

Informatica Enterprise Data Catalog

Vorteile

- Automatische Katalogisierung und Klassifizierung sämtlicher Datentypen im gesamten Unternehmen mit einem auf KI basierenden Datenkatalog
- Bereitstellung eines „System of Record“ für Metadaten über das gesamte Unternehmen hinweg mithilfe eines „Catalog of Catalogs“
- Automatische Extraktion detaillierter Metadaten aus verschiedensten Datenquellen, selbst aus komplexen Enterprise-Systemen
- Leistungsstarke, semantische Suchfunktionen (die der Google-Suche ähneln) zum Auffinden von Datenbeständen
- Ermittlung und Verständnis von Datenbeständen dank einer ganzheitlichen Ansicht, einschließlich Data Lineage, Beziehungen, Data Profiling-Statistiken und Scorecards zur Datenqualität
- Identifizierung von Domänen und Datenentitäten mit intelligenter Datenverwaltung
- Anreicherung von Datenbeständen mit verwalteten und Crowdsourcing-Kommentaren, Bewertungen und Rezensionen
- Automatische Zuordnung von Begriffen des Geschäftsglossars zu technischen Datensätzen
- Offene APIs zur Integration in Ihre Umgebung, um intelligente Metadaten zu finden – unabhängig vom Speicherort

Nutzen Sie das Potenzial Ihrer Daten mit einem intelligenten Datenkatalog

In der Geschäftswelt dreht sich alles um Daten: Datenorientierte Unternehmen schöpfen aus ihren Datenbeständen Umsatz und Gewinn. Der erste Schritt jeder Initiative zur digitalen Transformation besteht darin, Daten als Unternehmensressource zu verwalten, d. h. Bestandsaufnahme, Wertermittlung und Nutzenoptimierung – ganz ähnlich wie bei anderen wichtigen Kapital- und Betriebsinvestitionen auch.

Daten sind verschiedenartig und auf unterschiedliche Abteilungen, Anwendungen, Data Warehouses und Data Lakes verteilt (sowohl On-Premise als auch in der Cloud), was einen präzisen Überblick über sämtliche Daten und deren Speicherort erschwert. Und aufgrund der Zunahme von Datenquellen steigt auch die Komplexität der Datenlandschaft.

Informatica® Enterprise Data Catalog ist eine KI-basierte Lösung, die mit einer auf Machine Learning basierenden Discovery Engine Datenbestände unternehmensweit scannen und katalogisieren kann – über Cloud- und On-Premise-Anwendungen hinweg sowie für Big Data. Enterprise Data Catalog basiert auf der CLAIRE® Engine, die Metadaten nutzt, um intelligente Empfehlungen und Vorschläge zu erstellen und Aufgaben zum Datenmanagement zu automatisieren. So steigt die Produktivität von IT- und Business Usern bei der Verwaltung und Nutzung von Daten.

Informatica Enterprise Data Catalog bietet Datenanalysten und IT-Anwendern leistungsstarke semantische Suchfunktionen und dynamische Filtermöglichkeiten, Data Lineage, Profiling-Statistiken, Scorecards zur Datenqualität, ganzheitliche Ansichten von Beziehungen, Empfehlungen zur Datenähnlichkeit sowie ein integriertes Geschäftsglossar.

Dank der Funktionen zur Zusammenarbeit können Fachwissen und Crowd Sourcing in Kombination mit KI genutzt werden, um die Nutzererlebnisse zu verbessern und die Datenpflege zu automatisieren. So können Sie Unternehmensdatenbestände einfach und effizient verwalten, um deren Wert für das gesamte Unternehmen zu maximieren. Nutzer können Daten schnell finden und mühelos den Lebenszyklus von Geschäftsbegriffen, Definitionen, Referenzdaten und vielem mehr verwalten.

