

# Intelligente Datenkatalogisierung für Cloud Data Warehouses, Data Lakes und Lakehouses

## Hauptvorteile

- Maximierung des Business Value für Data Warehouses, Data Lakes und Lakehouses ohne Unterbrechung
- Schnelle Durchführung komplexer Initiativen zu Cloud-Migration und -Modernisierung
- Sicherstellung von Transparenz für alle Unternehmensdaten mit branchenführendem Metadaten-Management
- Bereitstellung von Funktionen, damit Nutzer relevante Daten in der Cloud problemlos finden

## Schnellere Data Discovery und Cloud Data Integration mit KI

Unternehmen transferieren ihre Daten und Analytics-Infrastruktur heutzutage immer schneller und öfter in die Cloud. Während die Vorteile der Cloud, wie Flexibilität, Skalierbarkeit und Kosteneffizienz, für alle Aspekte der Datenanalyse wichtig sind, sind diese jedoch insbesondere für Data Warehousing und Data Lakes dringend notwendig. Jahrelang waren Enterprise Data Warehouses richtige „Arbeitsmaschinen“, die Enterprise Analytics und Reporting-Systeme angetrieben haben. Jetzt ist es Zeit für einen Wechsel: Die Modernisierung von Data Warehouses in der Cloud, um die Performance zu erhöhen und mehr Wettbewerbsvorteile zu erlangen. Gleichzeitig nutzen Unternehmen Cloud Data Lakes, die ihre Cloud Data Warehouses ergänzen sollen. In jüngster Vergangenheit haben Unternehmen mit der Erstellung von Cloud Lakehouses begonnen, die die Vorteile von Data Warehouses und Data Lakes vereinen und somit das Rückgrad von BI-Dashboards bis zu erweiterten KI- und Machine Learning-Projekten bilden.

Doch die Nutzung von Cloud Data Warehouses, Data Lakes und Lakehouses führt auch zu neuen Herausforderungen. Obwohl die anfängliche Einrichtung eines neuen Cloud Data Warehouse oder Data Lake relativ einfach sein kann, ist ein hoher Grad an Strategie und Planung erforderlich, damit sich Ihre Investition auszahlt.

Unabhängig davon, ob Sie ein neues Cloud Data Warehouse, einen Data Lake oder ein Lakehouse erstellen oder Daten und Workloads in der Cloud modernisieren, ist es wichtig, Ihre derzeitige Datenlandschaft zu bewerten, um sicherzustellen, dass Sie die erforderlichen Tools und Best Practices haben, um Daten in der Cloud zu verwalten.

Informatica® Intelligent Cloud Lakehouse Data Management Solution for Cloud Analytics basiert auf branchenführender Datenintegration, Datenqualität, Data Governance und Metadaten-Management. Informatica Enterprise Data Catalog bietet die erforderliche Grundlage für das Metadaten-Management, damit Sie Ihre Datenlandschaft mithilfe von Data Discovery und End-to-End Transparenz, basierend auf Machine Learning, verstehen können. Informatica Intelligent Cloud Services<sup>SM</sup> (IICS) bietet marktführende iPaaS-Services für Cloud Data Integration, Cloud-Masseneinspeisung und Cloud-Datenqualität auf Enterprise-Niveau.

Mit Machine Learning-Funktionen auf Basis von umfassendem Metadaten-Management bietet Informatica Enterprise Data Catalog Unternehmen eine gemeinsame Metadaten-Grundlage für intelligente, automatisierte, End-to-End Transparenz und Data Lineage über Ihre gesamte Umgebung hinweg. Die Lösung ermöglicht es Ihnen, On-Premise Data Stores mit dem modernen Cloud-Ökosystem zu verbinden, um den Business Value zu maximieren, ohne dass Ihre Datennutzer dadurch bei ihrer Arbeit eingeschränkt werden.

## Die wichtigsten Funktionen

### Umfassende Metadaten-Konnektivität

Der intelligente Datenkatalog von Informatica lässt sich in alle führenden Cloud-Ökosysteme und moderne Data Warehouses, Data Lakes und Lakehouses integrieren, darunter auch Amazon Redshift, Amazon S3, Microsoft Azure Synapse Analytics, Google BigQuery, Snowflake und Databricks. Zudem bietet die Lösung einsatzfertige Metadaten-Konnektivität für Hunderte von Cloud-Anwendungen, On-Premise-Systeme, Enterprise- und Middleware-Anwendungen, Data Stores und analytische BI-Tools.

### Data Discovery basierend auf Machine Learning

CLAIRE<sup>®</sup>, die auf Metadaten basierende KI-Engine von Informatica, wendet KI und Machine Learning-Algorithmen auf technische, geschäftliche, betriebliche und Nutzungsdaten an. Sie versteht und analysiert komplexe Dateien und Metadaten automatisch, damit Sie Daten unabhängig vom Speicherort in hybriden und Multi-Cloud-Umgebungen ermitteln können.

