



Informatica®  
10.4.0

# Handbuch Neue Funktionen

Diese Software und die Dokumentation werden nur im Rahmen eines eigenen Lizenzvertrags zur Verfügung gestellt, der Beschränkungen für die Verwendung und Weitergabe enthält. Ohne ausdrückliche schriftliche Genehmigung der Informatica LLC darf kein Teil dieses Dokuments zu irgendeinem Zweck vervielfältigt oder übertragen werden, unabhängig davon, auf welche Art und Weise oder mit welchen Mitteln (elektronisch, mechanisch, durch Fotokopieren, Aufzeichnen usw.) dies geschieht.

U.S. GOVERNMENT RIGHTS Programs, software, databases, and related documentation and technical data delivered to U.S. Government customers are "commercial computer software" or "commercial technical data" pursuant to the applicable Federal Acquisition Regulation and agency-specific supplemental regulations. As such, the use, duplication, disclosure, modification, and adaptation is subject to the restrictions and license terms set forth in the applicable Government contract, and, to the extent applicable by the terms of the Government contract, the additional rights set forth in FAR 52.227-19, Commercial Computer Software License.

Informatica, the Informatica logo, Big Data Management, PowerCenter, and PowerExchange are trademarks or registered trademarks of Informatica LLC in the United States and many jurisdictions throughout the world. A current list of Informatica trademarks is available on the web at <https://www.informatica.com/trademarks.html>. Other company and product names may be trade names or trademarks of their respective owners.

Portions of this software and/or documentation are subject to copyright held by third parties. Required third party notices are included with the product.

Die in dieser Dokumentation enthaltenen Informationen können jederzeit ohne vorherige Ankündigung geändert werden. Wenn Sie Probleme in dieser Dokumentation finden, melden Sie sie uns unter [infa\\_documentation@Informatica.com](mailto:infa_documentation@Informatica.com).

Informatica-Produkte unterliegen einer Gewährleistung gemäß den Geschäftsbedingungen der Vereinbarungen, unter denen sie bereitgestellt werden. INFORMATICA STELLT DIE INFORMATIONEN IN DIESEM DOKUMENT OHNE MÄNGELGEWÄHR UND OHNE AUSDRÜCKLICHE ODER STILLSCHWEIGENDE GEWÄHRLEISTUNG JEDLICHER ART ZUR VERFÜGUNG. DIES GILT EINSCHLIESSLICH FÜR GEWÄHRLEISTUNGEN DER MARKTGÄNGIGKEIT, DER EIGNUNG FÜR EINEN BESTIMMTEN ZWECK UND GEWÄHRLEISTUNGEN ODER ZUSICHERUNGEN ÜBER DIE NICHTVERLETZUNG VON RECHTEN DRITTER.

# Inhalt

|   |           |
|---|-----------|
| <b>Einleitung .....</b>   | <b>8</b>  |
| Informatica-Ressourcen. ....  | 8         |
| Informatica-Netzwerk. ....  | 8         |
| Informatica-Wissensdatenbank. ....  | 8         |
| Informatica-Dokumentation. ....   | 8         |
| Informatica-Produktverfügbarkeitsmatrizen. ....                             | 9         |
| Informatica Velocity. ....  | 9         |
| Informatica Marketplace. ....   | 9         |
| Globaler Kundensupport von Informatica. ....                                | 9         |
| <br><b>Kapitel 1: Neue Funktionen (10.4.0).....</b>                         | <b>10</b> |
| CI/CD. ....   | 10        |
| Datenintegrationsdienst-REST-API. ....                                      | 10        |
| infacmd dis-Befehle. ....   | 12        |
| Reverse-Proxy-Server. ....  | 12        |
| infacmd roh-Befehle. ....   | 13        |
| Bereitstellung von Anwendungs-Patches. ....                                 | 13        |
| Mit Laufzeit-Anwendung verbinden. ....                                      | 14        |
| Ansicht „Objekt-Explorer“. ....   | 14        |
| Tags. ....  | 15        |
| Befehlszeilenprogramme. ....  | 15        |
| infacmd isp-Befehle (neue Funktionen in Version 10.4.0). ....               | 16        |
| Data Engineering Integration. ....  | 16        |
| Unterstützung neuer Datentypen. ....  | 16        |
| AWS Databricks-Integration. ....  | 17        |
| Cluster-Arbeitsabläufe für HDInsight-Zugriff auf ALDS Gen2-Ressourcen. .... | 17        |
| Databricks Delta Lake-Speicherzugriff. ....                                 | 17        |
| In Zuordnungen verwendete Knoten anzeigen. ....                             | 18        |
| Protokollaggregation. ....  | 18        |
| Analysieren von hierarchischen Daten auf der Spark-Engine. ....             | 18        |
| Profile und Stichprobenoptionen auf der Spark-Engine. ....                  | 18        |
| Python-Umwandlung. ....   | 19        |
| Sqoop. ....   | 19        |
| Data Engineering Streaming. ....  | 20        |
| Konfluente Schemaregistrierung in Streaming-Zuordnungen. ....               | 20        |
| Datenqualitätsumwandlungen in Streaming-Zuordnungen. ....                   | 20        |
| Ephemeral-Cluster in Streaming-Zuordnungen. ....                            | 20        |
| FileName-Port in Amazon S3. ....  | 21        |
| Microsoft Azure Data Lake Storage Gen2. ....                                | 21        |
| Streaming-Zuordnungen in Azure Databricks. ....                             | 21        |

|  |    |
|--|----|
| Dynamische Zuordnungen in Data Engineering Streaming . . . . .                           | 22 |
| Enterprise Data Catalog. . . . .   | 22 |
| Zuweisen von benutzerdefinierten Attributen zu Ressourcen und Klassen. . . . .           | 22 |
| Neue Ressourcen. . . . .   | 23 |
| Referenzressourcen und Referenzobjekte. . . . .  | 23 |
| Objekte aus Suchergebnissen exportieren. . . . .   | 23 |
| Herkunfts- und Auswirkungsfiler. . . . .   | 24 |
| Übersicht über Objektsteuerung. . . . .  | 24 |
| Regeln und Scorecards. . . . .   | 24 |
| Eindeutige Schlüsselableitung. . . . .   | 24 |
| Datendomänenerkennung für den CLOB-Dateityp. . . . .                                     | 24 |
| Data Discovery und Stichprobenoptionen auf der Spark-Engine. . . . .                     | 25 |
| Nachverfolgen der technischen Vorschau. . . . .  | 25 |
| Datenvorschau und Bereitstellung. . . . .  | 25 |
| Unterstützte Ressourcentypen für das Standalone Scanner Utility. . . . .                 | 26 |
| REST-APIs. . . . .   | 27 |
| Enterprise Data Preparation. . . . .   | 27 |
| Data Lake-Zugriffsverwaltung. . . . .  | 27 |
| Microsoft Azure Data Lake Storage als Datenquelle. . . . .                               | 27 |
| Dateien im Data Lake veröffentlichen. . . . .  | 27 |
| Dateien in den Data Lake hochladen. . . . .  | 28 |
| Informatica-Zuordnungen . . . . .  | 28 |
| Binden von Zuordnungsausgaben an Zuordnungsparameter. . . . .                            | 28 |
| CLAIRE-Empfehlungen und -Einblicke. . . . .  | 29 |
| Zuordnungsoptimierungsebene aktualisieren. . . . .                                       | 29 |
| Informatica-Umwandlungen. . . . .  | 30 |
| Adress-Validiererumwandlung. . . . .   | 30 |
| Arbeitsabläufe in Informatica. . . . .   | 30 |
| Erweiterte Eigenschaften für die Cluster-Erstellungsaufgabe mit Amazon EMR. . . . .      | 30 |
| Installation von Informatica. . . . .  | 30 |
| PostgreSQL. . . . .  | 31 |
| Systemprüfungstool vor der Installation (i10Pi) im automatischen Modus. . . . .          | 31 |
| Passwörter in der Eigenschaftsdatei für automatische Installation verschlüsseln. . . . . | 31 |
| Intelligentes Strukturmodell. . . . .  | 31 |
| Zusätzliche Eingabetypen. . . . .  | 31 |
| Modell zur Entwurfszeit aus Beispiel erstellen. . . . .                                  | 31 |
| Handhabung nicht identifizierter Daten. . . . .  | 32 |
| PowerCenter. . . . .   | 32 |
| Konnektivität. . . . .   | 32 |
| PowerExchange-Adapter. . . . .   | 33 |
| PowerExchange-Adapter für Informatica. . . . .   | 33 |
| PowerExchange-Adapter für PowerCenter. . . . .   | 36 |

|   |           |
|---|-----------|
| Sicherheit. . . . .   | 38        |
| Webanwendungen für die Verwendung von verschiedenen SAML-Identitätsanbietern konfigurieren. . . . . | 38        |
| <b>Kapitel 2: Neue Funktionen (10.2.2 HotFix 1).....</b>  | <b>39</b> |
| Befehlszeilenprogramme . . . . .  | 39        |
| infacmd Idm-Befehle. . . . .  | 39        |
| Enterprise Data Catalog. . . . .  | 39        |
| Azure Data Lake Storage Gen2. . . . .   | 40        |
| Verknüpfungen ohne Unterscheidung der Groß-/Kleinschreibung . . . . .                               | 40        |
| Enterprise Data Catalog Tableau Extension . . . . .   | 40        |
| Neue Ressourcen. . . . .  | 40        |
| Profiling-Warehouse-Scanner. . . . .  | 40        |
| REST-APIs. . . . .  | 41        |
| Suche Verbesserungen . . . . .  | 41        |
| Suche Registerkarten. . . . .   | 41        |
| Unterstützte Ressourcentypen für das Standalone Scanner Utility. . . . .                            | 42        |
| Technical Preview. . . . .  | 42        |
| <b>Kapitel 3: Neue Funktionen (10.2.2 Service Pack 1).....</b>                                      | <b>43</b> |
| Big Data Management. . . . .  | 43        |
| Sqoop. . . . .  | 43        |
| Big Data Streaming. . . . .   | 44        |
| Amazon S3-Ziel. . . . .   | 44        |
| TIME_RANGE-Funktion. . . . .  | 44        |
| Schreiben von Daten in mehrere HDFS-Dateien. . . . .  | 44        |
| Enterprise Data Catalog. . . . .  | 45        |
| Unterstützte Ressourcentypen für das Standalone Scanner Utility. . . . .                            | 45        |
| Enterprise Data Preparation. . . . .  | 45        |
| Zurücksetzen aller abgeleiteten Datentypen. . . . .   | 45        |
| Vorbereiten von Avro- und Parquet-Dateien. . . . .  | 45        |
| PowerExchange-Adapter für Informatica. . . . .  | 46        |
| PowerExchange for Hive. . . . .   | 46        |
| <b>Kapitel 4: Neue Funktionen (10.2.2).....</b>   | <b>47</b> |
| Anwendungsdienste. . . . .  | 47        |
| Massenerfassungsdienst. . . . .   | 47        |
| Metadaten-Zugriffsdienst. . . . .   | 47        |
| REST Operations Hub-Dienst. . . . .   | 48        |
| Big Data Management. . . . .  | 48        |
| Integration in Azure Databricks. . . . .  | 48        |
| Hierarchische Daten. . . . .  | 49        |
| Hohe Verfügbarkeit. . . . .   | 50        |

