muchos de los controladores y reguladores para la gestión y cumplimiento de riesgo ahora necesitan reportes y visibilidad intradía en tiempo real, llevando a las organizaciones a utilizar paradigmas informáticos de mayor rendimiento para reducir latencias. Como resultado, algunas organizaciones están evaluando y adoptando Apache Spark, con su arquitectura en memoria.

InterSystems IRIS incorpora conectividad paralela rápida y eficiente con Apache Spark, proporcionando un complemento fluido y de alto rendimiento para las arquitecturas de big data ya en uso en muchas instituciones financieras. La integración inteligente de los *shards* y la capa de enriquecimiento soportan tanto a las *batch inserts* (inserciones por lote) como a las *Individual inserts* (inserciones individuales), permitiendo que los *shards* de datos de InterSystems ingieran rápidamente los grandes conjuntos de datos.

La tecnología de InterSystems se integra directamente con Spark a través de un conector Spark nativo compatible con los *shards*. El conector presenta los *shards* de datos de InterSystems como particiones nativas para un rendimiento más alto. Esta profunda integración permite a las organizaciones aprovechar a la perfección la tecnología de InterSystems y optimizar las consultas de Spark para un mejor rendimiento.

Procesamiento transaccional analítico simultáneo usando datos en tiempo real

Mientras que algunos casos de uso de riesgo y cumplimiento requieren análisis de datos en lote, otros requieren que se ingieran y analicen datos en tiempo real con datos históricos o de referencia. Por ejemplo, pueden requerir monitoreo de transacciones y filtrado, cálculos de liquidez intradía o pre-operación, y otras aplicaciones en tiempo real de baja latencia.

InterSystems IRIS está optimizado para procesar simultáneamente cargas de trabajo transaccionales altas y un gran volumen de consultas analíticas sobre datos transaccionales, históricos o en *batch* (por lotes). Esto se hace sin sacrificar el rendimiento de ningún tipo de carga de trabajo. Es ideal para manejar los requerimientos en tiempo real y en *batch* (por lotes).

Interoperabilidad

InterSystems IRIS provee conectividad para una amplia gama de aplicaciones y fuentes de datos, incluyendo bases de datos, archivos planos, etc. También incluye una biblioteca integrada de adaptadores que entregan conectividad y transformación de datos para estándares industriales, protocolos y tecnologías tradicionales como REST, SOAP, HTTP/S, Kafka y JMS. InterSystems IRIS ofrece una puerta de enlace SQL que puede acceder y presentar datos y metadatos en bases de datos comunes y *data warehouse*, incluyendo Oracle, Sybase y DB2, además de motores SQL en motores Hadoop como Hive e Impala, para aplicaciones de clientes de InterSystems como tablas nativas. Además, la plataforma de datos entrega capacidades para aplicar transformaciones, un motor de flujos de trabajo, encriptación, resolución de entidades y un motor de procesamiento de lenguaje natural para trabajar con textos no estructurados.

InterSystems IRIS proporciona un acceso integral y unificado para operar con las herramientas de generación de informes, análisis y visualización de terceros que ya estén en uso y aplicaciones existentes para la gestión de riesgos y cumplimiento normativo. Entrega soporte ANSI SQL con optimizaciones de SQL probadas y demostradas en arquitecturas de datos completamente fragmentadas en *shards* y escaladas horizontalmente, además de la flexibilidad para soportar acceso de objeto, OLAP, y JSON/REST a los datos.

Análisis integrado de texto no estructurado

Los datos no estructurados, incluyendo texto libre en correos, documentos, mensajes de texto, Reportes de Actividad Sospechosa (SARs), datos externos de blogs y tweets, pueden entregar información valiosa para ayudar a los bancos a reducir los riesgos e identificar comportamientos fraudulentos o sospechosos.

El procesamiento de lenguaje natural es un componente integral de InterSystems IRIS. Estas funcionalidades nativas entregan un enfoque ascendente único que analiza los textos en base a lo que contiene el texto en sí. Puede funcionar con diccionarios y ontologías definidas por el cliente y entrega capacidades de análisis semántico natural integrado para analizar patrones y correlaciones en datos no estructurados. InterSystems IRIS también incluye funcionalidades para realizar exploración de datos, detección de señales y análisis de tendencias, perfiles y agrupaciones basados en contenido, extracción de información, categorización y mapeo de datos no estructurados. Estas pueden ser útiles al resumir y contextualizar grandes cantidades de texto libre para varias iniciativas de cumplimiento y vigilancia.



Procedencia y gobernabilidad de datos

Establecer la procedencia efectiva de los datos es la capacidad de describir la fuente de los datos y cómo cambian a medida que se mueven a través del canal de información. La gobernabilidad de datos es fundamental para las iniciativas de riesgo y cumplimiento. Por ejemplo, las regulaciones como la Regla 613 de SEC (Registro de Auditoría Consolidado) requieren que las organizaciones recopilen e identifiquen con precisión cada orden, cancelación, modificación y ejecución de operaciones para todas las acciones y opciones cotizadas en bolsa en todos los mercados estadounidenses.

Las diferentes aplicaciones que realizan distintas funciones pueden almacenar varias representaciones de los datos o modificar los datos; por ejemplo, separar una orden inicial grande en órdenes derivadas más pequeñas para la ejecución. Los analistas y los reguladores de cumplimiento deben confiar en las fuentes originales de datos, en los procesos y las transformaciones que se aplican. InterSystems IRIS entrega un soporte robusto para múltiples tipos de datos, incluyendo esquemas de objeto y SQL y sus funcionalidades flexibles para metadatos permiten la aplicación de trazabilidad y procedencia de datos adecuadas.

Conclusión

InterSystems IRIS Data Platform es un complemento potente y fluido para la infraestructura existente de las instituciones financieras para entregar una vista panorámica y segura de todos sus activos de datos de toda la empresa. Ofrece una capa de procesamiento de escalamiento horizontal de ultra alto rendimiento distribuido para manejar una serie de tareas complejas en *batch* (por lote) y en tiempo real, que son necesarias para consolidar la gestión de riesgos y las soluciones puntuales de cumplimiento normativo en una plataforma unificada e integral.

Al incorporar InterSystems IRIS en sus iniciativas de consolidación, las instituciones financieras pueden hacer consultas más rápidas y rentables sobre los datos de su empresa en aplicaciones, *data lakes*, *warehouses*, obteniendo información precisa y oportuna sobre su negocio, asegurando un mejor cumplimiento de las regulaciones de la industria, respondiendo con mayor rapidez a auditorías de reguladores financieros y analistas de cumplimiento, reduciendo los riesgos operacionales y normativos.

InterSystems es el motor detrás de las aplicaciones más importantes del mundo. En la salud, finanzas, gobierno y otros sectores donde la vida y la subsistencia de las personas estén en juego, InterSystems es el poder detrás de lo que es importante. Fundada en 1978, InterSystems es una compañía privada, cuyas oficinas centrales se encuentran en Cambridge, Massachusetts, EE. UU. y con oficinas en todo el mundo. Millones de personas en más de 80 países utilizan sus productos de software diariamente.

Para más información, visite:
[Financial.InterSystems.com](https://financial.intersystems.com)





[InterSystems.com/cl](https://www.intersystems.com/cl)