

# Wie Sie vollständige Data Observability erreichen, ohne Ihr Budget zu sprengen

Entdecken Sie einen neuen, entkoppelten Ansatz zur Überwachung von Datenpipelines. Stellen Sie die höchstmögliche Qualität sicher, ohne Ihre Systeme oder Ihr Budget zu belasten.

---

Von Ole Olesen-Bagneux  
Chief Evangelist, PhD

## Inhaltsverzeichnis

- 3 Warum benötigen Unternehmen Data Observability?
  - 4 Was ist Data Observability?
  - 5 Wie wird Data Observability mit Actian umgesetzt?
  - 7 Executive Summary: Actians Standpunkt zu Data Observability
  - 8 Actian Data Observability in Aktion
  - 8 Über Actian
- 

Das Ausmaß und die Geschwindigkeit heutiger Datenbewegungen überfordern herkömmliche Maßnahmen zur Qualitätskontrolle. Actian präsentiert einen modernen, intelligenten Ansatz für Data Observability. Dieser verhilft Unternehmen zu mehr Vertrauen, Zuversicht und Kosteneffizienz im Datenbetrieb.

## Über den Autor

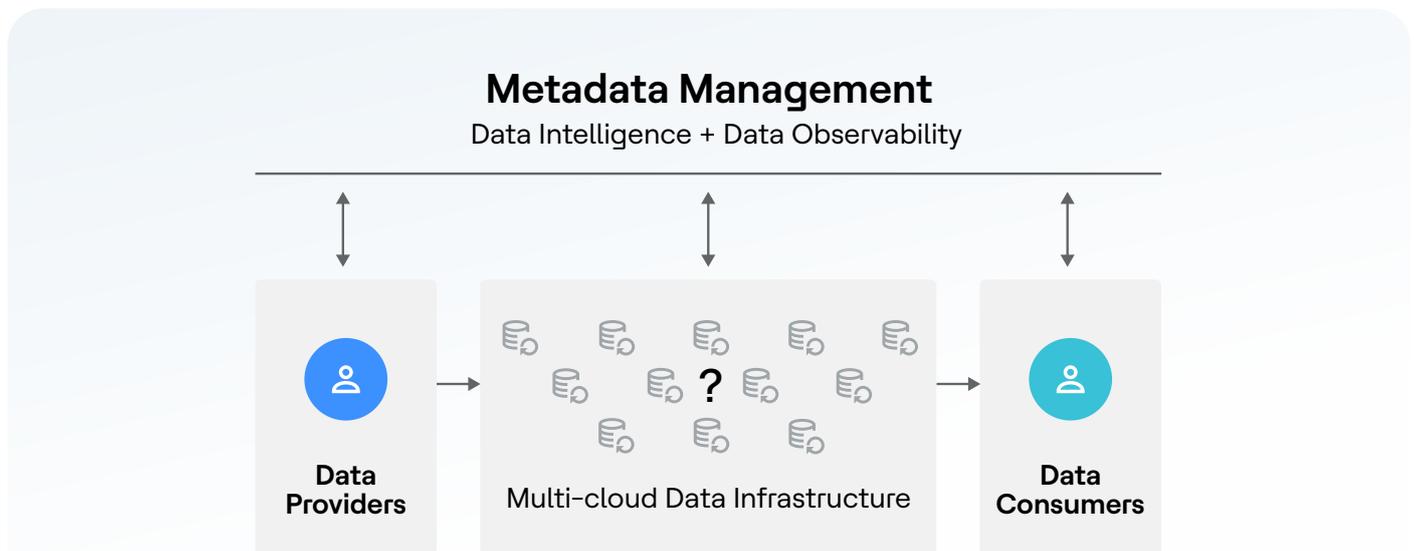


**Ole Olesen-Bagneux** ist ein weltweit anerkannter Vordenker im Bereich Metadatenmanagement und Unternehmensdatenarchitektur. In seiner Funktion als VP und Chief Evangelist bei Actian fördert er das Bewusstsein für moderne Ansätze der Datenintelligenz und regt deren Einführung an. Dabei stützt er sich auf sein umfangreiches Fachwissen in den Bereichen Datenmanagement, Metadaten, Datenkataloge und dezentrale Architekturen. Ole ist ein versierter Autor und Verfasser des 2023 bei O'Reilly erschienenen Buchs „The Enterprise Data Catalog“. Derzeit arbeitet er an Fundamentals of Metadata Management (O'Reilly, 2025). Er hat einen Dokortitel in Bibliotheks- und Informationswissenschaft der Universität Kopenhagen und verbindet durch sein einzigartige Perspektive die traditionelle Informationswissenschaft mit dem modernen Datenmanagement.