|  |    |
|--|----|
| Intelligentes Strukturmodell. . . . .  | 51 |
| Massenerfassung. . . . .   | 51 |
| Überwachung. . . . .   | 52 |
| Sicherheit. . . . .  | 52 |
| Ziele. . . . .   | 52 |
| Big Data Streaming . . . . .   | 53 |
| Azure Event Hubs-Datenobjekte. . . . .   | 53 |
| Kontenübergreifende IAM-Rolle in Amazon Kinesis-Verbindungen. . . . .  | 53 |
| Intelligentes Strukturmodell. . . . .  | 53 |
| Header-Ports für Big Data Streaming-Datenobjekte. . . . .  | 53 |
| AWS-Anmeldedatenprofile in Amazon Kinesis-Verbindungen. . . . .  | 54 |
| Strukturiertes Spark-Streaming. . . . .  | 54 |
| Fensterumwandlung. . . . .   | 54 |
| Befehlszeilenprogramme. . . . .  | 54 |
| infacmd dis-Befehle. . . . .   | 54 |
| infacmd ihs-Befehle. . . . .   | 55 |
| infacmd ipc-Befehle. . . . .   | 56 |
| infacmd ldm-Befehle. . . . .   | 56 |
| infacmd mi-Befehle. . . . .  | 57 |
| infacmd ms-Befehle. . . . .  | 57 |
| infacmd oie-Befehle. . . . .   | 58 |
| infacmd tools-Befehle. . . . .   | 58 |
| infasetup-Befehle. . . . .   | 58 |
| Enterprise Data Catalog. . . . .   | 59 |
| Automatische Zuweisung der Geschäftsbezeichnung zu einer Spalte. . . . .   | 59 |
| Benutzerzusammenarbeit bei Objekten. . . . .   | 59 |
| Erstellung der Anwendungsdienste von Enterprise Data Catalog unter Verwendung des<br>Installationsprogramms. . . . . | 59 |
| Dienstprogramm zur Validierung benutzerdefinierter Metadaten. . . . .  | 60 |
| Änderungsbenachrichtigungen. . . . .   | 60 |
| Business Glossary Zuweisungsbericht. . . . .   | 60 |
| Betriebssystemprofile. . . . .   | 60 |
| REST-APIs. . . . .   | 60 |
| Quellmetadatenfilter und Datenprofilfilter. . . . .  | 61 |
| Scanner-Dienstprogramm. . . . .  | 61 |
| Ressourcentypen. . . . .   | 61 |
| Enterprise Data Lake. . . . .  | 62 |
| Anwendung aktiver Regeln. . . . .  | 62 |
| Löschung doppelter Zeilen. . . . .   | 62 |
| Clusterung und Kategorisierung von Spaltendaten . . . . .  | 62 |
| CLAIRE-basierte Empfehlungen. . . . .  | 62 |
| Bedingte Aggregation. . . . .  | 63 |
| Datenmaskierung. . . . .   | 63 |

|   |    |
|---|----|
| Lokalisierung. . . . .  | 63 |
| Partitionierte Quellen und Ziele. . . . .                     | 63 |
| Hinzufügung von Kommentaren zu Rezeptschritten. . . . .       | 64 |
| Speicherung eines Rezepts als Mapping. . . . .                | 64 |
| Amazon S3, ADLS, WASB und MapR-FS als Datenquellen. . . . .   | 64 |
| Statistische Funktionen. . . . .                              | 64 |
| Funktionen für Datum und Uhrzeit. . . . .                     | 65 |
| Mathematische Funktionen. . . . .                             | 65 |
| Textfunktionen. . . . .                                       | 66 |
| Fensterfunktionen. . . . .                                    | 66 |
| Audit-Ereignisse bereinigen. . . . .                          | 67 |
| Spark-Ausführungs-Engine. . . . .                             | 67 |
| Informatica Developer. . . . .                                | 67 |
| Anwendungen. . . . .  | 67 |
| Informatica-Zuordnungen. . . . .                              | 67 |
| Datentypen. . . . .   | 68 |
| Zuordnungsausgaben. . . . .                                   | 68 |
| Zuordnungsparameter. . . . .                                  | 68 |
| Optimierungsebenen. . . . .                                   | 68 |
| Sqoop. . . . .  | 68 |
| Informatica-Umwandlungen. . . . .                             | 69 |
| Adress-Validiererumwandlung. . . . .                          | 69 |
| Updatestrategie-Umwandlung. . . . .                           | 72 |
| PowerExchange-Adapter für Informatica. . . . .                | 72 |
| PowerExchange for Amazon Redshift. . . . .                    | 73 |
| PowerExchange for Amazon S3. . . . .                          | 73 |
| PowerExchange for Google BigQuery. . . . .                    | 74 |
| PowerExchange for HBase. . . . .                              | 74 |
| PowerExchange for HDFS. . . . .                               | 75 |
| PowerExchange for Hive. . . . .                               | 75 |
| PowerExchange for MapR-DB. . . . .                            | 76 |
| PowerExchange for Microsoft Azure Blob Storage. . . . .       | 76 |
| PowerExchange for Microsoft Azure Cosmos DB SQL API. . . . .  | 77 |
| PowerExchange for Microsoft Azure Data Lake Store. . . . .    | 77 |
| PowerExchange for Microsoft Azure SQL Data Warehouse. . . . . | 77 |
| PowerExchange for Salesforce. . . . .                         | 78 |
| PowerExchange for Snowflake. . . . .                          | 78 |
| PowerExchange for Teradata Parallel Transporter API. . . . .  | 79 |

# Einleitung

Lesen Sie das *Informatica®-Handbuch für neue Funktionen und Verbesserungen*, um sich über die neuen Funktionen und Verbesserungen in Informatica-Produkten für diese Version zu informieren.

## Informatica-Ressourcen

Informatica stellt Ihnen über das Informatica-Netzwerk und andere Online-Portale zahlreiche Produktressourcen zur Verfügung. Nutzen Sie die Ressourcen, um Ihre Informatica-Produkte und -Lösungen optimal zu nutzen und von anderen Informatica-Benutzern und Fachspezialisten zu lernen.

### Informatica-Netzwerk

Das Informatica-Netzwerk bietet Zugriff auf zahlreiche Ressourcen, darunter die Informatica-Wissensdatenbank und der globale Kundensupport von Informatica. Um auf das Informatica-Netzwerk zuzugreifen, besuchen Sie <https://network.informatica.com>.

Als Mitglied des Informatica-Netzwerks haben Sie die folgenden Optionen:

- Durchsuchen Sie die Wissensdatenbank nach Produktressourcen.
- Zeigen Sie Informationen zur Produktverfügbarkeit an.
- Erstellen und überprüfen Sie Ihre Supportfälle.
- Ihr lokales Informatica-Netzwerk für Benutzergruppen suchen und mit anderen Benutzern zusammenarbeiten.

### Informatica-Wissensdatenbank

In der Informatica-Wissensdatenbank finden Sie Produktressourcen wie beispielsweise praktische Anleitungen, Best Practices, Videotutorials und Antworten auf häufig gestellte Fragen.

Zum Durchsuchen der Wissensdatenbank besuchen Sie <https://search.informatica.com>. Wenn Sie Fragen, Kommentare oder Ideen zur Wissensdatenbank haben, wenden Sie sich per E-Mail an das Team der Informatica-Wissensdatenbank unter [KB\\_Feedback@informatica.com](mailto:KB_Feedback@informatica.com).

### Informatica-Dokumentation

Verwenden Sie das Informatica-Dokumentationsportal, um in einer umfangreichen Dokumentationsbibliothek nach aktuellen und neuen Produktversionen zu suchen. Um das Dokumentationsportal zu erkunden, besuchen Sie <https://docs.informatica.com>



Wenn Sie Fragen, Kommentare oder Ideen zur Produktdokumentation haben, wenden Sie sich an das Informatica-Dokumentationsteam unter [infa\\_documentation@informatica.com](mailto:infa_documentation@informatica.com)

## Informatica-Produktverfügbarkeitsmatrizen

Produktverfügbarkeitsmatrizen (PAMs) geben die Versionen der Betriebssysteme, Datenbanken und Typen von Datenquellen und Zielen an, die in einer Produktversion unterstützt werden. Sie können die Informatica-PAMs unter <https://network.informatica.com/community/informatica-network/product-availability-matrices> durchsuchen.

## Informatica Velocity

Informatica Velocity ist eine Sammlung von Tipps und Best Practices, die von den Professionellen Informatica-Diensten entwickelt wurden und auf praktischen Erfahrungen aus Hunderten von Datenmanagementprojekten basieren. Informatica Velocity umfasst das gesammelte Wissen von Informatica-Beratern, die mit Unternehmen auf der ganzen Welt zusammenarbeiten, um erfolgreiche Datenmanagementlösungen zu planen, zu entwickeln, bereitzustellen und zu warten.

Die Informatica Velocity-Ressourcen finden Sie unter <http://velocity.informatica.com>. Wenn Sie Fragen, Anregungen oder Ideen zu Informatica Velocity haben, wenden Sie sich an die professionellen Informatica-Dienste unter [ips@informatica.com](mailto:ips@informatica.com).

## Informatica Marketplace

Informatica Marketplace ist ein Forum, das Lösungen zur Erweiterung und Verbesserung Ihrer Informatica-Implementierungen bereitstellt. Nutzen Sie die zahlreichen Lösungen von Informatica-Entwicklern und -Partnern im Marketplace, um Ihre Produktivität zu steigern und die Implementierungsdauer Ihrer Projekte zu verkürzen. Den Informatica Marketplace finden Sie unter <https://marketplace.informatica.com>.

## Globaler Kundensupport von Informatica

Sie können sich telefonisch oder über das Informatica-Netzwerk an ein Global Support-Center wenden.

Die Telefonnummer des globalen Kundensupports von Informatica vor Ort finden Sie auf der Informatica-Website unter folgender Verknüpfung:

<https://www.informatica.com/services-and-training/customer-success-services/contact-us.html>.

Um im Informatica-Netzwerk nach Online-Supportressourcen zu suchen, besuchen Sie

<https://network.informatica.com> und wählen Sie die eSupport-Option aus.

# KAPITEL 1

## Neue Funktionen (10.4.0)

Dieses Kapitel umfasst die folgenden Themen:

- [CI/CD, 10](#)
- [Befehlszeilenprogramme, 15](#)
- [Data Engineering Integration, 16](#)
- [Data Engineering Streaming , 20](#)
- [Enterprise Data Catalog, 22](#)
- [Enterprise Data Preparation, 27](#)
- [Informatica-Zuordnungen , 28](#)
- [Informatica-Umwandlungen, 30](#)
- [Arbeitsabläufe in Informatica, 30](#)
- [Installation von Informatica, 30](#)
- [Intelligentes Strukturmodell, 31](#)
- [PowerCenter, 32](#)
- [PowerExchange-Adapter, 33](#)
- [Sicherheit, 38](#)

## CI/CD

In diesem Abschnitt werden Verbesserungen bei CI/CD in Version 10.4.0 erläutert.

CI/CD bzw. Continuous Integration (fortlaufende Integration) und Continuous Delivery (fortlaufende Zustellung) ist eine Praxis, die die Integrations- und Zustellungsvorgänge in einer CI/CD-Pipeline automatisiert. In Version 10.4.0 können Sie die Verbesserungen in Ihre CI/CD-Pipeline integrieren, um zu optimieren, wie Objekte bereitgestellt, getestet und an die Produktionsumgebung zugestellt werden.

## Datenintegrationsdienst-REST-API

Ab Version 10.4.0 können Sie die Datenintegrationsdienst-REST-API nutzen, um Aufgaben in einer CI/CD-Pipeline zu automatisieren.

Zu den Aufgaben, die die REST-API automatisieren kann, zählen die Folgenden:

**Objekte abfragen.**

Fragen Sie Objekte ab, einschließlich Entwurfszeitobjekten in einem Modellrepository und Laufzeitobjekten, die an einen Datenintegrationsdienst bereitgestellt werden.

Sie können die Abfrage an andere REST-API-Anfragen übergeben. Sie können beispielsweise eine Abfrage an einen Versionskontrollvorgang übergeben, um die Versionskontrolle für eine bestimmte Gruppe von Objekten durchzuführen. Sie können auch eine Abfrage übergeben, um bestimmte Entwurfszeitobjekte an eine Anwendungs-Patch-Archivdatei bereitzustellen.

**Versionskontrollvorgänge durchführen.**

Führen Sie Versionskontrollvorgänge zum Einchecken, Auschecken, Rückgängigmachen eines Auscheckvorgangs oder Neuzuweisen eines ausgecheckten Entwurfszeitobjekts an einen anderen Entwickler durch.

**Tags verwalten.**

Verwalten Sie die Tags, die Entwurfszeitobjekten zugewiesen sind. Sie können ein neues Tag zuweisen oder Tags für ein Objekt ersetzen. Sie können auch Tags von einem Objekt entfernen.

**Anwendungen aktualisieren.**

Stellen Sie Entwurfszeitobjekte an eine Anwendungs-Patch-Archivdatei bereit und stellen Sie die Datei an einen Datenintegrationsdienst bereit, um eine bereitgestellte inkrementelle Anwendung zu aktualisieren.

**Anwendungen verwalten.**

Sie können eine Anwendung starten, deren Bereitstellung aufheben oder sie anhalten, je nach den Anforderungen des Projekts oder der Organisation.

**Bereitgestellte Zuordnungen ausführen.**

Führen Sie eine bereitgestellte Zuordnung aus, um die Anwendungsausgabe zu testen.

**Zuordnungen vergleichen.**

Vergleichen Sie zwei Zuordnungen in derselben Domäne.