Informatica Enterprise Data Catalog, eine Lösung, die einen auf KI-basierenden Datenkatalog nutzt, umfasst auch eine auf Machine Learning basierende Discovery Engine, um Datenbestände unternehmensweit zu scannen und zu katalogisieren – über Cloud- und On-Premise-Anwendungen hinweg sowie für Big Data.

Die wichtigsten Funktionen

Metadaten-APIs, die in Ihre Umgebung integriert werden können

Informatica Enterprise Data Catalog umfasst auf REST basierende APIs, die in Ihre Umgebung integriert werden können, so dass Kataloginhalte überall genutzt werden können. Unternehmen können Business Usern intelligente Metadaten, wie Anwendungen, BI-Berichte und Dashboards, bereitstellen. Nutzer können ausgewählte Kataloginhalte sowie dazugehörige Metadaten exportieren und teilen.

Enterprise Data Catalog Tableau Erweiterung für Self-Service Analytics

Die Tableau Erweiterung für Informatica Enterprise Data Catalog bietet agile Self-Service Analytics mit verwalteten Daten. So können Tableau Nutzer innerhalb der nativen Tableau Benutzeroberfläche auf sämtliche Ressourcen von Informatica Enterprise Data Catalog zugreifen. Business Usern, die Tableau Berichte nutzen, steht die intelligente Suchleiste innerhalb von Tableau zur Verfügung, um Datenbestände zu finden, die von Data Stewards zertifiziert sind. Zudem haben sie Zugriff auf betrieblichen und technischen Kontext, Datenqualität, Profiling-Statistiken und Empfehlungen von Kollegen. Nutzer können auch gemeinsam mit ihren Kollegen Bewertungen, Rezensionen und häufig gestellte Fragen zu den Datenbeständen erstellen – alles über die Benutzeroberfläche von Tableau.

Semantische Suchfunktion mit intelligenten Filtermöglichkeiten

Finden Sie mithilfe der leistungsstarken, semantischen Suchfunktionen mit intelligenten Filtermöglichkeiten relevante Datensätze für Ihre Analysen. Durch eine erweiterte Stichwortsuche mit Tokenabgleich lassen sich die relevantesten Datenbestände im Katalog ermitteln. Bei der semantischen Suche werden auch abgeleitete Datendomänen berücksichtigt, so dass kein Datenbestand unerkannt bleibt. Dank intelligenter Filtermöglichkeiten in Abhängigkeit von den Suchergebnissen können Nutzer die Suche nach den für sie wichtigen Datensätzen eingrenzen.

Self-Service-Datenbereitstellung

Nachdem Sie die relevanten Datensätze für Ihre Analysen gefunden haben, können Sie den Datensatz innerhalb von Informatica Enterprise Data Catalog mit wenigen Mausklicks in das gewünschte Ziel verschieben. Ihnen steht eine große Auswahl an Quellen und Zielen zur Verfügung, darunter auch Amazon Redshift, Azure Synapse Analytics (früher: Azure SQL Data Warehouse), Google BigQuery, Snowflake und BI-Tools wie Tableau. Die Funktion nutzt die Integration von Informatica Enterprise Data Catalog in Informatica Cloud Data Integration.

Data Lineage und Auswirkungsanalyse

Verfolgen Sie die Herkunft von Daten interaktiv mithilfe von Data Lineage-Ansichten auf verschiedenen Ebenen – von der übersichtlichen Ansicht auf Systemebene, die die Endpunkte hervorhebt, bis zu detaillierten Ansichten mit komplexen Informationen. In einer Drilldown-Lineage-Ansicht können beliebige Herkunftspfade erweitert werden, um die einzelnen Spalten und Metriken des Lineage-Diagramms anzuzeigen. Nutzer können detaillierte Auswirkungsanalysen für vor- und nachgelagerte Datenbestände durchführen.

