

# Cómo lograr una observabilidad completa de los datos sin gastar una fortuna

Descubra un enfoque nuevo y desacoplado para monitorizar las canalizaciones de datos y garantizar la calidad a escala, sin llevar al límite los sistemas ni el presupuesto.

---

Autor: Ole Olesen-Bagneux  
Evangelista Jefe

## Índice

- 3 ¿Por qué las organizaciones necesitan la observabilidad de los datos?
  - 4 ¿Qué es la observabilidad de los datos?
  - 5 ¿Cómo se lleva a cabo la observabilidad de los datos con Actian?
  - 7 Resumen ejecutivo: el punto de vista de Actian sobre la observabilidad de los datos
  - 8 Actian Data Observability en acción
  - 8 Acerca de Actian
- 

En la actualidad, la magnitud y velocidad a las que se mueven los datos están desbordando los controles de calidad tradicionales. Actian presenta un enfoque moderno e inteligente para la observabilidad de los datos, que aporta seguridad, confianza y rentabilidad a las operaciones de datos de las empresas.

## Acerca del autor

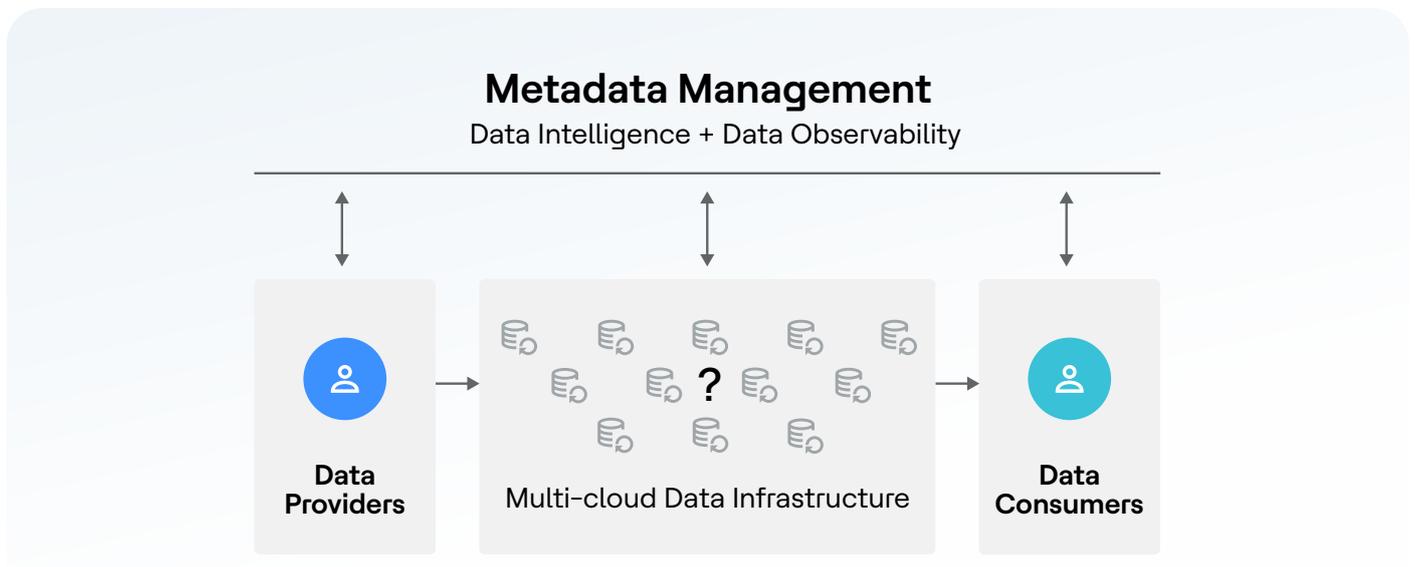


**Ole Olesen-Bagneux** es un referente mundial en gestión de metadatos y arquitectura de datos empresariales. Como Vicepresidente y Evangelista Jefe de Actian, promueve la concienciación del sector y la adopción de enfoques modernos en inteligencia de datos, aprovechando su amplia experiencia en gestión de datos, metadatos, catálogos de datos y arquitecturas descentralizadas. Como autor consumado, ha escrito «The Enterprise Data Catalog» (O'Reilly, 2023) y actualmente está trabajando en «Fundamentals of Metadata Management» (O'Reilly, 2025). Tiene un doctorado en Biblioteconomía y Ciencias de la Información por la Universidad de Copenhague y aporta una perspectiva única que une la informática tradicional con la gestión moderna de los datos.