## Warum benötigen Unternehmen Data Observability?

Die Antwort ist ganz einfach: Unternehmen speichern und verarbeiten mithilfe vieler verschiedener Technologien immense Datenmengen. Um sicherzustellen, dass diese Daten für die vorgesehenen Zwecke auch tatsächlich nützlich sind, müssen Unternehmen sie aufmerksam beobachten. Genau hierfür wurde die Disziplin und Technologie der Data Observability entwickelt. [1]

Dementsprechend werden einige Mitarbeiter mit der Bereitstellung von Daten beauftragt. Sie müssen diese transportieren und transformieren, um sie für andere Mitarbeiter und Datenkonsumenten nutzbar zu machen. Unternehmen bauen so im Laufe der Zeit eine große Multi-Cloud-Dateninfrastruktur auf, um die Kommunikation zwischen Datenanbietern und -verbrauchern zu erleichtern. Genau hier liegt jedoch das Problem: Diese Infrastruktur ist schwer zu kontrollieren, da sie aus einer Fülle heterogener Technologien besteht und enorme Datenmengen enthält, die auf unterschiedliche Weise über zahlreiche Lösungen hinweg transportiert werden. Das erschwert die Verwaltung. Um dieses Problem zu lösen, verlassen sich Unternehmen auf das Metadatenmanagement. [2].



Denken Sie einmal an die Systemverantwortlichen sowie an die Daten, die sich auf den Systemen Ihres Unternehmens befinden: Das können Daten aus Bereichen wie Personalwesen, Produktion, F&E und Finanzen sein. Diese sind in der Regel ausgesprochen wertvoll, da sie als Quelle für innovative Anwendungsfälle dienen. Basierend darauf, lassen sich verbesserte oder sogar neue Angebote bereitstellen. Dadurch sichern das Wachstum und die Zukunft Ihres Unternehmens.

Die Metadaten-schicht der Multi-Cloud-Dateninfrastruktur bietet Unternehmen die Möglichkeit, ihre Daten mithilfe einer Reihe von Funktionen zu verwalten, darunter Data Observability.

**Kurz gesagt: Unternehmen benötigen Data Observability, da heutzutage so viele Daten über so viele verschiedene Speicherlösungen transportiert werden, dass eine manuelle Kontrolle nicht mehr möglich ist. Data Observability stellt sicher, dass Datenverbraucher stets frische, korrekte und qualitativ hochwertige Daten erhalten.**

## Was ist Data Observability?

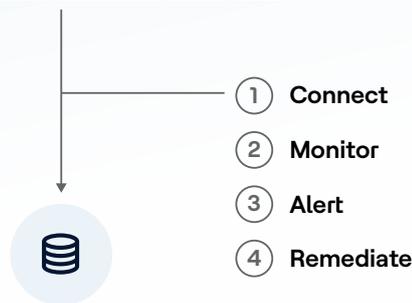
Werfen wir einen genaueren Blick darauf, was Data Observability eigentlich ist. **Data Observability beschreibt den Prozess der Untersuchung von Daten hinsichtlich ihrer Größe, Aktualität, Einzigartigkeit, Korrektheit, Vollständigkeit und Genauigkeit. Mithilfe einer Verbindung zu den Daten lassen sich diese Metriken [3] einrichten, verwalten und durchsetzen. Dafür werden sie präventiv überwacht. Darüber hinaus lassen sich Warnmeldungen einrichten, was die Problembehebung optimiert.**

**Transport von Daten in Pipelines.** Die Kernaktivität des Data Engineering ist der Transport von Daten durch Pipelines. Dafür werden ETL- (Extract-Transform-Load) oder ELT-Tools (Extract-Load-Transform) verwendet. [4] Data Observability hilft Dateningenieuren, die Performance von Pipelines in Bezug auf Umfang und Stabilität zu bewerten. Sie ermöglicht es, mit weniger und einfacheren Schritten die Kosten von Pipelines zu senken. Sie kann sogar davor warnen, falls geplante Pipelines brüchig werden oder ausfallen.