### Automatisierte End-to-End Data Lineage

Data Lineage ermöglicht grobe und detaillierte Überblicke, um zu verstehen, woher Ihre Daten stammen, wie Sie sich durch Ihr Unternehmen bewegen und wie Sie innerhalb des Unternehmens umgewandelt werden. Dank End-to-End Data Lineage über hybride und Multi-Cloud-Umgebungen hinweg können Sie Daten von der Quelle bis zum Ziel nachverfolgen, um zu erfahren, wie Daten verwendet werden und was davon abhängt. Mit Informatica Enterprise Data Catalog können Sie eine detaillierte Auswirkungsanalyse auf vor- und nachgelagerte Datenbestände durchführen. So können Sie die Auswirkungen der Migration über verschiedene Datenbestände, Ressourcen und Nutzer hinweg verstehen. Zusätzlich hilft Ihnen die Auswirkungsanalyse auf Basis detaillierter Data Lineage nachzuweisen, welche Kostenersparnisse sich aus dem Transfer bestimmter Datenbestände und Workloads in die Cloud ergeben.

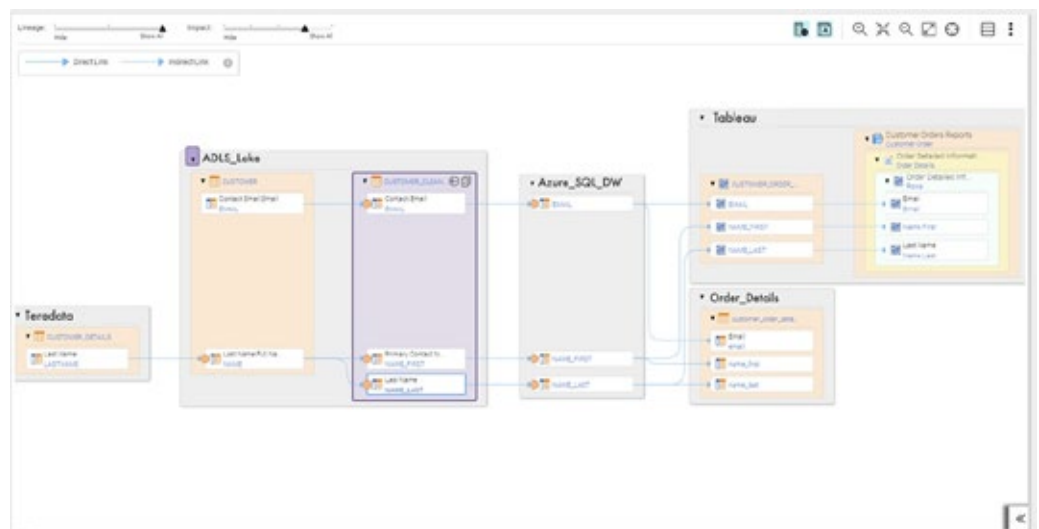


Abbildung 1. Anschauliche Darstellung von End-to-End Data Lineage über On-Premise-Systeme, Cloud Data Warehouses, Data Lakes und Lakehouses hinweg.

## Zusammenarbeit und Datenpflege mithilfe von Crowd Sourcing

Normalerweise sind verschiedene Stakeholder und Nutzer an der Migration von Daten in ein Cloud Data Warehouse, einen Data Lake oder ein Lakehouse beteiligt. Informatica Enterprise Data Catalog bietet Funktionen für die Zusammenarbeit und Datenpflege mithilfe von Crowd Sourcing, so dass Dateneigentümer und Fachgebietsexperten Datenbestände zertifizieren und die Daten identifizieren können, die verschoben werden sollten. Zudem können Datennutzer diese Datenbestände bewerten und prüfen und eigene Attribute und Anmerkungen hinzufügen, so dass zuverlässige Daten während und nach der Migration einfacher ermittelt werden können.

## Integration mit Cloud Data Management Services

Informatica Enterprise Data Catalog kann in IICS Data Management Services integriert werden, die wichtige Funktionen für Cloud Data Warehouses, Data Lakes und Lakehouses bieten. IICS Cloud Data Integration bietet hochperformante Funktionen für ETL, ELT, Dateneinspeisung, Synchronisierung und Replizierung für Multi-Cloud-Umgebungen. IICS Cloud Mass Ingestion unterstützt Streaming, Einspeisung von Dateien und Datenbanken in Data Warehouses, Data Lakes und Nachrichtenwarteschlangen. Mit IICS Cloud Data Quality können Sie Data Profiling durchführen, Daten bereinigen und standardisieren, um Probleme mit der Datenqualität schnell zu beheben und zu überwachen. Sie können Informatica Enterprise Data Catalog direkt in IICS nutzen, um Daten zu ermitteln und auszuwählen, Datenquellen und Mappings anzuzeigen, Datenqualitätsregeln und Scorecards aufzurufen und einen umfassenden Einblick in End-to-End Data Lineage zu erhalten.