## ¿Por qué las organizaciones necesitan la observabilidad de los datos?

La respuesta es sencilla: las empresas tienen inmensas cantidades de datos almacenados y transformados en muchas tecnologías diferentes. Para que estos datos sean útiles para los objetivos previstos, las empresas deben observarlos con suma atención. Por ello ha surgido la disciplina y la tecnología de la observabilidad de los datos. [1]

Certains collaborateurs sont donc chargés de fournir ces données, qu'ils doivent transporter et transformer afin En consecuencia, algunos empleados tienen la tarea de facilitar los datos necesarios, transportarlos y transformarlos para que sean accesibles para otros empleados: los usuarios de los datos. Para tender un puente entre los proveedores y los usuarios de los datos, las organizaciones crean a lo largo del tiempo una gran infraestructura de datos multinube, y ahí está el problema: esta infraestructura de datos es difícil de controlar, ya que abarca multitud de tecnologías de diferentes proveedores, con enormes cantidades de datos que se transportan de muchas maneras y con muchas soluciones. Esto es complicado de controlar. Para resolver el problema, las empresas recurren a la gestión de metadatos para gobernar su infraestructura de datos. [2]



Piense en los propietarios de los sistemas de su organización y en los datos que estos contienen: pueden pertenecer a los departamentos de recursos humanos, fabricación, I+D, finanzas, etc. Los datos en esos sistemas poseen un enorme valor: son la fuente para casos de uso innovadores para generar ofertas mejoradas e incluso nuevas, lo que garantiza el crecimiento y el futuro de su empresa.

La capa de metadatos que se encuentra sobre la infraestructura de datos multinube da a las empresas la inteligencia necesaria para gobernar sus datos, mediante un conjunto de capacidades relacionadas con los metadatos, entre ellas la observabilidad.

En resumen, las organizaciones necesitan la observabilidad de los datos porque es imposible controlarlos todos manualmente, dada la gran cantidad de ellos que circulan por multitud de soluciones de almacenamiento. La observabilidad de los datos garantiza que los usuarios reciban datos recientes, correctos y de alta calidad.

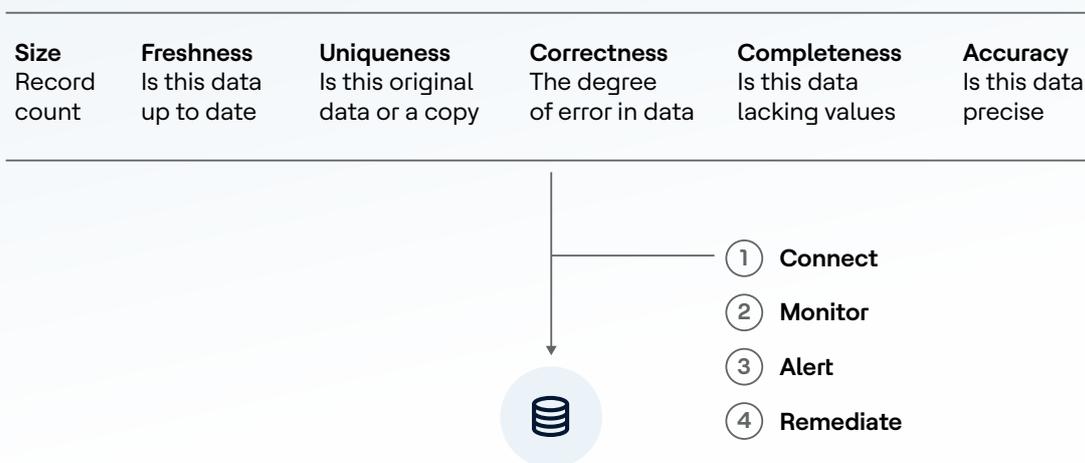
## ¿Qué es la observabilidad de los datos?

¿Pero qué es exactamente la observabilidad de los datos? Consiste en observar el tamaño, actualidad, singularidad, corrección, integridad y precisión de los datos. Para configurar, gestionar y aplicar estas métricas [3], hay que conectar los datos, monitorizarlos, alertar sobre ellos y, en última instancia, corregirlos.