Ganzheitliche Erkennung von Datenbeziehungen

Mithilfe einer ganzheitlichen Ansicht der Daten in einem Wissensdiagramm können Sie Unternehmensdaten und aussagekräftige Datenbeziehungen zügig suchen, ermitteln und nachvollziehen. Dabei lassen sich verwandte Datensätze sowie technische, geschäftliche, semantische und verwendungsbasierte Beziehungen automatisch ermitteln. In der ganzheitlichen Datenansicht werden verwandte Datensätze, Tabellen, Ansichten, Datendomänen, Berichte und Nutzer angezeigt. Dies erleichtert die progressive Ermittlung anderer relevanter Datensätze.

Automatisierte Klassifizierungen mit intelligenter Erkennung von Datendomänen und -entitäten

Datendomänen und -entitäten werden über alle strukturierten und unstrukturierten Datenbestände auf Feld-, Spalten- und Tabellenebene automatisch klassifiziert und identifiziert, beispielsweise Kunden-, Produkt- oder Auftragsdaten. Dies ist ein entscheidender Schritt für Unternehmen bei der Katalogisierung, Verwaltung und nutzbringenden Verwertung ihrer Datenbestände. Dank dieser klassifizierten Daten können bessere Empfehlungen für Suchvorgänge, für die Filterung von Suchergebnissen und das Geschäftsglossar erstellt werden. Informatica bietet über 60 vorgefertigte Datendomänen wie E-Mail-Adresse, Kreditkartennummer, Sozialversicherungsnummer, Land, Ort, URL und Unternehmensname. Darüber hinaus können Nutzer eigene Domänen hinzufügen. Die Klassifizierung der Datenbestände kann anhand von Datenregeln (Spalten mit Daten zum Abgleich einer spezifischen Logik gemäß Regel) oder Spaltennamensregeln (Spalten, die mit der Spaltennamenslogik gemäß Regel übereinstimmen) erfolgen.

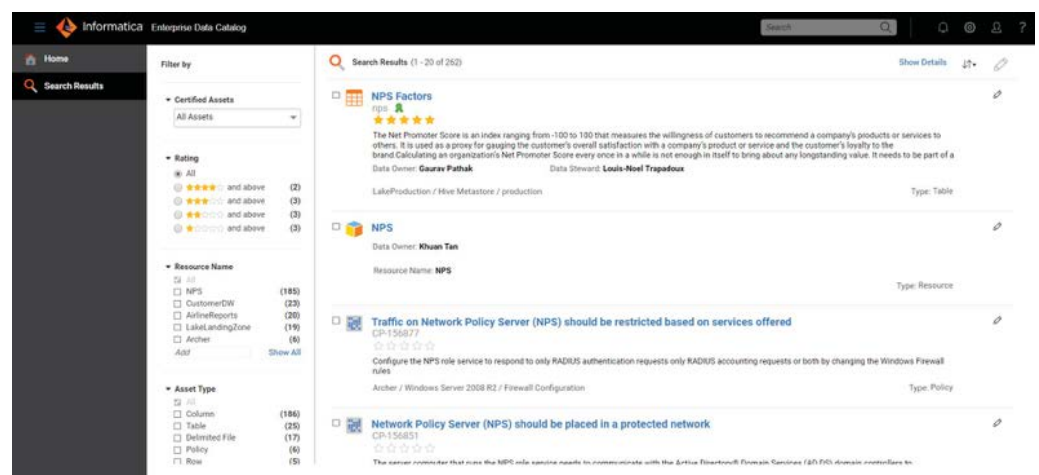


Abbildung 1: Intelligente semantische Suchfunktionen und dynamische Filtermöglichkeiten beschleunigen das Auffinden von Datensätzen. Zeigen Sie Bewertungen und zertifizierte Datensätze an.