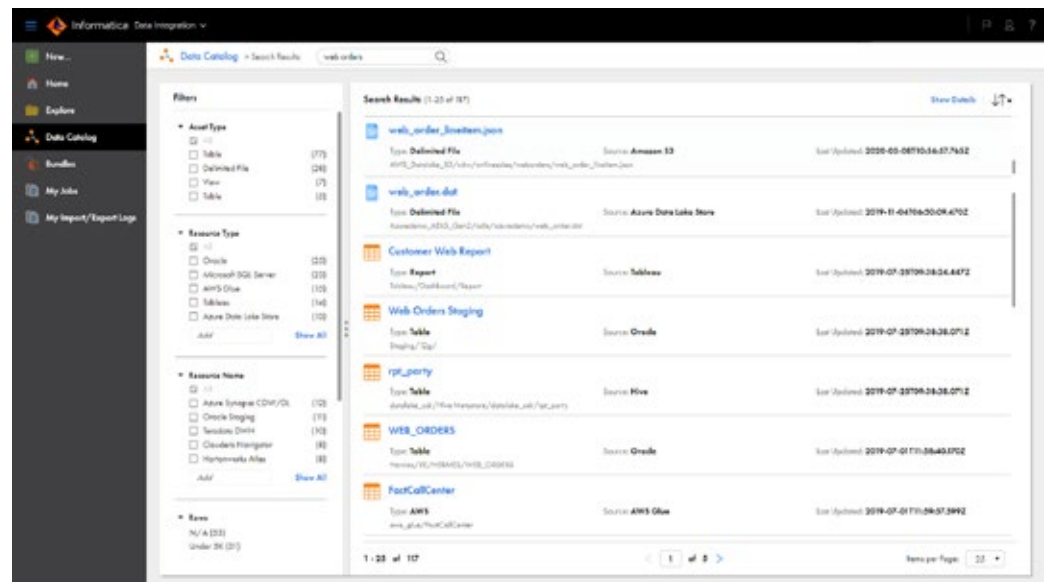


Abbildung 2. Ermittlung und Auswahl von Daten mithilfe von Informatica Enterprise Data Catalog über IICS.

### **Vorkonfigurierte Mappings**

Mit vorkonfigurierten Mapping-Templates, benutzerfreundlichen Assistenten für die Datenreplizierung, -synchronisierung und Masseneinspeisung für große Workloads können Sie mehr Daten in kürzerer Zeit in die Cloud transferieren. Mit der nahtlosen Integration zwischen Informatica Enterprise Data Catalog und IICS können Sie die Daten ermitteln, die migriert werden sollen, und sie mit einer einzigen, programmierfreien Schnittstelle problemlos in Ihr Cloud Data Warehouse, Ihren Data Lake oder Ihr Lakehouse transferieren. Nutzen Sie den visuellen Designer von IICS Cloud Data Integration zur Erstellung Ihres Datenkatalogs, um danach mit wenigen Mausklicks einfache bzw. komplexe Datenlasten zu erstellen, um verschiedene Endpunkte über Cloud- und On-Premise-Umgebungen zu verbinden, Daten umzuwandeln und analytische Workloads zu planen.

### **Hauptvorteile**

#### **Unterbrechungsfreie Maximierung des Business Value für Data Warehouses, Data Lakes und Lakehouses**

Untersuchen Sie Ihre Datenlandschaft, um zu verstehen, welche Daten migriert oder konsolidiert werden sollten und an welchen Stellen Sie den größten Business Value aus Cloud Data Warehouses, Data Lakes oder einem Lakehouse generieren können. Mit End-to-End Data Lineage erhalten Sie Einblick in Datenbeziehungen und Abhängigkeiten. Bieten Sie Datennutzern Self-Service-Zugriff, damit sie in der Cloud problemlos nach Daten für Analytics suchen können. Mit Informatica Enterprise Data Catalog können Sie sicherstellen, dass Strukturen und Schemata von Ihrem On-Premise Data Warehouse in die Cloud transferiert werden, so dass während und nach der Migration mögliche Unterbrechungen vermieden werden. Nach der Migration können Sie mithilfe von End-to-End Data Lineage-Ansichten prüfen, ob Daten korrekt zugeordnet wurden. Zudem wird skalierbare Data Governance durch automatisierte Data Discovery basierend auf Metadaten und Datenpflege mithilfe von Crowd Sourcing unterstützt, so dass zuverlässige Daten in der Cloud bereitgestellt werden können.