Sie können beispielsweise zwei Entwurfszeitzuordnungen oder eine Entwurfszeitzuordnung und eine Laufzeitzuordnung vergleichen.

Um die REST-API-Anfragen, die Sie verwenden können, und die Parameter für jede Anfrage anzuzeigen, sehen Sie die REST-API-Dokumentation über die Datenintegrationsdienst-Prozesseigenschaften oder die REST-Operations Hub Service-Eigenschaften im Administrator Tool ein.

Im Vergleich zu den infacmd-Befehlszeilenprogrammen gelten für die REST-API keine Setup-Anforderungen, und Sie können die REST-API in Umgebungen ausführen, in denen keine Informatica-Dienste auf dem Clientcomputer installiert sind.

Weitere Informationen zur REST-API finden Sie im Kapitel „Datenintegrationsdienst-REST-API“ im *Informatica 10.4.0-Anwendungsdienst-Handbuch*.

## infacmd dis-Befehle

Ab Version 10.4.0 können Sie mithilfe der infacmd dis-Befehle Vorgänge für Entwurfszeitobjekte, Laufzeitobjekte in einer Anwendung und bereitgestellte Zuordnungen ausführen. Mit den infacmd dis-Befehlen können Sie auch Datenintegrationsdienst-Hilfsprogramme ausführen.

Die folgende Tabelle beschreibt die neuen infacmd dis-Befehle:

| Befehl                 | Beschreibung   |
|------------------------|--|
| compareMapping         | Vergleicht zwei abgefragte Zuordnungen. Fragen Sie die Zuordnungen ab, um Zuordnungseigenschaften, Umwandlungseigenschaften und Ports innerhalb von Umwandlungen zu vergleichen. Um Entwurfszeitzuordnungen abzufragen, geben Sie das Entwurfszeit-Modellrepository an. Um Laufzeitzuordnungen abzufragen, geben Sie kein Modellrepository an. In der Abfrage wird der Datenintegrationsdienst verwendet, den Sie zum Ausführen des Befehls angeben. |
| deployObjectsToFile    | Stellt Entwurfszeitobjekte in einer Anwendungs-Patch-Archivdatei bereit.   |
| queryDesignTimeObjects | Fragt Entwurfszeitobjekte aus einem Modellrepository-Dienst ab.  |
| queryRunTimeObjects    | Fragt Laufzeitobjekte ab, die in einem Datenintegrationsdienst bereitgestellt sind, und gibt eine Liste von Objekten zurück.   |
| replaceAllTag          | Ersetzt Tags durch die angegebenen Tags für die abgefragten Objekte im Modellrepository-Dienst.  |
| tag                    | Weist ein Tag zu den abgefragten Objekten im Modellrepository-Dienst zu.   |
| untag                  | Entfernt ein Tag aus den abgefragten Objekten im Modellrepository-Dienst.  |
| listPatchNames         | Listet alle Patches auf, die auf eine inkrementelle Anwendung angewendet wurden.   |

Weitere Informationen finden Sie im Kapitel „infacmd dis-Befehlsreferenz“ der *Informatica® 10.4.0-Befehlsreferenz*.

## Reverse-Proxy-Server

Ab Version 10.4.0 können Sie mit dem REST Operations Hub den Reverse-Proxy-Server starten, der den Lastausgleich für die Anfragen der Datenintegrationsdienst-API über Knoten in einem Datenintegrationsdienst-Raster hinweg durchführt.

Der REST Operations Hub ist rasterfähig.

Weitere Informationen zum Reverse-Proxy-Server finden Sie im Kapitel „Systemdienste“ im *Informatica 10.4.0-Anwendungsdienst-Handbuch*.

## infacmd roh-Befehle

Ab Version 10.4.0 können Sie die folgenden infacmd roh-Befehle zum Aktualisieren des REST Operations Hub Service-Prozesses verwenden.

In der folgenden Tabelle werden die neuen infacmd roh-Befehle beschrieben:

| Befehle                         | Beschreibung  |
|---------------------------------|---|
| listReverseProxyServerOptions   | Listet die Reverse-Proxy-Server-Eigenschaften auf.                    |
| listServiceProcessOptions       | Listet die Prozesseigenschaften des REST Operations Hub-Dienstes auf. |
| updateServiceOptions            | Aktualisiert die Eigenschaften für den REST Operations Hub Service.   |
| updateReverseProxyServerOptions | Aktualisiert die Reverse-Proxy-Server-Eigenschaften in der Domäne.    |

Ab Version 10.4.0 wurden die folgenden infacmd roh-Befehle umbenannt:

- listROHProperties in listProcessProperties.
- updateROHService in updateServiceProcessOptions.

**Hinweis:** Aktualisieren Sie alle Skripts, in denen der vorherige Befehlsname verwendet wird.

Weitere Informationen finden Sie im Kapitel „infacmd roh-Befehlsreferenz“ der *Informatica 10.4.0-Befehlsreferenz*.

## Bereitstellung von Anwendungs-Patches

Ab Version 10.4.0 verfügt die Bereitstellung von Anwendungs-Patches über neue Funktionen in den folgenden Bereichen:

### Assistent zur inkrementellen Bereitstellung

Ab Version 10.4.0 können Sie die folgenden Aufgaben im Assistenten zur inkrementellen Bereitstellung durchführen:

- Optional können Sie eine Patch-Beschreibung eingeben.
- Optional können Sie wählen, ob Sie Statusinformationen beibehalten oder verwerfen möchten. Statusinformationen verweisen auf Zuordnungseigenschaften und die Eigenschaften von Laufzeitobjekten, wie z. B. Zuordnungsausgaben oder die Sequenzgeneratorumwandlung.

Weitere Informationen zum Assistenten zur inkrementellen Bereitstellung finden Sie im Kapitel zur Bereitstellung von Anwendungs-Patches im *Informatica 10.4.0 Developer Tool-Handbuch*.

Weitere Informationen zu Statusinformationen finden Sie im Kapitel „Anwendungsbereitstellung“ im *Informatica 10.4.0 Developer Tool-Handbuch*.

### Patch-Verlauf

Ab Version 10.4.0 zeigt der Patch-Verlauf im Assistenten zur inkrementellen Bereitstellung sowohl den Patch-Namen als auch die Patch-Beschreibung der Patches an, die zum Aktualisieren der inkrementellen Anwendung bereitgestellt wurden. Die Uhrzeit der Patch-Erstellung wird am Beginn der Patch-Beschreibung angefügt.

Zusätzlich können Sie das Administrator Tool verwenden, um den Patch-Verlauf für eine bereitgestellte inkrementelle Anwendung anzuzeigen.

Weitere Informationen zum Patch-Verlauf finden Sie im Kapitel zur Bereitstellung von Anwendungs-Patches im *Informatica 10.4.0 Developer Tool-Handbuch*.

Weitere Informationen zu den bereitgestellten Anwendungen finden Sie im Kapitel zu Datenintegrationsdienst-Anwendungen im *Informatica 10.4.0-Anwendungsdiensthandbuch*.

### Anwendungs-Patch-Archivdateien

Ab Version 10.4.0 speichern Anwendungs-Patch-Archivdateien zusätzlich die Patch-Beschreibung und geben an, ob Statusinformationen beibehalten werden.

Weitere Informationen zum Patch-Verlauf finden Sie im Kapitel zur Bereitstellung von Anwendungs-Patches im *Informatica 10.4.0 Developer Tool-Handbuch*.

### infacmd tools-Befehle

Ab Version 10.4.0 verfügt der Befehl „infacmd tools patchApplication“ über die folgende neue Option:

| Option                          | Argument   | Beschreibung   |
|---------------------------------|------------|--|
| -RetainStateInformation<br>-rsi | True False | Optional. Gibt an, ob die Statusinformationen beibehalten oder verworfen werden. Statusinformationen verweisen auf Zuordnungseigenschaften und die Eigenschaften von Laufzeitobjekten, wie z. B. Zuordnungsausgaben oder die Sequenzgeneratorumwandlung. |

Weitere Informationen finden Sie im Kapitel „infacmd tools-Befehlsreferenz“ in der *Informatica 10.4.0-Befehlsreferenz*.

## Mit Laufzeit-Anwendung verbinden

Ab Version 10.4.0 können Sie eine Verbindung zu einer Laufzeitanwendung im Developer Tool herstellen. Nachdem Sie eine Verbindung zu einer Laufzeitanwendung hergestellt haben, können Sie die Anwendung in der Ansicht **Objekt-Explorer** erweitern und schreibgeschützte Kopien der Laufzeitobjekte im **Editor** öffnen.

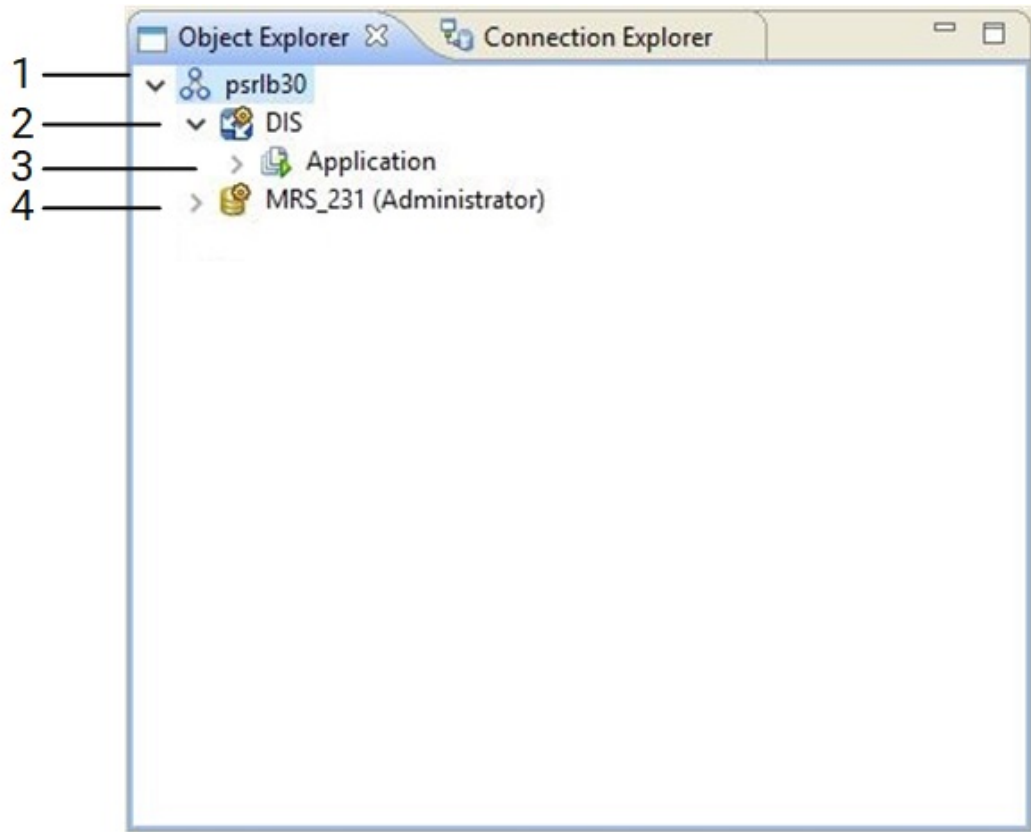
Nachdem Sie eine Verbindung zu einer Laufzeitanwendung hergestellt haben, werden mit den Suchvorgängen, die Sie im Developer Tool durchführen, Laufzeitobjekte in der Anwendung gefunden.

Weitere Informationen zum Herstellen einer Verbinden mit einer Laufzeitanwendung und zum Anzeigen der Laufzeitobjekte finden Sie im Kapitel „Anwendungsbereitstellung“ im *Informatica 10.4.0-Developer Tool-Handbuch*

## Ansicht „Objekt-Explorer“

Ab Version 10.4.0 zeigt die Ansicht **Objekt-Explorer** im Developer Tool die Domäne an, nachdem Sie eine Verbindung zu einem Modellrepository oder einer Laufzeitanwendung hergestellt haben. Sie können die Domäne erweitern, um die Entwurfszeitobjekte im Modellrepository oder die Laufzeitobjekte in der Laufzeitanwendung anzuzeigen.

Die folgende Abbildung zeigt die Ansicht **Objekt-Explorer**:



1. Domäne
2. Datenintegrationsdienst
3. Laufzeitanwendung
4. Modellrepository

Weitere Informationen zur Benutzeroberfläche im Developer Tool finden Sie im Kapitel „Informatica Developer“ im *Informatica 10.4.0 Developer Tool-Handbuch*.