**Transporte de los datos por las canalizaciones.** Transportar los datos por canalizaciones llamadas herramientas ETL (extraer-transformar-cargar) o ELT (extraer-cargar-transformar) es la actividad central de la ingeniería de datos. [4] Para ayudar a los ingenieros de datos, la observabilidad de los datos puede evaluar el rendimiento de las canalizaciones en cuanto a alcance y estabilidad.



## Data Observability



La observabilidad de los datos posibilita:

### Canalizaciones lógicas de los datos (transporte de datos):

- Rendimiento de la canalización de datos
- Coste de la canalización de los datos
- Predicciones sobre la canalización de los datos

### Escalar las arquitecturas de datos:

- Productos de datos fiables
- Implementación más sencilla de los contratos de datos
- Mayor éxito en los casos de uso de la IA

### Gobernanza y cumplimiento

- Aumentar la calidad de los datos
- Controlar el uso de información de identificación personal y de datos confidenciales
- Mejorar el catálogo de datos (el «buscador») en la capa de metadatos

Veamos estos puntos con más detalle.

La observabilidad permite reducir el coste de la canalización, ya que hay menos pasos y son más sencillos, e incluso avisa si la canalización planeada va a fallar o a ser frágil.

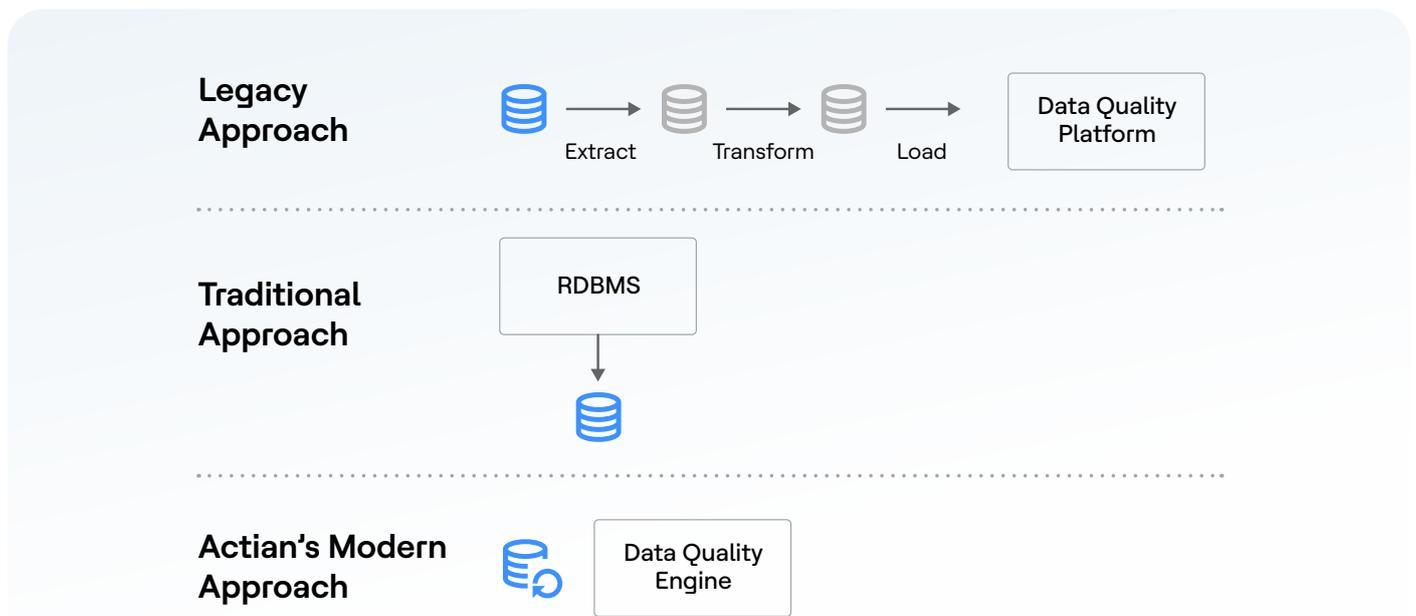
**Escalar las arquitecturas de datos.** Las organizaciones se están pasando a los productos de datos, ya que los datos gestionados como producto son más escalables que los que se gestionan en grandes almacenes empresariales y lagos de datos. La observabilidad es un componente fundamental para escalar las arquitecturas con canalizaciones sólidas para unos productos de datos de alta calidad y fáciles de consumir gracias a los contratos de datos automatizados, para que al final su empresa triunfe con la IA. [5]