Zusammenarbeit und Datenpflege mithilfe von Crowd Sourcing

Mit Informatica Enterprise Data Catalog können Datenanalysten und Data Scientists problemlos relevante und zuverlässige Daten für Analytics finden, indem sie KI, menschliches Fachwissen und die Vorteile der Zusammenarbeit nutzen. Dateneigentümer und Fachgebietsexperten können Datensätze zertifizieren. Datennutzer können Bewertungen und Rezensionen für Datensätze erstellen, um Crowd Sourcing zu nutzen.

Nutzer können Datensätze verfolgen, die für sie von Interesse sind und über Änderungen informiert werden. Auf einer Frage- und Antwort-Plattform können Fachgebietsexperten gängige Fragen von Nutzern beantworten. Zudem können Nutzer individuelle Attribute und Anmerkungen zu Datensätzen hinzufügen, um die Zusammenarbeit zwischen IT- und Betrieb noch zu vertiefen und die Suchergebnisse zu verbessern.

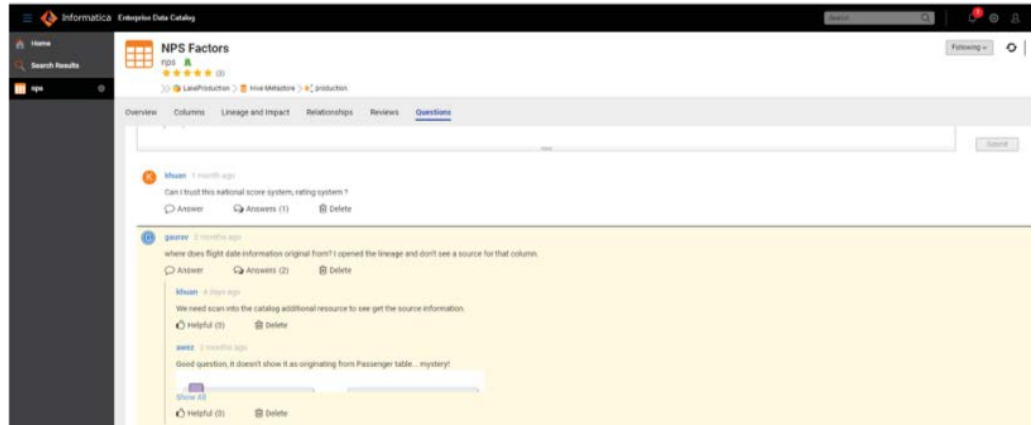


Abbildung 2: Mit einer Funktion für Fragen und Antworten kann die Zusammenarbeit verbessert werden.

Integrierte Datenqualität

Data Profiling-Statistiken, Regeln zur Sicherstellung der Datenqualität, Scorecards und Metrikengruppen lassen sich zusammen mit technischen Metadaten anzeigen, um die Qualität der Datenbestände vor ihrer Verwendung zu Analyse Zwecken zu bewerten. Diese Profiling-Statistiken umfassen Wertverteilungen, Muster sowie Inferenzen von Datentypen und Datendomänen.

Automatische Zuordnung der Begriffe des Geschäftsglossars

Informatica Enterprise Data Catalog ermöglicht den einfachen Import von Geschäftsglossarbeständen, wie Begriffen, Richtlinien und Klassifizierungen aus Informatica Axon™. Zudem können Sie die Daten durch geschäftliche Kontextinformationen anreichern, indem Geschäftsbegriffe mit den richtigen technischen Metadaten verknüpft werden. Informatica Enterprise Data Catalog ermöglicht die automatische Zuordnung von Geschäftsbegriffen mit technischen Metadaten, so dass der aufwändige, manuelle Prozess für Data Governance wegfällt. So können geschäftliche und IT-Verwalter gemeinsam geschäftliche Metadaten verwalten und dabei von der effizienten Automatisierung menschlicher Arbeitsabläufe profitieren. Informatica Enterprise Data Catalog unterstützt auch den Import von Geschäftsglossarbeständen aus Informatica Business Glossary und den Tools anderer Anbieter.