## Tags

Ab Version 10.4.0 haben Tags die folgende Funktion:

- Wenn Sie eine Zuordnung bereitstellen, die einem Tag zugeordnet ist, wird das Tag an die Laufzeitversion der Zuordnung im Datenintegrationsdienst verteilt.
- Wenn Sie die bereitgestellte Zuordnung anhand eines Anwendungs-Patches aktualisieren, wird der Patch-Name der Laufzeitversion der Zuordnung als Tag zugeordnet.

Weitere Informationen zu Tags finden Sie im Kapitel „Informatica Developer“ im *Informatica 10.4.0 Developer Tool-Handbuch*.

## Befehlszeilenprogramme

In diesem Abschnitt werden neue Befehle in Version 10.4.0 erläutert.

## infacmd isp-Befehle (neue Funktionen in Version 10.4.0)

In der folgenden Tabelle werden neue infacmd isp-Befehle beschrieben:

| Befehl                  | Beschreibung   |
|-------------------------|--|
| addCustomLDAPType       | Fügt einen benutzerdefinierten LDAP-Typ hinzu, der einen LDAP-Verzeichnisdienst definiert.   |
| listAllCustomLDAPTypes  | Listet die Konfigurationsinformationen für alle benutzerdefinierten LDAP-Typen auf, die von der angegebenen Domäne verwendet werden. |
| listAllLDAPConnectivity | Listet die Konfigurationsinformationen für alle LDAP-Konfigurationen auf, die von der angegebenen Domäne verwendet werden.           |
| listCustomLDAPType      | Listet die Konfigurationsinformationen für einen benutzerdefinierten LDAP-Typ auf.   |
| removeCustomLDAPType    | Entfernt den angegebenen benutzerdefinierten LDAP-Typ aus der angegebenen Domäne.  |
| removeLDAPConnectivity  | Entfernt die angegebene LDAP-Konfiguration aus der angegebenen Domäne.   |
| updateCustomLDAPType    | Aktualisiert den angegebenen benutzerdefinierten LDAP-Typ.   |
| updateLDAPConnectivity  | Aktualisiert die angegebene LDAP-Konfiguration.  |

Weitere Informationen finden Sie in der *Informatica 10.4.0-Befehlsreferenz*.

## Data Engineering Integration

In diesem Abschnitt werden neue Data Engineering Integration-Funktionen in Version 10.4.0 erläutert.

### Unterstützung neuer Datentypen

Ab Version 10.4.0 können Sie die folgenden neuen Datentypen für komplexe Dateien verwenden:

- Wenn Sie eine Zuordnung verwenden, die komplexe Avro- und Parquet-Dateiobjekte in der nativen Umgebung oder in der Hadoop-Umgebung liest oder schreibt, können Sie die folgenden Datentypen verwenden:
  - Datum
  - Dezimal
  - Zeitstempel
- Sie können den Uhrzeit-Datentyp verwenden, um komplexe Avro- oder Parquet-Dateiobjekte in der nativen Umgebung oder auf der Blaze-Engine zu lesen und zu schreiben.
- Sie können anwendbare Datum-, Uhrzeit-, Zeitstempel- und Dezimal-Datentypen verwenden, wenn Sie eine Zuordnung auf der Databricks-Spark-Engine ausführen.

Die neuen Datentypen gelten für die folgenden Adapter:

- PowerExchange for HDFS
- PowerExchange for Amazon S3



- PowerExchange for Google Cloud Storage
- PowerExchange for Microsoft Azure Blob Storage
- PowerExchange for Microsoft Azure Data Lake Storage Gen1
- PowerExchange for Microsoft Azure Data Lake Storage Gen2

Weitere Informationen zu Datentypen finden Sie im Kapitel „Datentyppräferenz“ im *Data Engineering Integration 10.4.0-Benutzerhandbuch*.

## AWS Databricks-Integration

Ab Version 10.4.0 können Sie die Informatica-Domäne mit Databricks on AWS integrieren.

Sie können AWS Databricks nutzen, um Zuordnungen mit den folgenden Funktionen auszuführen:

- Sie können Zuordnungen mit Amazon Simple Storage Service (S3)- und Amazon Redshift-Quellen und -Zielen innerhalb der Databricks-Umgebung ausführen.
- Sie können Cluster-Arbeitsabläufe entwickeln, um anhand von Databricks on AWS Ephemeral-Cluster zu erstellen.
- Sie können die Python-Umwandlung zu einer Zuordnung hinzufügen, die für die Ausführung auf der Databricks-Spark-Engine konfiguriert ist.  
Die Python-Umwandlung wird für die technische Vorschau unterstützt.

AWS Databricks unterstützt die gleichen Datentypen wie Azure Databricks.

Weitere Informationen finden Sie in den folgenden Handbüchern:

*Data Engineering 10.4.0-Integrationshandbuch*

*Data Engineering 10.4.0-Administratorhandbuch*

*Data Engineering Integration 10.4.0-Benutzerhandbuch*

*Informatica 10.4.0 Developer-Arbeitsablaufhandbuch*

## Cluster-Arbeitsabläufe für HDInsight-Zugriff auf ALDS Gen2-Ressourcen

Ab Version 10.4.0 können Sie einen Cluster-Arbeitsablauf erstellen, der auf einem Azure HDInsight-Cluster ausgeführt wird, um auf ADLS Gen2-Ressourcen zuzugreifen.

Weitere Informationen zu Cluster-Arbeitsabläufen finden Sie im *Informatica Data Engineering Integration 10.4.0-Benutzerhandbuch*.

## Databricks Delta Lake-Speicherzugriff

Ab Version 10.4.0 können Sie auf Databricks Delta Lake-Speicher als Quellen und Ziele zugreifen.

Zuordnungen können auf Delta Lake-Ressourcen auf den AWS- und Azure-Plattformen zugreifen.

Informationen zum Konfigurieren des Zugriffs auf Delta Lake-Tabellen finden Sie unter *Data Engineering Integration Guide*. Informationen zum Erstellen von Zuordnungen für den Zugriff auf Delta Lake-Tabellen finden Sie unter *Data Engineering Integration User Guide*.

## In Zuordnungen verwendete Knoten anzeigen

Ab Version 10.4.0 können Sie die maximale Zahl der Clusterknoten anzeigen, die von einer Zuordnung während einer bestimmten Zeitdauer verwendet werden.

Mit der REST Operations Hub-API `ClusterStats(startTimeInmillis=[value], endTimeInmillis=[value])` können Sie die maximale Anzahl an Hadoop-Knoten für eine Clusterkonfiguration anzeigen, die von einer Zuordnung während einer bestimmten Zeitdauer verwendet wird.

Weitere Informationen zur REST-API finden Sie im Referenzkapitel zur Überwachung der REST-API im *Data Engineering 10.4.0-Administratorhandbuch*.

## Protokollaggregation

Ab Version 10.4.0 können Sie aggregierte Protokolle für bereitgestellte Zuordnungen erhalten, die in der Hadoop-Umgebung ausgeführt werden.

Sie können die aggregierten Clusterprotokolle für eine Zuordnung basierend auf der Job-ID im Monitoring Tool erfassen, bzw. verwenden Sie den Befehl `infacmd ms fetchAggregatedClusterLogs`. Sie können eine .zip- oder tar.gz-Datei der aggregierten Clusterprotokolle für eine Zuordnung basierend auf der Job-ID erhalten und die komprimierte aggregierte Protokolldatei in ein Zielverzeichnis schreiben.

Weitere Informationen finden Sie im *Informatica 10.4.0-Administratorhandbuch*.

## Analysieren von hierarchischen Daten auf der Spark-Engine

Ab 10.4.0 können Sie komplexe Funktionen verwenden, um bis zu 5 MB Daten Midstream in einer Zuordnung zu analysieren.

Die Spark-Engine kann Roh-Zeichenfolgenquelldaten anhand der folgenden komplexen Funktionen analysieren:

- PARSE\_JSON
- PARSE\_XML

Die komplexen Funktionen analysieren JSON- oder XML-Daten in der Quellzeichenfolge und generieren strukturierte Zieldaten.

Weitere Informationen finden Sie im Kapitel „Hierarchische Datenverarbeitung“ im *Informatica Data Engineering Integration 10.4.0-Benutzerhandbuch*.

Weitere Informationen zu komplexen Funktionen finden Sie im Kapitel „Funktionen“ im *Developer-Referenzhandbuch für die Umwattungssprache von Informatica 10.4.0*.

## Profile und Stichprobenoptionen auf der Spark-Engine

Ab Version 10.4.0 können Sie Profile ausführen und Stichprobenoptionen für die Spark-Engine wählen.

### Profilerstellung auf der Spark-Engine

Sie können Profile auf der Spark-Engine im Informatica Developer Tool und im Informatica Analyst Tool erstellen und ausführen. Sie können die Datendomänenerkennung ausführen und Scorecards auf der Spark-Engine erstellen.

## Stichprobenoptionen auf der Spark-Engine

Sie können die folgenden Stichprobenoptionen wählen, um Profile auf der Spark-Engine auszuführen:

- Mit der Stichprobenoption **Limit n** wird ein Profil basierend auf der Anzahl der Zeilen im Datenobjekt ausgeführt. Wenn Sie ein Profil in der Hadoop-Umgebung ausführen möchten, sammelt die Spark-Engine Stichproben aus mehreren Partitionen des Datenobjekts und überträgt die Stichproben an einen einzelnen Knoten, um die Stichprobengröße zu berechnen. Sie können die Stichprobenoptionen für „limit n“ nicht auf Profile mit erweitertem Filter anwenden. Die Funktion wird für Oracle-, SQL Server- und DB2-Datenbanken unterstützt.
- Mit der Stichprobenoption **Random percentage** wird ein Profil basierend auf dem Prozentsatz der Zeilen im Datenobjekt ausgeführt.

Weitere Informationen zu den Profilen und Stichprobenoptionen auf der Spark-Engine finden Sie im *Informatica 10.4.0 Data Discovery-Handbuch*.

## Python-Umwandlung

Ab Version 10.4.0 hat die Python-Umwandlung die folgende Funktion:

### Aktiver Modus

Sie können eine aktive Python-Umwandlung erstellen. Als aktive Umwandlung kann die Python-Umwandlung die Anzahl der sie durchlaufenden Zeilen ändern. Beispielsweise kann die Python-Umwandlung mehrere Ausgabezeilen aus einer einzelnen Eingabezeile generieren, oder die Umwandlung kann eine einzige Ausgabezeile aus mehreren Eingabezeilen generieren.

Weitere Informationen finden Sie im Kapitel „Python-Umwandlung“ im *Informatica Data Engineering Integration 10.4.0-Benutzerhandbuch*.

### Partitionierte Daten

Sie können Python-Code ausführen, um eingehende Daten basierend auf dem Standard-Partitionierungsschema der Daten zu verarbeiten, oder Sie können die Daten erneut partitionieren, bevor der Python-Code ausgeführt wird. Um die Daten erneut zu partitionieren, bevor der Python-Code ausgeführt wird, wählen Sie einen oder mehrere Eingabeports als Partitionsschlüssel aus.

Weitere Informationen finden Sie im Kapitel „Python-Umwandlung“ im *Informatica Data Engineering Integration 10.4.0-Benutzerhandbuch*.

## Sqoop

Ab Version 10.4.0 können Sie die folgenden Sqoop-Argumente in der JDBC-Verbindung konfigurieren:

- --update-key
- --update-mode
- --validate
- --validation-failurehandler
- --validation-threshold
- --validator
- --mapreduce-job-name
- --bindir
- --class-name
- --jar-file

- --outdir
- --package-name

Weitere Informationen zum Konfigurieren dieser Sqoop-Argumente finden Sie in der Sqoop-Dokumentation.

## Data Engineering Streaming

In diesem Abschnitt werden neue Data Engineering Streaming-Funktionen in Version 10.4.0 erläutert.

### Konfluente Schemaregistrierung in Streaming-Zuordnungen

Ab Version 10.4.0 können Sie Confluent Kafka als Quellen und Ziele in Streaming-Zuordnungen verwenden, die die Schemaregistrierung nutzen.

Sie können Confluent Kafka zum Speichern und Abrufen von Apache Avro-Schemas in Streaming-Zuordnungen verwenden. Bei der Schemaregistrierung wird Kafka als zugrunde liegender Speichermechanismus verwendet.

Weitere Informationen finden Sie im *Data Engineering Streaming 10.4.0-Benutzerhandbuch*.

### Datenqualitätsumwandlungen in Streaming-Zuordnungen

Ab Version 10.4.0 können Sie die Datenqualitätsumwandlungen in Streaming-Zuordnungen verwenden.