**Gobernanza y cumplimiento.** La observabilidad juega un papel clave en la gobernanza de los datos, ya que garantiza su calidad mediante políticas que pueden aplicarse con supervisión, incluido el uso de información de identificación personal y datos confidenciales. La observabilidad puede incluso integrarse en el «buscador» de la capa de metadatos (el catálogo de datos), de modo que los usuarios finales solo descubran datos fiables. [6]

## ¿Cómo se lleva a cabo la observabilidad de los datos con Actian?

En Actian, defendemos un enfoque completamente nuevo y moderno para la observabilidad de los datos. Generalmente, la observabilidad de los datos se hace con un enfoque tradicional y algunas organizaciones todavía la realizan con sistemas heredados. El diagrama contiguo muestra los diferentes enfoques.

**El enfoque tradicional.** El enfoque tradicional de observabilidad usa un sistema de administración de bases de datos relacionales (RDBMS, por sus siglas en inglés) encima de la base de datos para consultar y observar los datos, igual que cualquier otro tipo de consulta realizada con el RDBMS. Este enfoque es mejor que el heredado, pero sigue presentando inconvenientes considerables. Hay un límite computacional en el número de consultas que se pueden ejecutar y, para que sean rentables, estas no pueden abarcar todos los datos.



Para entender por qué destaca la solución de Actian, vamos a explicar los enfoques heredados, tradicionales y modernos para la observabilidad de los datos.

**El enfoque heredado.** El enfoque heredado en observabilidad de datos consiste en configurar una canalización clásica de datos que los extrae, transforma y carga en una plataforma de calidad de datos, donde se realizan diversas verificaciones automáticas y manuales. Este sistema es muy laborioso, no está sincronizado con los datos que observa (puede haber cambios que dejen obsoletas las conclusiones de la observación) y no es escalable, porque exige demasiado trabajo.

Por lo tanto, los datos se observan de forma incompleta, ya que el sistema solo puede observar una muestra, por lo que la observabilidad es intrínsecamente poco fiable. No se puede escalar hasta un nivel que permita observar todos los datos en cada fuente. Además, hay muchas limitaciones en los tipos de fuentes de datos que se pueden observar, ya que se excluyen los semiestructurados, como los scripts JSON anidados y archivos YAML. Por último, este enfoque requiere muchas consultas, por lo que resulta caro. ¿De verdad quiere una observabilidad de datos poco fiable, incompleta y costosa?