Intelligente Datenähnlichkeit

Erweiterte statistische und Machine-Learning-Algorithmen identifizieren ähnliche Daten und Teilmengen von Daten. Diese leistungsstarke Funktion erleichtert Nutzern das Auffinden der relevantesten und zuverlässigsten Daten. Wenn beispielsweise ein Telekommunikationsanalyst die Kundenabwanderung analysiert, könnte er Daten zur Prepaid-Aktivität von Kunden im aktuellen Quartal abfragen. Informatica Enterprise Information Catalog kann eine bereinigte Version der Daten (Substitutionsdaten), Daten zur Kundenaktivität im Vorquartal (vereinbare Daten) und eine Kundendetailtabelle (verknüpfbare Daten) zur Anreicherung des Datensatzes empfehlen.

Universelle Metadatenkonnektivität mit Advanced Scanners

Sie können Metadaten aus sämtlichen Datenquellen im Unternehmen extrahieren, beispielsweise aus Datenbanken, Data Warehouses, cloudbasierten Data Lakes, BI-Tools, Hadoop-Clustern, NoSQL und komplexen Unternehmenssystemen, darunter auch veraltete und Mainframe-Systeme, ETL-Tools verschiedener Anbieter, SQL-Dialekten und verschiedenen Unternehmensanwendungen – über hybride und Multi-Cloud-Umgebungen hinweg.

Mit Enterprise Data Catalog Advanced Scanners können Sie jedes Script, jeden Ablauf und jeden Prozess visuell darstellen, um die Logik und den internen Datenfluss zu verstehen. Sie erhalten umfassende Data Lineage auf Spaltenebene, darunter auch einen Überblick über sämtliche Quellen samt Details. Sie können sowohl statischen als auch dynamischen Code scannen und Sprache verarbeiten, um automatisch Data Lineage-Informationen zu erhalten.

Die folgenden Datenquellen zur Extraktion von Metadaten werden beispielsweise unterstützt:

- **Datenbanken und Data Warehouses:** Oracle, Microsoft SQL Server, SQL Scripts, Sybase ASE, IBM Netezza, Teradata, JDBC, SAP HANA, SAP BW, SAP BW/4HANA, gespeicherte Abläufe
- **Big Data:** Cloudera Navigator, Hive (Cloudera/Hortonworks/MapR/IBM BigInsights/EMR), HDFS, Hortonworks Atlas, Cassandra, MongoDB, Kafka
- **Mainframes:** DB2 z/OS, DB2 i5/OS, COBOL, JCL
- **BI und Analytics:** SAP BusinessObjects, Tableau, Microsoft Power BI, Cognos, MicroStrategy, OBIEE, QlikView, Qlik Sense, Microsoft SSRS und SSAS, SAS
- **ETL:** Informatica PowerCenter, Informatica Data Engineering Integration, Informatica Intelligent Cloud Services, Informatica Data Integration Hub, Microsoft SSIS, Oracle Data Integrator, IBM InfoSphere DataStage, AWS Glue
- **Geschäftsglossar:** Informatica Axon Data Governance, Informatica Business Glossary
- **Datenmodellierung:** Erwin Data Modeler, SAP PowerDesigner
- **Unternehmensanwendungen:** Salesforce, Oracle, Workday, Informatica MDM, SAP ECC
- **Dateisysteme:** Microsoft SharePoint, Microsoft OneDrive, Windows/Linux-Dateisysteme
- **Dateiformate:** MS Excel, MS Word, MS PowerPoint, Adobe PDF, Flat Files, CSV, Delimited, XML, JSON, Avro, Parquet
- **Cloud-Plattformen:** AWS S3, AWS Redshift, Azure SQL DB, Azure SQL DW, Azure ADLS, Azure ADLS Gen 2, Azure Blob, Google Cloud Storage, Google BigQuery, Snowflake