Sie können die folgenden Datenqualitätsumwandlungen in Streaming-Zuordnungen verwenden, um den Datenqualitätsvorgang auf die Streaming-Daten anzuwenden:

- Adress-Validiererumwandlung
- Klassifiziererumwandlung
- Parser-Umwandlung
- Standardisierer-Umwandlung

Weitere Informationen finden Sie im *Data Engineering Streaming 10.4.0-Benutzerhandbuch*.

### Ephemeral-Cluster in Streaming-Zuordnungen

Ab Version 10.4.0 können Sie einen Arbeitsablauf ausführen, um einen Ephemeral-Cluster zu erstellen, der Zuordnungen und andere Aufgaben in einem Cloudplattform-Cluster ausführt.

Um den Datenprozess an der Stelle wieder aufzunehmen, an der ein Cluster gelöscht wurde, können Sie Streaming-Zuordnungen auf dem Ephemeral-Cluster ausführen, indem Sie einen externen Speicher und ein Prüfpunkt-Verzeichnis angeben.

Weitere Informationen finden Sie im *Data Engineering Streaming 10.4.0-Benutzerhandbuch*.

## FileName-Port in Amazon S3

Ab Version 10.4.0 wird beim Erstellen eines Datenobjekt-Schreibvorgangs für Amazon S3-Dateien der FileName-Port automatisch angezeigt.

Zur Laufzeit erstellt der Datenintegrationsdienst getrennte Verzeichnisse für jeden Wert im FileName-Port und fügt die Zieldateien innerhalb der Verzeichnisse hinzu.

Weitere Informationen finden Sie im *Data Engineering Streaming 10.4.0-Benutzerhandbuch*.

## Microsoft Azure Data Lake Storage Gen2

Ab Version 10.4.0 können Sie Microsoft Azure Data Lake Storage Gen2 als Ziel in Streaming-Zuordnungen verwenden.

Azure Data Lake Storage Gen2 basiert auf Azure Blob Storage. Azure Data Lake Storage Gen2 hat sowohl die Funktionen von Azure Data Lake Storage Gen1 als auch von Azure Blob Storage. Sie können Azure Databricks Version 5.4 oder Azure HDInsight Version 4.0 verwenden, um auf die in Azure Data Lake Storage Gen2 gespeicherten Daten zuzugreifen.

Weitere Informationen finden Sie im *Data Engineering Streaming 10.4.0-Benutzerhandbuch*.

## Streaming-Zuordnungen in Azure Databricks

Ab Version 10.4.0 können Sie Streaming-Zuordnungen im Azure Databricks-Dienst in Microsoft Azure Cloud Services ausführen.

### Quellen und Ziele

Sie können Streaming-Zuordnungen mit den folgenden Quellen und Zielen innerhalb der Databricks-Umgebung ausführen:

- Microsoft Azure Event Hubs

- Azure Data Lake Storage Gen2 (ADLS Gen2)

### Umwandlungen

Sie können einer Databricks-Streaming-Zuordnung die folgenden Umwandlungen hinzufügen:

- Aggregator

- Ausdruck

- Filter

- Joiner

- Normalisierer

- Rang

- Router

- Union

- Fenster

### Datentypen

Folgende Datentypen werden unterstützt:

- Array

- Bigint

Datum/Uhrzeit

Dezimal

Doppelt

Ganzzahl

Map

Struct

Text

String

#### **Arbeitsabläufe**

Sie können Cluster-Arbeitsabläufe entwickeln, um in der Databricks-Umgebung Ephemeral-Cluster zu erstellen. Verwenden Sie Azure Data Lake Storage Gen1 (ADLS Gen1) und Azure Data Lake Storage Gen2 (ADLS Gen2) zum Erstellen von Ephemeral-Clustern in der Databricks-Umgebung.

Weitere Informationen zu Streaming-Zuordnungen in Azure Databricks finden Sie im *Data Engineering Streaming 10.4.0-Benutzerhandbuch*.

## Dynamische Zuordnungen in Data Engineering Streaming

Ab Version 10.4.0 ist Unterstützung für dynamische Zuordnungen in Data Engineering Streaming für die technische Vorschau verfügbar.

Sie können Confluent Kafka-Datenobjekte als dynamische Quellen und Ziele in einer Streaming-Zuordnung verwenden.

Funktionen für die technische Vorschau werden zu Bewertungszwecken unterstützt, sind aber nicht garantiert und nicht produktionsreif. Informatica empfiehlt, diese Funktionen nur außerhalb von Produktionsumgebungen einzusetzen. Informatica beabsichtigt, die Preview-Funktionen in kommenden Versionen für die Produktion bereitzustellen. Ob dies aber tatsächlich geschieht, hängt von verschiedenen marktbezogenen oder technischen Faktoren ab. Weitere Informationen erhalten Sie beim globalen Kundensupport von Informatica.

## Enterprise Data Catalog

In diesem Abschnitt werden die neuen Funktionen von Enterprise Data Catalog in Version 10.4.0 beschrieben.

### Zuweisen von benutzerdefinierten Attributen zu Ressourcen und Klassen

Ab Version 10.4.0 können Sie benutzerdefinierte Attribute zu bestimmten Ressourcen und Klassen im Katalog oder zu allen Ressourcen und Klassen im Katalog hinzufügen. Sie können diese Aufgabe ausführen, wenn Sie ein benutzerdefiniertes Attribut erstellen oder ändern.

Weitere Informationen finden Sie im *Informatica 10.4.0 Catalog Administrator-Handbuch*.

## Neue Ressourcen

Ab Version 10.4.0 stehen in Enterprise Data Catalog die folgenden neuen Ressourcen zur Verfügung:

- AWS Glue
- Microsoft Power BI
- Apache Cassandra
- Snowflake
- Google Cloud Storage

Aus allen obigen Ressourcen können Sie Metadaten-, Beziehungs- und Herkunftsinformationen extrahieren. Weitere Informationen finden Sie im *Konfigurationshandbuch zu Informatica 10.4.0 Enterprise Data Catalog Scanner*.

## Referenzressourcen und Referenzobjekte

Ab Version 10.4.0 können Sie Ressourcen konfigurieren, um Metadaten über Datenquellen oder andere Ressourcen im Katalog zu extrahieren, der von der Ressource referenziert wird. Beispiele sind Quell- und Zieltabellen in PowerCenter-Zuordnungen oder Quelltabellen und -Dateien aus Tableau-Berichten. Diese referenzierten Datenquellen werden als Referenzressourcen bezeichnet, und die in den Datenquellen enthaltenen Objekte werden als Referenzobjekte bezeichnet. Sie können die Referenzressourcen und Referenzobjekte in Enterprise Data Catalog-Anwendungen anzeigen. Um die vollständige Herkunft in Enterprise Data Catalog anzuzeigen, können Sie Verbindungszuordnungen für die Referenzressourcen durchführen. Sie können Verbindungszuordnungen zwischen Referenzressourcen und den Ressourcen, die Sie in Enterprise Data Catalog konfigurieren, für die Datenquellen der Referenzressourcen durchführen.

Sie können die folgenden Ressourcen konfigurieren, um Metadaten über Datenquellen oder andere Ressourcen im Katalog zu extrahieren, der von der Ressource referenziert wird:

- PowerCenter
- AWS Glue
- Tableau Server
- Coudera Navigator
- Apache Atlas
- Informatica Intelligent Cloud Services
- Informatica Platform
- SQL Server Integration Service

Weitere Informationen finden Sie im *Informatica 10.4.0 Catalog-Administratorhandbuch* und im *Informatica 10.4.0 Enterprise Data Catalog-Benutzerhandbuch*.

## Objekte aus Suchergebnissen exportieren

Ab Version 10.4.0 können Sie alle Objekte, die in den Suchergebnissen angezeigt werden, in die CSV-Datei (Datei mit kommasetrennten Werten) exportieren und die CSV-Datei zurück in Enterprise Data Catalog importieren. Exportieren Sie Objekte, nachdem Sie die Suchergebnisse auf einer beliebigen Suchregisterkarte optimiert haben.

Weitere Informationen finden Sie im Kapitel *Objektaufgaben* im *Informatica 10.4.0 Enterprise Data Catalog-Benutzerhandbuch*.

## Herkunfts- und Auswirkungsfiler

Ab Version 10.4.0 können Sie Filter für die Ansicht „Herkunft und Auswirkung“ erstellen. Verwenden Sie Filter, um die Ergebnisse in der Ansicht „Herkunft und Auswirkung“ einzuschränken. Erstellen Sie benutzerdefinierte Filter mit verschiedenen Kombinationen von Filteroptionen, gestützt auf die verfügbaren Objekttypen in der Ansicht „Herkunft und Auswirkung“.

Weitere Informationen finden Sie im Kapitel *Anzeigen von Herkunft und Auswirkung* im *Informatica 10.4.0 Enterprise Data Catalog-Benutzerhandbuch*.

## Übersicht über Objektsteuerung

Ab Version 10.4.0 können Sie Ablaufsteuerungsobjekte auf der Registerkarte **Übersicht über Objektsteuerung** für Tabellen- und Dateiobjekttypen anzeigen. Eine Ablaufsteuerung wendet Einschränkungen oder Bedingungen auf den Datenablauf an. Beispielsweise kann eine SQL-Abfrage Einschränkungen in der WHERE-Klausel enthalten, oder Zuordnungen können Umwandlungen enthalten.

Weitere Informationen finden Sie im Kapitel *Anzeigen von Herkunft und Auswirkung* im *Informatica 10.4.0 Enterprise Data Catalog-Benutzerhandbuch*.

## Regeln und Scorecards

Ab Version 10.4.0 können Sie Regeln und Scorecard-Ergebnisse aus dem Profiling Warehouse extrahieren und sie im Enterprise Data Catalog anzeigen. Sie können die Spaltenergebnisse für eine Regel und Scorecard anzeigen, um die Datenqualität der Datenquelle zu messen.

Weitere Informationen finden Sie im Kapitel *Anzeigen von Objekten* im *Informatica 10.4.0 Enterprise Data Catalog-Benutzerhandbuch*.

## Eindeutige Schlüsselableitung

Ab Version 10.4.0 können Sie Kandidaten für eindeutige Schlüssel aus den Spalten im Datenquellenobjekt generieren. Sie können die Ableitungsergebnisse für eindeutige Schlüssel anzeigen, um Datenqualitätsprobleme zu verstehen, indem Sie die Spalten identifizieren, die doppelte Werte aufweisen.

Sie können die Ableitungsergebnisse für eindeutige Schlüssel akzeptieren oder ablehnen. Nachdem Sie die Ableitung für eindeutige Schlüssel akzeptiert oder abgelehnt haben, können Sie die Ableitung des eindeutigen Schlüssels zurücksetzen, um den abgeleiteten Status wiederherzustellen.

Weitere Informationen finden Sie im Kapitel *Anzeigen von Objekten* im *Informatica 10.4.0 Enterprise Data Catalog-Benutzerhandbuch*.

## Datendomänenerkennung für den CLOB-Dateityp

Ab Version 10.4.0 können Sie die Datendomänenerkennung für den CLOB-Dateityp durchführen.

Weitere Informationen finden Sie im Kapitel *Enterprise Data Catalog-Konzepte* im *Informatica 10.4.0 Enterprise Data Catalog-Administratorhandbuch*.



# Data Discovery und Stichprobenoptionen auf der Spark-Engine

Ab Version 10.4.0 können Sie Profile zum Erkennen von Datendomänen ausführen und Stichprobenoptionen für die Spark-Engine wählen.

## Datendomänenerkennung auf der Spark-Engine

Sie können die Datendomänenerkennung auf der Spark-Engine durchführen.

## Stichprobenoptionen auf der Spark-Engine

Sie können die folgenden Stichprobenoptionen wählen, um Datendomänen auf der Spark-Engine zu erkennen:

- Mit der Stichprobenoption **Limit n** wird ein Profil basierend auf der Anzahl der Zeilen im Datenobjekt ausgeführt. Wenn Sie Datendomänen in der Hadoop-Umgebung erkennen möchten, sammelt die Spark-Engine Stichproben aus mehreren Partitionen des Datenobjekts und überträgt die Stichproben an einen einzelnen Knoten, um die Stichprobengröße zu berechnen.
- Mit der Stichprobenoption **Random percentage** wird ein Profil basierend auf dem Prozentsatz der Zeilen im Datenobjekt ausgeführt.

Weitere Informationen finden Sie im Kapitel *Enterprise Data Catalog-Konzepte* im *Informatica 10.4.0 Enterprise Data Catalog-Administratorhandbuch*.