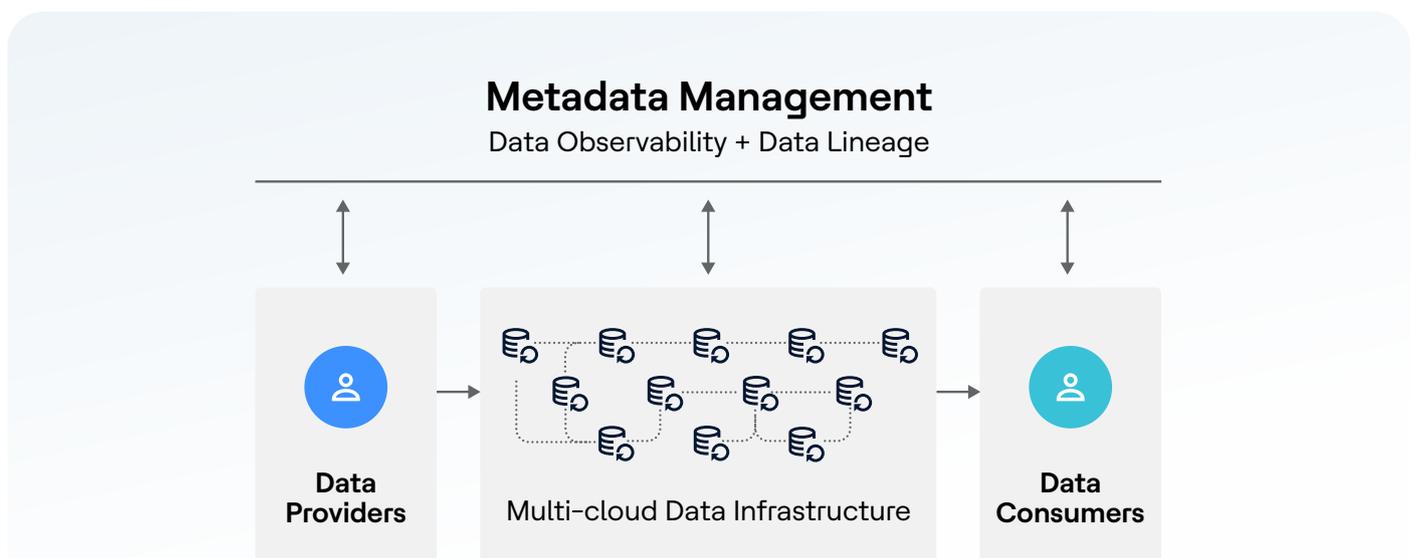
## El enfoque moderno de Actian

Para observar los datos, Actian Data Observability configura un pequeño motor de calidad de datos desacoplado junto a los datos. Este motor supervisa de continuo los datos en la memoria, estando junto a los datos y no a través del sistema de administración de bases de datos relacionales encima. [7] De esta manera, los datos no se transportan para ser observados, como en el enfoque heredado. Ni tampoco se consulta la base de datos con SQL para observarlos, como en el enfoque tradicional. Al utilizar un enfoque moderno con un motor de calidad de datos desacoplado, Actian Data Observability puede conectarse de forma nativa a todas sus fuentes de datos estructuradas y semiestructuradas, de archivos JSON, Parquet y YAML a la mayoría de los formatos de tabla abiertos y flujos, que generalmente no se pueden observar con el enfoque tradicional. Una de nuestras mayores ventajas es cómo Actian Data Observability hace que la calidad de los datos sea completamente independiente de los sistemas subyacentes. Este enfoque independiente del sistema simplifica mucho la adopción por clientes que usan diversas pilas tecnológicas. Además, esta arquitectura desacoplada garantiza el futuro de los procesos de calidad, al eliminar los problemas de migración si los sistemas cambian más adelante. Dado que el motor de calidad de datos funciona al lado de estos, se reduce la presión sobre los sistemas operativos, ya que no es necesario realizar consultas en sus bases de datos de forma continua.

Esto también reduce mucho el coste de la observabilidad de los datos, ya que la memoria se monitoriza de forma condensada y precisa. Por último, este enfoque permite una observabilidad total de todos los datos en las fuentes, ya que no debe tener en cuenta el coste de consultarlos. Esto le brinda una observabilidad completa de los datos, fácil de configurar y ejecutar, rentable y, ante todo, exhaustiva. Un último componente clave hace que Actian Data Observability sea tan potente.

## La solución completa: el linaje de datos para la observabilidad

Por sí sola, una solución de observabilidad de los datos no puede rastrear el movimiento de los datos entre las distintas fuentes. Solo observa directamente una o varias fuentes sin compararlas. Por lo tanto, no puede rastrear cómo cambia la calidad de los datos a medida que estos se mueven por la infraestructura, a menos que se mapeen a mano las fuentes. Por eso, Actian Data Observability incorpora un linaje de datos que los rastrea automáticamente a través de las diferentes fuentes. Esta combinación da una observabilidad perfecta, del proveedor al usuario de los datos.



## Resumen ejecutivo: el punto de vista de Actian sobre la observabilidad de los datos

En Actian, creemos que la observabilidad de los datos puede hacerse de acuerdo con los valores tecnológicos que defendemos: debe ser fiable, flexible y fácil de usar. Eso es lo que ofrecemos con Actian Data Observability.

Como jefe de datos o responsable de tecnología de datos debe tenerlo claro: su organización merece una solución de observabilidad de datos moderna, rentable y completa. Justo lo que ofrece Actian Data Observability, que destaca en un mercado dominado por tecnologías que, por desgracia, son caras, complicadas y sencillamente incompletas.

No puede confiar en las soluciones tradicionales de observabilidad de datos para proyectos estratégicos de datos de los que depende el futuro de su empresa. Las tecnologías tradicionales de observabilidad de datos no son escalables y probablemente ni siquiera logren una adopción significativa dentro de su organización. Sin embargo, estresarán los sistemas de origen y serán costosas.

Con Actian Data Observability y las capacidades de linaje de datos de Actian, su empresa puede generar confianza en los datos en todo su entorno informático, al tiempo que mantiene los sistemas de producción funcionando sin problemas y con costes mínimos.