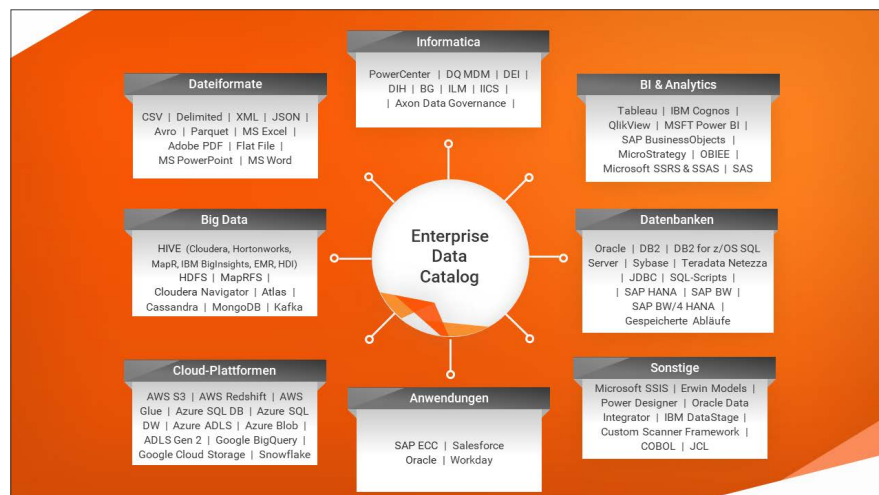


Abbildung 3: Informatica Enterprise Data Catalog unterstützt universelle Metadatenkonnektivität.

Sicherheit auf Ressourcenebene

Lese- und Schreibberechtigungen für Nutzer und Gruppen können auf Ressourcenebene erteilt werden, so dass Nutzer benutzerdefinierte Attribute anzeigen oder bearbeiten, Domänen pflegen und Geschäftsglossarbegriffe verknüpfen können.

Unternehmensweite Umsetzung

Informatica Enterprise Data Catalog ist für die Umsetzung in Big-Data-Umgebungen geeignet und kann mehrere Millionen von Datensätzen über Hunderte von Datenquellen hinweg scannen. Die parallele Metadatenaufnahme und die ultraschnelle verteilte Indexierung ermöglichen eine schnelle Aktualisierung der Kataloginhalte und eine beispiellose Suchperformance sowie eine hohe Verfügbarkeit mit Fehlertoleranz für Rund-um-die-Uhr-Implementierungen. Mit dem auf Spark basierendem Data Profiling können Sie riesige Datenmengen analysieren, um Ihre Unternehmensdaten besser zu verstehen.

Einheitliche Administration

Verwaltung und Monitoring von Katalogressourcen, Zeitplänen für die Metadatenextraktion, Profiling-Läufen und weiteren Aufgaben erfolgen über eine einheitliche Administrationskonsole. Ein Dashboard zur Aufgabenkontrolle enthält Widgets für das Monitoring von Aufgaben sowie Ressourcenansichten. E-Mail-Benachrichtigungen unterstützen Administratoren bei der proaktiven Behebung von Katalogproblemen.



Abbildung 4: Umfassende, ganzheitliche Ansichten der Datenbeziehungen erleichtern die Nachvollziehbarkeit.

Vorteile

Intelligente Katalogisierung verschiedenster Datentypen unternehmensweit

Informatica Enterprise Data Catalog ermittelt zahlreiche Datentypen sowie deren Beziehungen im gesamten Unternehmen. Vorgefertigte Scanner erfassen Metadaten aus Datenbanken, Data Warehouses, Data Lakes, Cloud Data Stores, Anwendungen, BI-Tools, ETL-Tools, Metadatenkatalogen von Drittanbietern, NoSQL usw. Sämtliche Metadaten werden in einer hochgradig skalierbaren grafischen Datenbank indiziert und katalogisiert, die auf schnelle Aktualisierungen, intelligente Suchvorgänge und schnelle Abfragen ausgelegt ist.