# Nachverfolgen der technischen Vorschau

Enterprise Data Catalog Version 10.4.0 enthält Funktionen, die für die technische Vorschau zur Verfügung stehen.

Funktionen für die technische Vorschau sind noch nicht produktionsreif und werden daher zwar unterstützt, aber nicht garantiert. Informatica empfiehlt, diese Funktionen nur außerhalb von Produktionsumgebungen einzusetzen. Informatica beabsichtigt, die Vorschaufunktionen in kommenden GA-Versionen für die Produktion bereitzustellen. Ob dies aber tatsächlich geschieht, hängt von verschiedenen marktbezogenen oder technischen Faktoren ab. Weitere Informationen erhalten Sie beim globalen Kundensupport von Informatica.

- Ab Version 10.4.0 können Sie die kompakte Ansicht „Herkunft und Auswirkung“ anzeigen. Die kompakte Ansicht „Herkunft und Auswirkung“ zeigt das Herkunfts- und Auswirkungsdiagramm zusammengefasst auf Ressourcenebene an.

Weitere Informationen finden Sie im Kapitel *Anzeigen von Herkunft und Auswirkung* im *Informatica 10.4.0 Enterprise Data Catalog-Benutzerhandbuch*.

- Ab Version 10.4.0 können Sie Metadaten aus SAP Business Warehouse-, SAP BW/4HANA-, IBM InfoSphere DataStage- und Oracle Data Integrator-Quellen extrahieren, wenn diese zur Laufzeit nicht erreichbar oder offline sind.

Weitere Informationen finden Sie im *Informatica 10.4.0 Catalog Administrator-Handbuch*.

- Ab Version 10.4.0 können Sie Metadaten aus den SAP Business Warehouse- und SAP BW/4HANA-Datenquellen extrahieren.

Weitere Informationen finden Sie im *Konfigurationshandbuch zu Informatica 10.4.0 Enterprise Data Catalog Scanner*.

# Datenvorschau und Bereitstellung

Ab Version 10.4.0 können Sie die Datenbereitstellung durchführen, nachdem Sie die Datenerkennung im Katalog abgeschlossen haben. Die Datenbereitstellung hilft Ihnen dabei, Daten zur weiteren Analyse in ein

Ziel zu verschieben. Sie können eine Datenvorschau für relationale Quellen anzeigen, um die Daten zu bewerten, bevor Sie sie an das Ziel verschieben.

Weitere Informationen zur Datenvorschau und zur Bereitstellung von Daten finden Sie im *Informatica 10.4.0 Catalog-Administratorhandbuch* und im *Informatica 10.4.0 Enterprise Data Catalog-Benutzerhandbuch*.

## Unterstützte Ressourcentypen für das Standalone Scanner Utility

Ab Version 10.4 können Sie Metadaten aus den folgenden externen Quellen extrahieren, wenn zur Laufzeit nicht auf sie zugegriffen werden kann oder sie offline sind:

- Amazon Redshift
- Amazon S3
- Apache Cassandra
- Axon
- Azure Data Lake Store
- Azure Microsoft SQL Data Warehouse
- Azure Microsoft SQL Server
- Business Glossary
- Benutzerdefinierte Herkunft
- Datenbankskripts
- Erwin
- Glue
- Google BigQuery
- Google Cloud Storage
- Informatica Cloud Service
- IBM Cognos
- Microsoft Azure Blob Storage
- Microsoft SQL Server Integration Services
- PowerBI
- QlikView Business Intelligence
- Salesforce
- SAP HANA
- SAP PowerDesigner
- Snowflake
- Workday

Weitere Informationen finden Sie im Kapitel „Metadaten-Extraktion aus Offline- und nicht zugreifbaren Ressourcen“ im *Administratorhandbuch zu Informatica 10.4 Enterprise Data Catalog*.

## REST-APIs

Ab Version 10.4 können Sie die folgenden Informatica Enterprise Data Catalog REST-APIs verwenden:

- Datenbereitstellungs-REST-APIs. Zusätzlich zu den vorhandenen REST-APIs können Sie anzeigen, ob die Datenbereitstellung für den Benutzer verfügbar ist, und die Ressourcen auflisten, die die Datenbereitstellung unterstützen.
- Herkunftsfiler-REST-APIs. Sie können einen Herkunftsfiler erstellen, aktualisieren, auflisten oder löschen.
- Modellinformations-REST-APIs. Zusätzlich zu den vorhandenen REST-APIs können Sie die vordefinierten Schieberegler-Facets, Schieberegler-Facet-Definitionen und Herkunftsfilerdefinitionen auflisten.
- REST-API für Modelländerungen. Zusätzlich zu den vorhandenen REST-APIs können Sie eine Schieberegler-Facet-Definition erstellen, aktualisieren und löschen.
- Überwachungsinformations-REST-APIs. Sie können Jobs senden oder auflisten; dazu zählen Jobs vom Objektexporttyp, Objektimporttyp, Ressourcenexporttyp und Suchexporttyp.
- REST-API zur Zählung von untergeordneten Objekten. Sie können die Gesamtzahl der untergeordneten Elemente für ein Objekt auflisten.
- Produktinformations-REST-APIs. Sie können die Details zu Enterprise Data Catalog auflisten, darunter die Release-Version, die Build-Version und das Build-Datum.

Weitere Informationen zu den REST-APIs finden Sie im *Informatica 10.4 Enterprise Data Catalog REST-API-Benutzerhandbuch*.

## Enterprise Data Preparation

In diesem Abschnitt werden neue Funktionen für Enterprise Data Preparation in Version 10.4.0 erläutert.

### Data Lake-Zugriffsverwaltung

Ab Version 10.4.0 können Sie den Zugriff auf den Data Lake verwalten, indem Sie Berechtigungen für die Enterprise Data Catalog-Ressourcen, die Hive-Schemas und HDFS-Speicherorte im Data Lake repräsentieren, den Enterprise Data Preparation-Benutzern und -Benutzergruppen zuweisen.

Wenn Sie einem Benutzer oder einer Benutzergruppe Berechtigungen für bestimmte Schemas oder Speicherorte gewähren, zeigt die Anwendung nur die Schemas und Speicherorte an, für die der Benutzer über Berechtigungen verfügt, wenn er einen Import- Veröffentlichungs- oder Upload-Vorgang durchführt.

Weitere Informationen finden Sie im *Enterprise Data Preparation 10.4.0-Administratorhandbuch*.

### Microsoft Azure Data Lake Storage als Datenquelle

Ab Version 10.4.0 können Sie einen Microsoft Azure Data Lake Storage Gen2-Speicherort als Datenquelle verwenden.

### Dateien im Data Lake veröffentlichen

Ab Version 10.4.0 können Sie vorbereitete Daten in einer Datei an einem HDFS-Speicherort im Data Lake veröffentlichen.

Wenn Sie Daten veröffentlichen, können Sie den Datentyp auswählen, in den Sie die Daten im Data Lake schreiben möchten. Wenn Sie z. B. die Daten als Datei mit kommagetrennten Werten veröffentlichen möchten, schreibt die Anwendung die Daten als CSV-Datei in den Data Lake.

Weitere Informationen finden Sie im *Enterprise Data Preparation 10.4.0-Benutzerhandbuch*.

## Dateien in den Data Lake hochladen

Ab Version 10.4.0 können Sie mehrere Dateiformate von Ihrem lokalen Laufwerk in den Data Lake hochladen.

Sie können aus den folgenden Upload-Optionen wählen:

### **Laden Sie eine Datei direkt in den Data Lake hoch.**

Sie können eine kommagetrennte Datei, eine Avro-Datei, eine JSON-Datei oder eine Parquet-Datei im UTF-8-Format direkt von Ihrem lokalen Laufwerk in den Data Lake hochladen, ohne eine Vorschau der Daten anzuzeigen. Diese Option können Sie wählen, wenn Sie eine Datei ohne Datenvorschau hochladen möchten.

### **Lassen Sie die Dateistruktur von CLAIRE analysieren und laden Sie dann die Datei in den Data Lake hoch.**

Sie können die Daten in einer kommagetrennten Datei oder in einem Microsoft Excel-Arbeitsblatt in den Data Lake hochladen. Beim Hochladen der Datei verwendet Enterprise Data Preparation die eingebettete CLAIRE-Erkennungs-Engine zum Bestimmen der Dateistruktur und zum Anzeigen einer Datenvorschau.

Wenn Sie diese Option zum Hochladen einer Excel-Arbeitsmappe verwenden, erkennt die CLAIRE-Engine die Arbeitsblätter und Tabellen im Arbeitsblatt. Sie können das Arbeitsblatt bzw. die Tabelle auswählen, für das/die Sie eine Vorschau anzeigen möchten.

**Hinweis:** Das Hochladen eines Excel-Arbeitsblatts ist eine Vorschaufunktion.

### **Definieren Sie die Dateistruktur und laden Sie dann die Datei in den Data Lake hoch.**

Sie können die Daten in einer kommagetrennten Datei von Ihrem lokalen Laufwerk in den Data Lake hochladen. Beim Hochladen der Datei können Sie eine Vorschau der Daten anzeigen, die Dateistruktur angeben und die Spaltenattribute entsprechend Ihren Anforderungen konfigurieren. Wählen Sie diese Option, wenn Sie die Spaltenattribute vor dem Hochladen der Datei ändern müssen.

Weitere Informationen finden Sie im *Enterprise Data Preparation 10.4.0-Benutzerhandbuch*.

## Informatica-Zuordnungen

In diesem Abschnitt werden neue Funktionen für Informatica-Zuordnungen in Version 10.4.0 beschrieben.

## Binden von Zuordnungsausgaben an Zuordnungsparameter

Ab Version 10.4.0 können Sie Zuordnungsausgaben in Zuordnungen beibehalten und binden, die für die Ausführung in der nativen Umgebung oder auf der Spark-Engine bereitgestellt werden.

Erstellen Sie eine Zuordnungsausgabe. Binden Sie die Ausgabe an einen Zuordnungsparameter, um den Wert in nachfolgenden Ausführungen der Zuordnung zu verwenden. Wenn Sie die Zuordnung ausführen, übergibt der Datenintegrationsdienst den Wert der Zuordnungsausgabe an den Zuordnungsparameter. Um Zuordnungsausgaben beizubehalten, müssen Sie einen Laufzeit-Instanznamen angeben, indem Sie für den Befehl „`infacmd ms runMapping`“ die Option „`-RuntimeInstanceName`“ verwenden.

Das Developer Tool enthält jetzt eine Spalte **Bindung** in der Ansicht **Eigenschaften** der Zuordnung, um eine Zuordnungsausgabe an einen Parameter zu binden.

Weitere Informationen zu Zuordnungsausgaben in bereitgestellten Zuordnungen finden Sie im Kapitel „Zuordnungsausgaben“ im *Informatica 10.4.0 Developer-Zuordnungshandbuch*.

## infacmd ms-Befehle

In der folgenden Tabelle werden neue und aktualisierte infacmd ms-Befehle beschrieben:

| Befehl                        | Beschreibung   |
|-------------------------------|--|
| deleteMappingPersistedOutputs | Neuer Befehl, der alle persistenten Zuordnungsausgaben für eine bereitgestellte Zuordnung löscht. Geben Sie die zu löschenden Ausgaben anhand des Namens der Anwendung und des Namens der Laufzeitinstanz der Zuordnung an. Um bestimmte Ausgaben zu löschen, verwenden Sie die Option -OutputNamesToDelete. |
| getMappingStatus              | Aktualisierter Befehl, der jetzt den Jobnamen zurückgibt. Wenn Sie einen Laufzeit-Instanznamen in infacmd ms runMapping definiert haben, wird der Laufzeit-Instanzname als Jobname verwendet.  |
| listMappingPersistedOutputs   | Neuer Befehl, der die persistenten Zuordnungsausgaben für eine bereitgestellte Zuordnung auflistet. Die Ausgaben werden basierend auf dem Namen der Anwendung und dem Namen der Laufzeitinstanz der Zuordnung aufgelistet.   |

Weitere Informationen finden Sie im Kapitel „infacmd ms-Befehlsreferenz“ der *Informatica 10.4.0-Befehlsreferenz*.

## CLAIRE-Empfehlungen und -Einblicke

Ab Version 10.4.0 können Sie künstliche Intelligenz von CLAIRE nutzen, um Zuordnungen während der Entwicklung zu optimieren und zu korrigieren.