## Data Observability

Size	Freshness	Uniqueness	Correctness	Completeness	Accuracy
Record count	Is this data up to date	Is this original data or a copy	The degree of error in data	Is this data lacking values	Is this data precise



Data Observability ermöglicht:

- ✓ **Logische Datenpipelines (Datentransport):**
  - Leistung der Datenpipelines
  - Kosten der Datenpipelines
  - Vorhersagen für Datenpipelines
- ✓ **Skalierbare Datenarchitekturen:**
  - Zuverlässige Datenprodukte
  - Einfachere Implementierung von Datenverträgen
  - Größerer Erfolg für KI-Anwendungsfälle
- ✓ **Governance und Compliance**
  - Erhöhte Datenqualität
  - Die Verwendung von personenbezogenen und sensiblen Daten lässt sich leichter identifizieren
  - Verbesserung des Datenkatalogs (der „Suchmaschine“) in der Metadatenschicht

**Skalierbare Datenarchitekturen.** Unternehmen bewegen sich verstärkt in Richtung Datenprodukte, da Daten sich auf diese Weise besser skalieren lassen als solche, die in großen Data Warehouses und Data Lakes gespeichert sind. Data Observability ist bei der Skalierung von Datenarchitekturen eine Schlüsselkomponente. Sie bietet solide Pipelines für hochwertige Datenprodukte, die sich durch automatisierte Datenverträge einfach nutzen lassen. Dies verhilft Ihrem Unternehmen letztendlich mit KI zum Erfolg. [5]

**Governance und Compliance.** Data Observability ist für die Data Governance von entscheidender Bedeutung. Sie stellt die Datenqualität durch Richtlinien sicher, die mithilfe von Überwachung durchgesetzt werden. Das gilt auch im Hinblick auf die Verwendung personenbezogener oder sensibler Daten. Darüber hinaus kann Observability in die „Suchmaschine“ der Metadatenschicht, d.h. den Datenkatalog, integriert werden, sodass Endbenutzer nur Zugang zu vertrauenswürdigen Daten haben. [6]

Schauen wir uns diese genauer an.

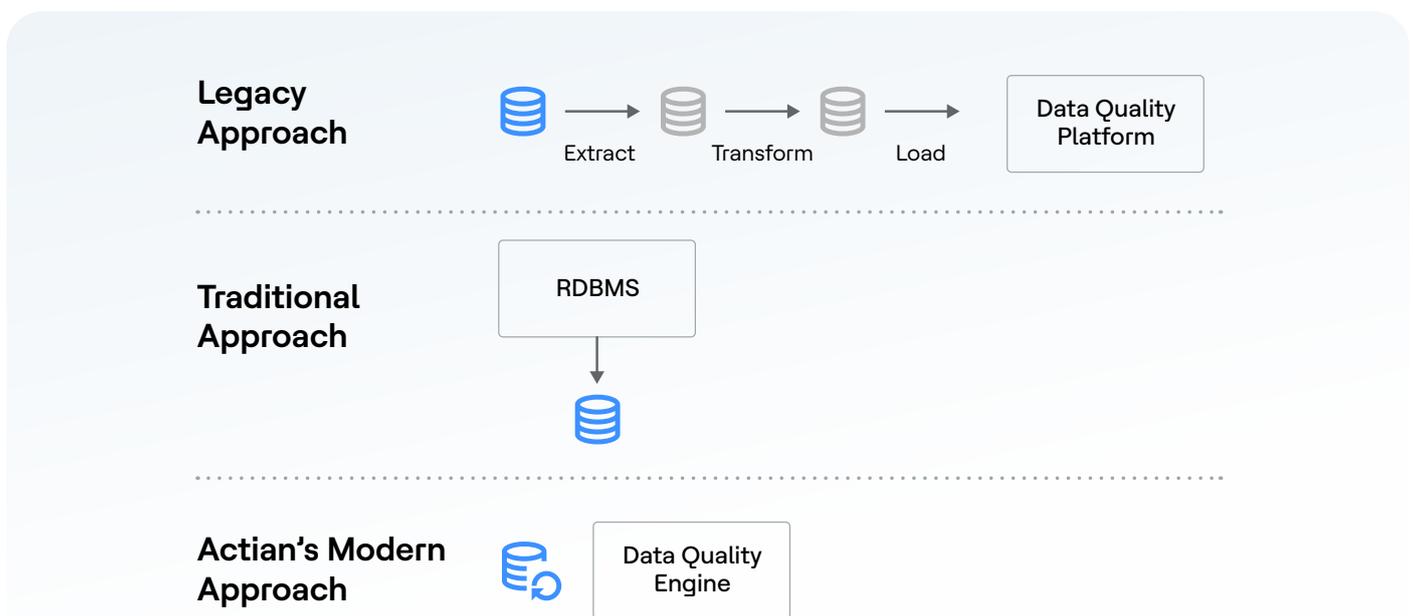
## Wie wird Data Observability mit Actian umgesetzt?