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 innovativer Vorreiter zu etablieren – völlig unabhängig von Ihrer Branche, Kategorie oder Nische. 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.

Je mehr Daten erstellt und im gesamten Unternehmen verteilt werden, desto unvermeidlicher wird die Entstehung ähnlicher und duplizierter Datensätze. Informatica Enterprise Data Catalog ermittelt mithilfe erweiterter statistischer und Machine-Learning-Algorithmen ähnliche Daten und Teilmengen von Daten, so dass Nutzer leichter die relevantesten und zuverlässigsten Daten auffinden können.

Leistungsstarke, semantische Suchfunktionen (die der Google-Suche ähneln) zum Auffinden von Datenbeständen

Die Suche nach den benötigten Daten in Hunderten Unternehmenssystemen kann schnell aussichtslos erscheinen. Die einzige Erfolgsaussicht besteht dabei in leistungsstarken semantischen Suchfunktionen auf der Grundlage umfassender Metadaten-Erkenntnisse in einer skalierbaren Infrastruktur. Informatica Enterprise Data Catalog enthält diese semantischen Suchfunktionen mit intelligenten Filtermöglichkeiten zur weiteren Eingrenzung der Suchergebnisse. Da Informatica geschäftliche, technische und betriebliche Metadaten eindeutig miteinander verknüpft, können Business User ihre Daten anhand von Geschäftsbegriffen suchen und anschließend eine ganzheitliche Beziehungsansicht nach verwandten Datenbeständen durchsuchen.

Ermittlung und besseres Verständnis von Datenbeständen dank ganzheitlicher Beziehungsansichten und Data Lineage

Die Redewendung „Was du nicht messen kannst, kannst du nicht lenken“ lässt sich auch auf die Verwaltung von Datenbeständen übertragen. Um Daten optimal nutzen zu können, müssen Sie wissen, welche Daten vorhanden sind, woher sie stammen, wie sie umgewandelt wurden und wie zuverlässig sie sind. Informatica Enterprise Data Catalog beantwortet nicht nur diese Fragen, sondern bietet zudem zusammenfassende und detaillierte Ansichten zu Data Lineage, Profiling-Statistiken, Scorecards zur Datenqualität sowie ganzheitliche Beziehungsansichten, die ein klares Bild Ihrer Daten liefern.

Anreicherung von Datenbeständen mit Geschäftskontext durch verwaltete und Crowdsourcing-Anmerkungen

Informatica Enterprise Data Catalog setzt auf die maximale Wiederverwendung von Daten und deren Wert, indem Unternehmensdatenbestände automatisch bis auf Feld-/Spaltenebene klassifiziert werden. Um den Wert der Daten noch weiter zu steigern, erfasst Informatica Enterprise Data Catalog neben den Nutzern der Daten und dem Zweck der Datennutzung auch Crowdsourcing-Tags und Anmerkungen, Bewertungen und Rezensionen. Mithilfe dieser „Weisheit der Vielen“ lassen sich Daten anreichern und pflegen, so dass sie noch wertvoller für das Unternehmen werden. Informatica Enterprise Data Catalog lässt sich in Informatica Axon integrieren, damit Geschäftsglossarbestände, wie Begriffe, Definitionen und Richtlinien problemlos aus Axon importiert werden können. Diese geschäftlichen Metadaten werden automatisch mit technischen und betrieblichen Metadaten verknüpft, so dass Geschäftsanalysten, Datenverwalter und andere Nutzer die Datenbestände schnell auffinden, nachvollziehen und gemeinsam verwenden können.

Weitere Informationen

Unter <https://www.informatica.com/de/products/data-catalog/enterprise-data-catalog.html> erfahren Sie mehr über Informatica Enterprise Data Catalog.



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

IN06_0620_03238

© Copyright Informatica LLC 2020. Informatica, das Logo von Informatica, CLAIRE, Axon, PowerCenter, Big Data Management und Informatica Cloud 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.