Wenn Sie Empfehlungen aktivieren, wird CLAIRE automatisch während der Entwicklung von Zuordnungen ausgeführt, und es werden Empfehlungen angezeigt, mit denen Sie die Zuordnungen korrigieren und optimieren können.

Sie können auch CLAIRE-Analysen für die Zuordnungen in einem Projekt oder Projektordner ausführen. Wenn Sie eine Gruppe von Zuordnungen analysieren, zeigt CLAIRE Einblicke zu Ähnlichkeiten zwischen Zuordnungen an.

Weitere Informationen zu Empfehlungen und Einblicken finden Sie unter *Data Engineering Integration User Guide*.

## Zuordnungsoptimierungsebene aktualisieren

Ab Version 10.4.0 können Sie mit dem Befehl „infacmd ms UpdateOptimizationDefaultLevel“ die Optimierungsebene der Zuordnungen in einer Anwendung auf die Standardebene (Automatisch) zurückzusetzen.

Wenn Sie den Befehl ausführen, müssen Sie einen Anwendungsnamen angeben.

„UpdateOptimizationDefaultLevel“ legt die Optimierungsebene für alle Zuordnungen in der Anwendung fest.

Weitere Informationen finden Sie in der *Informatica 10.4.0-Befehlsreferenz* und im *Informatica 10.4.0 Developer-Zuordnungshandbuch*.

# Informatica-Umwandlungen

In diesem Abschnitt werden neue Funktionen in Informatica-Umwandlungen in Version 10.4.0 erläutert.

## Adress-Validiererumwandlung

In diesem Abschnitt werden neue Funktionen für die Adress-Validiererumwandlung erläutert.

Die Adress-Validiererumwandlung enthält zusätzliche Adressfunktionen für die folgenden Länder:

### USA

Ab Version 10.4 erkennt der Adressvalidierer MC als alternative Version von MSC bzw. Mail Stop Code in einer US-Adresse.

Umfassende Informationen zu den Funktionen und Vorgängen der Engine für die Adressenüberprüfungssoftware in Version 10.4 finden Sie im *Entwicklerhandbuch für Informatica Address Verification 5.15.0*.

# Arbeitsabläufe in Informatica

In diesem Abschnitt werden neue Informatica Arbeitsablauf-Funktionen in Version 10.4.0 erläutert.

## Erweiterte Eigenschaften für die Cluster-Erstellungsaufgabe mit Amazon EMR

Ab Version 10.4.0 stehen neue erweiterte Eigenschaften zur Verfügung, wenn Sie eine Cluster-Erstellungsaufgabe mit einer Amazon EMR-Verbindung in einem Arbeitsablauf konfigurieren.

Sie können die folgenden Funktionen für eine Cluster-Erstellungsaufgabe konfigurieren:

- EBS-Volumengröße auf Root-Gerät. Die Anzahl GB des EBS-Volume-Root-Geräts.
- Benutzerdefinierte AMI-ID. ID eines benutzerdefinierten Amazon Linux Amazon Machine Image (AMI).
- Sicherheitskonfiguration. Der Name einer Sicherheitskonfiguration für die Authentifizierung und Verschlüsselung auf dem Cluster.

Weitere Informationen finden Sie im *Data Engineering Integration 10.4.0-Benutzerhandbuch* und im *Informatica® 10.4.0 Developer-Arbeitsablaufhandbuch*.

# Installation von Informatica

In diesem Abschnitt werden neue Installationsfunktionen in 10.4.0 erläutert.

## PostgreSQL

Ab Version 10.4.0 können Sie die PostgreSQL-Datenbank als Domänenkonfigurations-Repository, Modellrepository und PowerCenter-Repository verwenden. Für Enterprise Data Preparation können Sie die PostgreSQL-Datenbank nur für den zusätzlichen Modellrepository-Dienst verwenden.

Sie können auch die psql-Clientanwendungsversion 10.6 installieren, damit PostgreSQL unter Linux oder Windows funktioniert.

Weitere Informationen zu PostgreSQL finden Sie in den *Informatica 10.4.0-Installationshandbüchern*.

## Systemprüfungstool vor der Installation (i10Pi) im automatischen Modus

Ab Version 10.4.0 können Sie das Systemprüfungstool vor der Installation (i10pi) im automatischen Modus ausführen. Sie können i10pi im automatischen Modus ausführen, um zu überprüfen, ob der Rechner die Systemvoraussetzungen für eine Installation ohne Benutzereingriff erfüllt.

Weitere Informationen zum Ausführen von i10Pi im automatischen Modus finden Sie im *Informatica 10.4.0-Installationshandbuch*.

## Passwörter in der Eigenschaftsdatei für automatische Installation verschlüsseln

Ab Version 10.4.0 umfasst das Informatica-Installationsprogramm ein Hilfsprogramm, das Sie zum Verschlüsseln von Passwörtern verwenden können, die Sie in der Eigenschaftsdatei festlegen. Es enthält die erforderlichen Optionen zum Installieren der Dienste im automatischen Modus.

Wenn Sie das Installationsprogramm im automatischen Modus ausführen, entschlüsselt das Installations-Framework die verschlüsselten Passwörter.

Weitere Informationen finden Sie im *Informatica-Installations- und Konfigurationshandbuch*.

# Intelligentes Strukturmodell

In diesem Abschnitt werden neue Funktionen rund um intelligente Strukturmodelle in Version 10.4.0 beschrieben.

## Zusätzliche Eingabetypen

Ab Version 10.4.0 kann Intelligent Structure Discovery ORC-, Avro- und Parquet-Eingaben verarbeiten.

Weitere Informationen finden Sie im *Data Engineering Integration 10.4.0-Benutzerhandbuch*.

## Modell zur Entwurfszeit aus Beispiel erstellen

Ab Version 10.4.0 kann Intelligent Structure Discovery ein Modell gestützt auf eine Beispieldatei erstellen, die Sie beim Erstellen eines komplexen Datenobjekts auswählen. Sie können das Modell anschließend in Informatica Intelligent Cloud Services Data Integration optimieren.

Diese Funktion wird für XML-, JSON-, ORC-, AVRO- und Parquet-Beispieldateien unterstützt.

Weitere Informationen finden Sie im *Data Engineering Integration 10.4.0-Benutzerhandbuch*.

## Handhabung nicht identifizierter Daten

Ab Version 10.4.0 ordnet Intelligent Structure Discovery die Ausgabe nicht identifizierter Daten im Modell im strukturierten JSON-Format an.

Weitere Informationen finden Sie im *Data Engineering Integration 10.4.0-Benutzerhandbuch*.

## PowerCenter

In diesem Abschnitt werden neue PowerCenter-Funktionen in Version 10.4.0 erläutert.

### HTTP-Umwandlung

Ab Version 10.4.0 umfasst die HTTP-Umwandlung auch die folgenden Methoden für den endgültigen URL-Aufbau: SIMPLE PATCH, SIMPLE PUT und SIMPLE DELETE.

Mit der SIMPLE PATCH-Methode können Sie eine teilweise Aktualisierung durchführen, und die Eingabedaten müssen kein vollständiger Textkörper zu sein. Sie können damit Daten vom Eingabeport als Patch für die Ressource aktualisieren.

Mit der SIMPLE PUT-Methode können Sie ein Dokument vollständig ersetzen. Sie können Daten aus einem Eingabeport als einzelnen Datenblock für den HTTP-Server erstellen. Wenn die Daten bereits vorhanden sind, können Sie Daten von einem Eingabeport als einzelner Datenblock für den HTTP-Server aktualisieren.

Mit der SIMPLE DELETE-Methode können Sie Daten vom HTTP-Server löschen.

Sie können auch die Basis-URL für die HTTP-Umwandlung parametrisieren.

Zuvor konnten Sie den endgültigen URL-Aufbau nur mit den folgenden beiden Methoden angeben: SIMPLE GET und SIMPLE POST. Sie konnten auch nicht die endgültige URL für die HTTP-Umwandlung parametrisieren.

Weitere Informationen finden Sie im Kapitel „HTTP-Umwandlung“ im *Informatica 10.4.0-PowerCenter-Umwandlungshandbuch*.

## Konnektivität

In diesem Abschnitt werden neue Konnektivitätsfunktionen in Version 10.4.0 erläutert.

### Imitierte Benutzernamen für Oracle-Verbindungen

Ab Version 10.4.0 können Sie den imitierten Benutzernamen in der Oracle-Verbindung in PowerCenter angeben, um eine Verbindung zu Oracle-Quellen und -Zielen herzustellen. Wenn Sie einen imitierten Benutzer angeben, können Sie im Namen dieses Benutzers eine Verbindung zu Oracle herstellen und Lese- oder Schreibvorgänge ausführen.

Weitere Informationen finden Sie im *Informatica 10.4.0 PowerCenter-Grundlagen-Arbeitsablaufhandbuch*.

### Unterstützung für SAP HANA-Datenbank-Modellierungsansichten

Ab Version 10.4.0 können Sie Daten aus den folgenden Typen von SAP HANA-Datenbank-Modellierungsansichten lesen:

- Analyseansichten



- Attributansichten
- Berechnete Ansichten

Weitere Informationen finden Sie im *Informatica 10.4.0 PowerCenter Designer-Handbuch*.

## PowerExchange-Adapter

In diesem Abschnitt werden die neuen PowerExchange-Adapter-Funktionen in Version 10.4.0 erläutert.

### PowerExchange-Adapter für Informatica

In diesem Abschnitt werden neue Funktionen der Informatica-Adapter in Version 10.4.0 erläutert.

#### PowerExchange for Amazon Redshift

Ab Version 10.4.0 enthält PowerExchange for Amazon Redshift die folgenden Funktionen:

- Sie können Zuordnungen in der AWS Databricks-Umgebung ausführen.
- Sie können einen Clusterregionsnamen in der Verbindungseigenschaft für die Clusterregion auswählen, auch wenn Sie den Clusterregionsnamen in der JDBC-URL-Verbindungseigenschaft angeben.
- Sie können Nullwerte beibehalten, wenn Sie Daten aus Amazon Redshift lesen.
- Sie können die Anzahl der Staging-Dateien pro Batch angeben, wenn Sie Daten in Amazon Redshift schreiben.
- Sie können die Datensatzreihenfolge beibehalten, wenn Sie Daten aus einer CDC-Quelle in ein Amazon Redshift-Ziel schreiben.

Weitere Informationen finden Sie im *Informatica 10.4.0 PowerExchange for Amazon Redshift-Benutzerhandbuch*.

#### PowerExchange for Amazon S3

Ab Version 10.4.0 enthält PowerExchange for Amazon S3 die folgenden Funktionen:

- Sie können Zuordnungen in der AWS Databricks-Umgebung ausführen.
- Sie können die vorläufigen Sicherheitsanmeldedaten nutzen, indem Sie AssumeRole für den Zugriff auf die AWS-Ressourcen verwenden.
- Sie können den Datenformattyp und das Schema in den Eigenschaften des Lese- und Schreibvorgangs zur Laufzeit parametrisieren.

Weitere Informationen finden Sie im *Informatica 10.4.0 PowerExchange for Amazon S3-Benutzerhandbuch*.

#### PowerExchange for Google BigQuery

Ab Version 10.4.0 bietet PowerExchange for Google BigQuery die folgenden Funktionen:

- Sie verwenden einen Google Dataproc-Cluster, um Zuordnungen auf der Spark-Engine auszuführen.
- Sie können die Leistung der Zuordnung steigern, indem Sie die Zuordnung im optimierten Spark-Modus ausführen. Wenn Sie den optimierten Spark-Modus zum Lesen von Daten verwenden, können Sie die Anzahl der zu verwendenden Partitionen angeben. In den erweiterten Eigenschaften für den Lese- und

Schreibvorgang können Sie angeben, ob Sie die Zuordnung im Modus **Generisch** oder **Optimiert** ausführen möchten. Der optimierte Spark-Modus erhöht die Zuordnungsleistung.

- Sie können außerdem ein SQL-Override konfigurieren, um die zum Extrahieren von Daten aus der Google BigQuery-Quelle verwendete SQL-Standardabfrage zu überschreiben.
- Sie können Daten des NUMERIC-Datentyps in Google BigQuery lesen oder schreiben. Der NUMERIC-Datentyp ist ein exakter numerischer Wert mit einer Genauigkeit von 38 Stellen und einer Skalierung von 9 Dezimalstellen. Wenn Sie den NUMERIC-Datentyp lesen oder schreiben, wird er vom Datenintegrationsdienst dem Dezimal-Umwandlungsdantentyp zugewiesen, und die zulässige Genauigkeit ist maximal 38 und die Skalierung maximal 9.