Bei Actian setzen wir uns für einen brandneuen, modernen Ansatz zur Data Observability ein. Data Observability basiert meist auf einem herkömmlichen Ansatz. Einige Unternehmen nutzen hierfür immer noch eine veraltete Konfiguration. Die verschiedenen Ansätze sind im folgenden Diagramm dargestellt.

Um die Vorzüge der Actian-Lösung hervorzuheben, vergleichen wir den herkömmlichen Legacy-Ansatz mit modernen Ansätzen zur Data Observability.

**Der herkömmliche Ansatz.** Bei herkömmlichen Ansätzen zur Data Observability werden Daten mithilfe eines relationalen Datenbankmanagementsystems (RDBMS) innerhalb einer Datenbank abgefragt und beobachtet. Dieses Verfahren hat gegenüber dem Legacy-Ansatz zwar Vorteile, weist aber auch erhebliche Nachteile auf. So ist die Anzahl der möglichen Abfragen rechnerisch begrenzt. Aus Gründen der Kosteneffizienz lassen sich Abfragen außerdem nicht auf allen Daten ausführen.

Da die Daten nur durch Stichproben beobachtet werden, sind die Ergebnisse unvollständig. Das macht Ihre Data Observability grundsätzlich unzuverlässig.



**Der Legacy-Ansatz.** Dieser Ansatz beinhaltet die Einrichtung einer klassischen Datenpipeline. Diese extrahiert und transformiert die Daten und lädt sie in eine Datenqualitätsplattform. Dort werden diese dann einer Kombination aus automatisierten und manuellen Prüfungen unterzogen. Diese Vorgehensweise ist sehr zeitaufwändig und wird nicht mit den beobachteten Daten synchronisiert. Dadurch können Änderungen auftreten, durch die die beobachteten Daten veralten. Aufgrund ihres hohen Arbeitsaufwands ist die Methode nicht skalierbar.

Das Ganze lässt sich nicht so weit skalieren, dass alle Daten jeder Quelle beobachtet werden können. Zudem gibt es erhebliche Einschränkungen bei den Arten von Datenquellen, die beobachtet werden können. So sind beispielsweise halbstrukturierte Daten wie verschachtelte JSON-Skripte und YAML-Dateien ausgeschlossen. Schließlich ist dieser Ansatz abfrageintensiv und deshalb teuer. Wollen Sie tatsächlich eine unzuverlässige, unvollständige und kostspielige Data Observability?

## Der moderne Ansatz von Actian

Actian Data Observability nutzt eine kleine, entkoppelte Datenqualitäts-Engine, die neben den Daten selbst eingerichtet wird, um diese zu beobachten. Diese Engine überwacht die Daten kontinuierlich im Speicher, statt über das oben genannte relationale Datenbankmanagementsystem. [7] Im Gegensatz zum herkömmlichen Ansatz werden die Daten nicht transportiert, um sie zu beobachten. Für diesen Prozess wird die Datenbank auch nicht, wie beim herkömmlichen Ansatz, mit SQL abgefragt. Dank dieses modernen Ansatzes kann Actian Data Observability eine native Verbindung zu all Ihren strukturierten und halbstrukturierten Datenquellen herstellen. Dazu gehören JSON-, Parquet- und YAML-Dateien sowie die meisten offenen Tabellenformate und Streams. Mit dem herkömmlichen Ansatz lassen diese sich in der Regel nicht beobachten. Eine der größten Stärken unseres Ansatzes ist, dass Actian Data Observability die Datenqualität völlig unabhängig von den darunter liegenden Systemen gewährleistet. Dieser systemunabhängige Ansatz vereinfacht den Einführungsprozess für Kunden mit unterschiedlichen technischen Systemen erheblich. Außerdem macht diese entkoppelte Architektur die Qualitätsprozesse zukunftssicherer, indem sie Migrationsprobleme bei einem Systemwechsel vermeidet. Da die Datenqualitäts-Engine in unmittelbarer Nähe zu den Daten arbeitet, entlastet sie außerdem die operativen Systeme, weil Abfragen nicht ständig an Datenbanken gerichtet werden müssen.