### Referencias

[1] Hablaremos de la observabilidad de los datos basándonos en la definición, con la que estamos de acuerdo, que ofrece Gartner® en «Quick Answer: "What is Data Observability"», Ankush Jain, 8 de junio de 2022: «La observabilidad de los datos es la capacidad de una organización para tener una amplia visibilidad de su panorama de datos y las dependencias de datos multicapa (como canalizaciones, infraestructura y aplicaciones de datos) en todo momento, para identificar, controlar, prevenir, escalar y corregir las interrupciones en los datos rápido y según acuerdos de nivel servicio de datos aceptables».

[2] El diagrama se basa en la figura 1-1. de Pietheine Strengtholt: «Building Medallion Architectures» (O'Reilly, 2025). La observabilidad de los datos también es una capacidad necesaria en las arquitecturas Medallion.

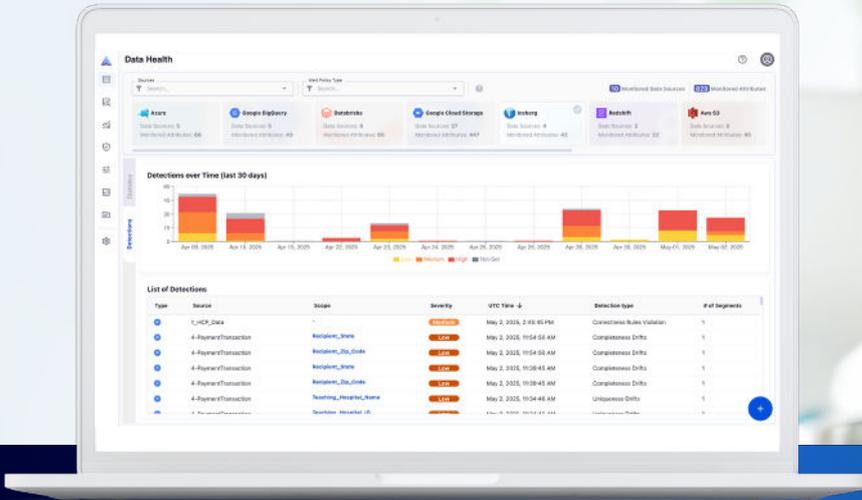
[3] Aquí no profundizaremos en los detalles matemáticos de las métricas. Una fuente valiosa sobre el tema es Andy Petrella: «Fundamentals of Data Observability» (O'Reilly, 2023), capítulo 4.

[4] El trabajo clave sobre ingeniería de datos es Joe Reis y Matthew Housley: «Fundamentals of Data Engineering» (O'Reilly, 2022).

[5] Los estudios muestran que más del 60 % de los proyectos de IA fracasarán antes de 2026 porque los datos no están preparados para la IA. Véase Gartner® «A Journey Guide to Delivering AI Success through "AI-Ready" Data» de Ehtisham Zaidi, Roxane Edilali, 18 de octubre de 2024.

[6] Ole Olesen-Bagneux: «The Enterprise Data Catalog», (O'Reilly, 2023), capítulo 7.

[7] Los detalles técnicos sobre esta configuración son importantes, pero demasiado complejos para este libro blanco. Encontrará más información en [actian.com/data-observability](https://actian.com/data-observability).



## ¿Quiere ver cómo funciona Actian Data Observability?

Empiece a reducir costes, aumentar la confianza y simplificar la gobernanza en toda su arquitectura de datos... más rápido de lo que pensaba.

Con el enfoque moderno y desacoplado de Actian, tendrá una plena visibilidad de la calidad de sus datos, sin superar el límite de los sistemas ni del presupuesto. Da igual si escala productos de datos, se prepara para la IA o aplica la gobernanza a gran velocidad, Actian le ayuda a tomar el control con confianza.

Visite [actian.com](https://actian.com) y solicite una demostración para ver cómo puede transformar Actian la forma de observar los datos, confiar en ellos y gestionarlos.

## Acerca de Actian

Actian permite a las empresas gestionar y controlar los datos a gran escala con total confianza. Las organizaciones confían en las soluciones de gestión de datos e inteligencia de datos de Actian para agilizar los entornos de datos complejos y acelerar la entrega de datos listos para la IA. Diseñadas para ser flexibles, las soluciones de Actian se integran a la perfección y funcionan de forma fiable en entornos locales, en la nube e híbridos. Obtenga más información sobre Actian, la división de datos de HCLSoftware, en [actian.com](https://actian.com).