#### **Schnelle Durchführung komplexer Initiativen zu Cloud-Migration und -Modernisierung**

Verkürzen Sie den Zeitaufwand für Cloud-Migration und -Modernisierung. Ermitteln Sie die zu migrierenden oder konsolidierenden Daten, erstellen Sie dann mithilfe eines visuellen Mapping Designers einfache bzw. komplexe Datenintegrationslasten, die verschiedene Endpunkte über Cloud- und On-Premise-Umgebungen hinweg mithilfe vorgefertigter Transformationen verbinden, um die Entwicklung zu beschleunigen.

#### **Sicherstellung von Transparenz für alle Unternehmensdaten mit branchenführendem Metadaten-Management**

Sie müssen Ihre Daten zunächst einmal finden und verstehen, bevor Sie Business Value daraus ableiten können. Die Lösung von Informatica bietet branchenführende Funktionen für das Metadaten-Management, so dass Sie Hunderte von Konnektoren für cloudbasierte und On-Premise-Systeme nutzen können. Dank der auf Metadaten basierenden Intelligenz können Sie mithilfe der Lösung alle Orte innerhalb Ihrer Unternehmenssysteme ausfindig machen, an denen Daten gespeichert und verwendet werden. Und diese Daten können Sie dann nahtlos in Ihr Enterprise Cloud Data Warehouse, Ihren Data Lake oder Ihr Lakehouse transferieren. Zudem erfasst und speichert die Lösung wichtige Metadaten zu Notfallplänen und Beziehungen zwischen Datensätzen und Teildatensätzen, um die Verbindungen zwischen ihnen während und nach der Datenmigration und -Modernisierung in der Cloud aufrechtzuerhalten.

## Informationen zu Informatica

Die digitale Transformation verändert unsere Erwartungshaltung: besserer Service und schnellere Lieferung zu geringeren Kosten. Unternehmen müssen sich neu orientieren, um wettbewerbsfähig zu bleiben. Dabei spielen Daten eine zentrale Rolle.

Als führender Anbieter für Enterprise Cloud Data Management unterstützt Informatica Sie dabei, sich als intelligenter Vorreiter zu etablieren – völlig unabhängig davon, in welcher Branche, Kategorie oder Nische Sie tätig sind. Wir ermöglichen es Ihnen, agiler zu werden, neue Wachstumsmöglichkeiten wahrzunehmen und Innovationen voranzutreiben. Informatica ist zu 100 % auf Daten fokussiert, und bietet Unternehmen vielseitige Lösungen, um sich am Markt durchzusetzen.

Entdecken Sie jetzt das gesamte Angebot von Informatica, um das komplette Potenzial Ihrer Daten zu nutzen und so die nächste intelligente Innovation auf den Weg zu bringen.

## Bereitstellung von Funktionen, damit Nutzer relevante, zuverlässige Daten in der Cloud problemlos finden

Wenn Daten im Rahmen einer Migration transferiert werden, können Datennutzer die Daten, die sie zur Erledigung ihrer Aufgaben benötigen, mitunter nicht mehr finden. Gespeicherte Links und wiederkehrende Berichte funktionieren eventuell nicht mehr. Und Datennutzer nutzen eventuell weiterhin Daten aus einer ihnen vertrauten On-Premise-Quelle, obwohl die Cloud die bessere Alternative darstellt. Datenanalysten und Data Scientists müssen ebenfalls wissen, wo genau sich Daten befinden und einschätzen können, welche Datenbestände für ihre Zwecke am besten geeignet sind. Mithilfe von Informatica Enterprise Data Catalog können Nutzer problemlos alle Daten finden, die sie benötigen, ohne die IT um Hilfe zu bitten, so dass Engpässe und Ineffizienz vermieden werden.

## Die nächsten Schritte

In den Ressourcen [Informatica Enterprise Data Catalog – Datenblatt](#) und der Lösungsübersicht [Modernisieren Sie Ihr Data Warehouse und Ihren Data Lake in der Cloud](#) erfahren Sie mehr zu diesen Themen.



Hauptsitz Ingersheimer Str. 10, 70499 Stuttgart Tel: +49 (0) 711 139 84-0 Gebührenfrei in den USA: 1.800.653.3871

IN17\_0620\_03625

© Copyright Informatica LLC 2020. Informatica, das Logo von Informatica, Informatica Intelligent Cloud Services und CLAIRE sind Marken oder eingetragene Marken von Informatica LLC in den USA und in anderen Ländern. Die aktuelle Liste mit Marken von Informatica ist hier zu finden: <https://www.informatica.com/de/trademarks.html>. Alle weiteren Firmen- und Produktbezeichnungen können Handelsnamen oder Marken ihrer jeweiligen Eigentümer sein. Die in diesem Dokument enthaltenen Informationen können sich ohne vorherige Ankündigung ändern und werden „wie gesehen“ und ohne jegliche ausdrückliche oder stillschweigende Gewährleistung bereitgestellt.