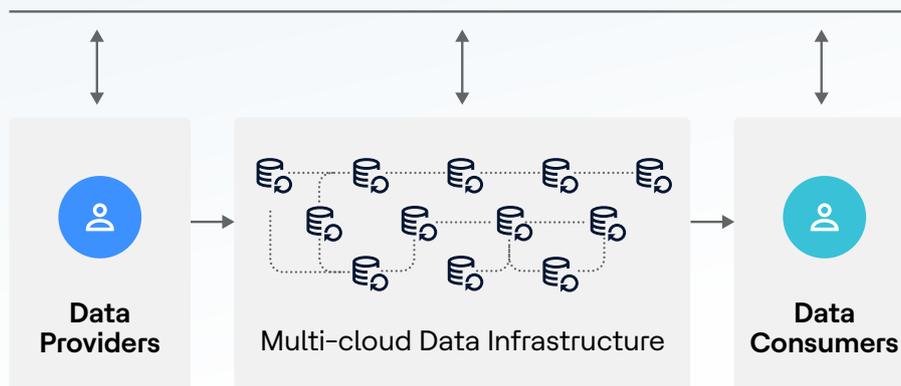
Dank der präzisen und komprimierten Überwachung des Speichers werden die Kosten für die Data Observability erheblich reduziert. Mit diesem Ansatz lässt sich daher eine umfassende Beobachtung sämtlicher Datenquellen realisieren, da die Kosten für die Datenabfrage keine Rolle spielen. Dies ermöglicht eine vollständige Data Observability, die sich einfach einrichten und ausführen lässt. Sie ist kostengünstig und vor allem vollständig. Eine letzte Schlüsselkomponente macht Actian Data Observability so leistungsstark:

## Die Komplettlösung – Data-Lineage für Data Observability

Eine Data-Observability-Lösung kann Datenbewegungen zwischen Datenquellen nicht eigenständig verfolgen. Sie prüft lediglich eine oder mehrere Quellen direkt, ohne diese zu vergleichen. Um die Änderungen der Datenqualität zu verfolgen, wenn sich die Daten durch die Dateninfrastruktur bewegen, müssen Sie die Quellen daher manuell zuordnen. Aus diesem Grund verfügt Actian Data Observability über eine Forward-Data-Lineage, mit der sich Daten automatisch über diverse Datenquellen hinweg verfolgen lassen. In Kombination mit Actian Data Observability lassen sich Daten vom Anbieter bis hin zum Verbraucher perfekt beobachten.

## Metadata Management

Data Observability + Data Lineage



## Zusammenfassung: Welchen Standpunkt Actian zur Data Observability vertritt

Bei Actian vertreten wir die Meinung, dass Data Observability auf den technologischen Werten basieren sollte, die uns am Herzen liegen: Sie muss vertrauenswürdig, flexibel und benutzerfreundlich sein. Genau das bieten wir mit Actian Data Observability.

Wenn Sie Datenverantwortlicher oder Entscheidungsträger für Datentechnologie sind, sollten Sie Folgendes wissen: Ihr Unternehmen verdient eine moderne, kosteneffektive und umfassende Data-Observability-Lösung. Actian Data Observability ist genau das, was Actian Data Observability und hebt sich damit von einem Markt ab, der von teuren, komplizierten und unvollständigen Technologien dominiert wird.

Herkömmliche Data-Observability-Lösungen können die strategischen Datenprojekte, von denen die Zukunft Ihres Unternehmens abhängt, nicht liefern. Diese Technologien belasten Ihre Quellsysteme, sind kostspielig und nicht skalierbar. Daher wird die Akzeptanz in Ihrem Unternehmen höchstwahrscheinlich nicht sehr hoch sein.

Durch die Kombination von Actian Data Observability mit den Data-Lineage-Funktionen von Actian kann Ihr Unternehmen in seiner gesamten IT-Landschaft Datenvertrauen aufbauen. So können Sie sicherstellen, dass die Produktionssysteme mit minimalem Aufwand reibungslos laufen.



### Literaturverzeichnis

[1] Wir sprechen über Data Observability und stützen uns dabei auf die folgende Definition von Gartner®: „Data Observability ist die Fähigkeit eines Unternehmens, jederzeit einen umfassenden Überblick über seine Datenlandschaft und die mehrschichtigen Datenabhängigkeiten (wie Datenpipelines, Dateninfrastruktur und Datenanwendungen) zu haben. Ziel ist es, Datenausfälle innerhalb akzeptabler Daten-SLAs schnell zu erkennen, zu kontrollieren, zu verhindern, zu eskalieren und zu beheben.“ (Ankush Jain, 8. Juni 2022, „Quick Answer: Was ist Data Observability?“)

[2] Das Diagramm ist inspiriert von Abbildung 1-1. in Pietheine Strengtholt: Building Medallion Architectures (O'Reilly, 2025) – Data Observability ist auch für Medallion-Architekturen eine notwendige Fähigkeit.

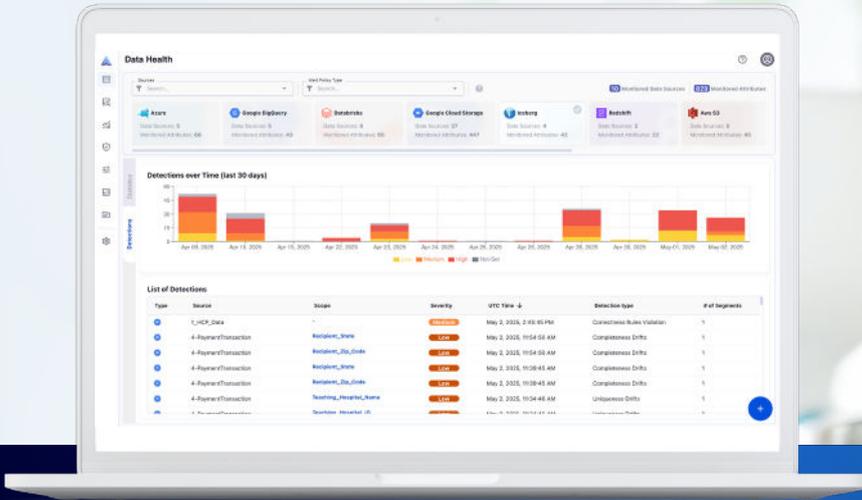
[3] Wir werden uns hier nicht mit den mathematischen Details der Metriken befassen. Weitere Informationen zu diesem Thema finden Sie in Andy Petrellas „Fundamentals of Data Observability“ (O'Reilly, 2023), Kapitel 4.

[4] Ein bahnbrechendes Werk über Data Engineering ist Joe Reis und Matthew Housleys „Fundamentals of Data Engineering“ (O'Reilly, 2022).

[5] Studien zeigen, dass mehr als 60 % der KI-Projekte bis 2026 scheitern werden, da die Daten nicht KI-fähig sind. Siehe: Gartner®: „A Journey Guide to Delivering AI Success through ‚AI-Ready‘ Data“ von Ehtisham Zaidi und Roxane Edilali (18. Oktober 2024).

[6] Ole Olesen-Bagneux, The Enterprise Data Catalogue (O'Reilly, 2023), Kapitel 7.

[7] Die technischen Details dieser Einrichtung sind zwar wichtig, jedoch zu umfangreich, um sie in dieses Whitepaper aufzunehmen. Weitere Informationen finden Sie unter [actian.com/data-observability](https://actian.com/data-observability)



## Möchten Sie Actian Data Observability in Aktion sehen?

Beginnen Sie damit, Ihre Kosten zu senken, das Vertrauen zu erhöhen und die Governance in Ihrer gesamten Datenarchitektur zu vereinfachen – schneller, als Sie es für möglich halten.

Mit dem modernen, entkoppelten Ansatz von Actian erhalten Sie volle Transparenz über die Qualität Ihrer Daten – ohne Ihre Systeme oder Ihr Budget zu belasten. Ganz gleich, ob Sie Datenprodukte skalieren, sich auf KI vorbereiten oder eine schnelle Governance durchsetzen wollen – Actian hilft Ihnen, die Kontrolle zu behalten.

Besuchen Sie [actian.com](https://actian.com) und fragen Sie nach einer **Demo**, um zu sehen, wie Actian die Art und Weise verändern kann, wie Sie Ihre Daten beobachten, verändern sowie verwalten und das Vertrauen in diese stärken können.

## Über Actian

Actian unterstützt Unternehmen dabei, umfangreiche Datenmengen sicher und effizient zu verwalten und zu steuern. Organisationen nutzen Datenmanagement- und Data-Intelligence-Lösungen von Actian, um komplexe Datenumgebungen erfolgreich zu optimieren und KI-fähige Daten schneller bereitzustellen. Die Lösungen von Actian sind flexibel und nahtlos integrierbar. Sie arbeiten problemlos in On-Premises-, Cloud- und Hybrid-Umgebungen. Erfahren Sie mehr über Actian, die Datenabteilung von HCLSoftware, unter [actian.com](https://actian.com).