Weitere Informationen finden Sie im *Informatica 10.4.0 PowerExchange for Google BigQuery-Benutzerhandbuch*.

## PowerExchange for Google Cloud Storage

Ab Version 10.4.0 bietet PowerExchange for Google Cloud Storage die folgenden Funktionen:

- Sie verwenden einen Google Dataproc-Cluster, um Zuordnungen auf der Spark-Engine auszuführen.
- Sie können die folgenden erweiterten Eigenschaften für Datenobjekt-Lesevorgänge in Google Cloud Storage konfigurieren, wenn Sie Daten aus einer Google Cloud Storage-Quelle lesen:

### Google Cloud Storage-Pfad

Überschreibt den Google Cloud Storage-Pfad zu der Datei, die Sie im Google Cloud Storage-Datenobjekt ausgewählt hatten.

Verwenden Sie das folgende Format:

```
gs://<bucketname> oder gs://<bucketname>/<ordnername>
```

### Name der Quelldatei

Überschreibt den Google Cloud Storage-Quelldateinamen, der im Google Cloud Storage-Datenobjekt angegeben ist.

### Ist Verzeichnis

Liest alle verfügbaren Dateien in dem Ordner, der in der erweiterten Eigenschaft des Datenobjekt-Lesevorgangs für **Google Cloud Storage-Pfad** angegeben ist.

Weitere Informationen finden Sie im *Informatica 10.4.0 PowerExchange for Google Cloud Storage-Benutzerhandbuch*.

## PowerExchange for Microsoft Azure Blob Storage

Ab Version 10.4.0 enthält PowerExchange for Microsoft Azure Blob Storage die folgenden Funktionen:

- Sie können den Datenformattyp und das Schema in den Eigenschaften des Lese- und Schreibvorgangs zur Laufzeit parametrisieren.
- Sie können die Authentifizierung über Signaturen für gemeinsamen Zugriff verwenden, während Sie eine Microsoft Azure Blob Storage-Verbindung herstellen.

Weitere Informationen finden Sie im *Informatica 10.4.0 PowerExchange for Microsoft Azure Blob Storage-Benutzerhandbuch*.

## PowerExchange for Microsoft Azure SQL Data Warehouse

Ab Version 10.4.0 können Sie Daten aus einem Microsoft Azure SQL Data Warehouse-Endpunkt lesen oder auf einen Microsoft Azure SQL Data Warehouse-Endpunkt schreiben, der sich in einem virtuellen Netzwerk (VNet) befindet.

Weitere Informationen finden Sie im *Informatica 10.4.0 PowerExchange for Microsoft Azure SQL Data Warehouse-Benutzerhandbuch*.

## PowerExchange for Salesforce

Ab Version 10.4.0 bietet PowerExchange for Salesforce die folgenden Funktionen:

- Sie können unter Verwendung von Version 45.0, 46.0 und 47.0 der Salesforce-API eine Salesforce-Verbindung aufbauen und auf Salesforce-Objekte zugreifen.
- Sie können die Primärschlüsselsegmentierung für Abfragen in einem freigegebenen Objekt aktivieren, das einen Freigabeeintrag im übergeordneten Objekt darstellt. Die Primärschlüsselsegmentierung wird für freigegebene Objekte nur dann unterstützt, wenn das übergeordnete Objekt unterstützt wird. Wenn Sie beispielsweise eine Abfrage für „CaseHistory“ durchführen möchten, muss die Primärschlüsselsegmentierung für das übergeordnete Objekt „Case“ unterstützt werden.
- Sie können Zuweisungsregeln zum erneuten Zuweisen von Attributen in Datensätzen verwenden, wenn Sie Datensätze für Lead- oder Case-Zielobjekte mithilfe der Standard-API einfügen, aktualisieren oder ein Upsert für sie durchführen.

Weitere Informationen finden Sie im *Informatica 10.4.0 PowerExchange for Salesforce-Benutzerhandbuch*.

## PowerExchange for SAP NetWeaver

Ab Version 10.4.0 enthält PowerExchange for SAP NetWeaver die folgenden Funktionen:

- Sie können HTTPS-Streaming für SAP Table Reader-Zuordnungen konfigurieren.
- Sie können Daten aus ABAP CDS-Ansichten mit SAP Table Reader lesen, wenn Sie die SAP NetWeaver-Systemversion 7.50 oder höher verwenden.
- Sie können Daten aus SAP-Tabellen mit Feldern lesen, die die folgenden Datentypen aufweisen:
  - DF16\_DEC
  - DF32\_DEC
  - DF16\_RAW
  - DF34\_RAW
  - INT8
  - RAWSTRING
  - SSTRING
  - STRING

Weitere Informationen finden Sie im *Informatica PowerExchange for SAP NetWeaver 10.4.0-Benutzerhandbuch*.

## PowerExchange for Snowflake

Ab Version 10.4.0 enthält PowerExchange for Snowflake die folgenden Funktionen:

- Sie können Snowflake-Zuordnungen in der Databricks-Umgebung ausführen.
- Sie können Snowflake-Objekte als dynamische Quellen und Ziele in einer Zuordnung verwenden.
- Sie können ein Snowflake-Ziel erstellen, indem Sie die Option **Ziel erstellen** verwenden.
- Sie können eine Zielschemastrategie für ein Snowflake-Ziel in einer Zuordnung konfigurieren. Sie können aus den verfügbaren Optionen auswählen, um entweder das vorhandene Zielschema beizubehalten oder

ein Ziel zu erstellen, falls keines vorhanden ist. Sie können auch Optionen der Zielschemastrategie als Parameterwert angeben.

- In den erweiterten Zieleigenschaften von Snowflake können Sie einen Dateinamen und -pfad für abgelehnte Datensätze angeben, den der Datenintegrationsdienst zum Schreiben von Datensätzen verwendet, die beim Schreiben ins Ziel abgelehnt wurden.
- Wenn der ODBC-Provider-Typ in der Snowflake-ODBC-Verbindung Snowflake ist, können Sie die Pushdown-Optimierung konfigurieren, um die Umwandlungslogik in die Snowflake-Datenbank zu übertragen.
- Sie können Daten vom Dezimal-Datentyp mit einer Genauigkeit und Skalierung von 38 Stellen lesen und schreiben, indem Sie das benutzerdefinierte Kennzeichen `EnableSDKDecimal38` in den Eigenschaften des Datenintegrationsdiensts konfigurieren.

Weitere Informationen finden Sie im *Informatica 10.4.0 PowerExchange for Snowflake-Benutzerhandbuch*.

## PowerExchange for HDFS

Ab Version 10.4.0 enthält PowerExchange for HDFS die folgenden Funktionen:

- Sie können den Datenformattyp und das Schema in den Eigenschaften des Lese- und Schreibvorgangs zur Laufzeit parametrisieren.
- Sie können das Schema eines komplexen Datei-Datenobjekts für einen Lese- oder Schreibvorgang formatieren.

Weitere Informationen finden Sie im *Informatica 10.4.0 PowerExchange for HDFS-Benutzerhandbuch*.

## PowerExchange-Adapter für PowerCenter

In diesem Abschnitt werden neue PowerCenter-Adapterfunktionen in Version 10.4.0 erläutert.

### PowerExchange for Google BigQuery

Ab Version 10.4.0 können Sie Daten des NUMERIC-Datentyps in Google BigQuery lesen oder schreiben. Der NUMERIC-Datentyp ist ein exakter numerischer Wert mit einer Genauigkeit von 38 Stellen und einer Skalierung von 9 Dezimalstellen. Wenn Sie den NUMERIC-Datentyp lesen oder schreiben, wird er vom PowerCenter-Integrationsdienst dem Dezimal-Umwandlungsdatentyp zugewiesen, und die zulässige Genauigkeit ist maximal 28 und die Skalierung maximal 9.

Weitere Informationen finden Sie im *Informatica 10.4.0 PowerExchange for Google BigQuery-Benutzerhandbuch für PowerCenter*.

### PowerExchange for Google Cloud Storage

Ab Version 10.4.0 können Sie die folgenden erweiterten Eigenschaften für Datenobjekt-Lesevorgänge in Google Cloud Storage konfigurieren, wenn Sie Daten aus einer Google Cloud Storage-Quelle lesen:

#### **Google Cloud Storage-Pfad**

Überschreibt den Google Cloud Storage-Pfad zu der Datei, die Sie im Google Cloud Storage-Datenobjekt ausgewählt hatten.

Verwenden Sie das folgende Format:

```
gs://<bucketname> oder gs://<bucketname>/<ordnername>
```

### Name der Quelldatei

Überschreibt den Google Cloud Storage-Quelldateinamen, der im Google Cloud Storage-Datenobjekt angegeben ist.

### Ist Verzeichnis

Liest alle verfügbaren Dateien in dem Ordner, der in der erweiterten Eigenschaft des Datenobjekt-Lesevorgangs für **Google Cloud Storage-Pfad** angegeben ist.

Weitere Informationen finden Sie im *Informatica 10.4.0 PowerExchange for Google Cloud Storage-Benutzerhandbuch für PowerCenter*.

## PowerExchange for Greenplum

Ab Version 10.4.0 können Sie PowerExchange for Greenplum zum Lesen von Daten aus Greenplum verwenden. Sie können spezifische Sitzungseigenschaften für Greenplum-Quellen konfigurieren, um zu bestimmen, wie Daten aus Greenplum extrahiert werden.

Wenn Sie eine Greenplum-Sitzung zum Lesen von Daten ausführen, ruft der PowerCenter-Integrationsdienst den parallelen Dateiserver der Greenplum-Datenbank, gpfdist (das Dateiverteilungsprogramm von Greenplum) auf, um Daten zu lesen.

Weitere Informationen finden Sie im *Informatica 10.4.0 PowerExchange for Greenplum-Benutzerhandbuch für PowerCenter*.

## PowerExchange for JD Edwards EnterpriseOne

Ab Version 10.4.0 können Sie mithilfe von Version 9.2 der JD Edwards EnterpriseOne-API eine JD Edwards EnterpriseOne-Verbindung herstellen und auf JD Edwards EnterpriseOne-Objekte zugreifen.

Weitere Informationen finden Sie im *Informatica 10.4.0 PowerExchange for JD Edwards EnterpriseOne-Benutzerhandbuch für PowerCenter*.

## PowerExchange for Kafka

Ab Version 10.4.0 können Sie die folgenden SSL-Eigenschaften konfigurieren, um eine sichere Verbindung zu einem Kafka-Broker zu ermöglichen:

- SSL-Modus
- SSL-TrustStore-Dateipfad
- SSL-TrustStore-Passwort
- SSL-KeyStore-Dateipfad
- SSL-KeyStore-Passwort

Sie können den Kafka-Nachrichten-Broker für die Verwendung der Kafka-Broker-Version 0.10.1.1 und höher konfigurieren.

Weitere Informationen finden Sie im *Informatica PowerExchange for Kafka 10.4.0-Benutzerhandbuch für PowerCenter*.

## PowerExchange for Salesforce

Ab Version 10.4.0 können Sie mithilfe von Version 46.0 und 47.0 der Salesforce-API eine Salesforce-Verbindung herstellen und auf Salesforce-Objekte zugreifen.

Weitere Informationen finden Sie im *Informatica 10.4.0 PowerExchange for Salesforce-Benutzerhandbuch für PowerCenter*.

## PowerExchange for SAP NetWeaver

Ab Version 10.4.0 können Sie PowerExchange for SAP Dynamic ABAP Table Extractor zum Lesen von Daten aus SAP-Tabellen und ABAP Core Data Services(CDS)-Ansichten über HTTP/HTTPS-Streaming verwenden. Sie können Daten aus ABAP CDS-Ansichten mit PowerExchange for SAP Dynamic ABAP Table Extractor lesen, wenn Sie die SAP NetWeaver-Systemversion 7.50 oder höher verwenden.

Weitere Informationen finden Sie im *Informatica PowerExchange for SAP NetWeaver 10.4.0-Benutzerhandbuch*.

## Sicherheit

In diesem Abschnitt werden neue Sicherheitsfunktionen in Version 10.4.0 erläutert.

### Webanwendungen für die Verwendung von verschiedenen SAML-Identitätsanbietern konfigurieren

Ab Version 10.4.0 können Sie Informatica-Webanwendungen, die in einer Domäne ausgeführt werden, für die Verwendung von verschiedenen SAML-Identitätsanbietern (Security Assertion Markup Language) konfigurieren. Beispielsweise können Sie Informatica Administrator für die Verwendung von Active Directory-Verbindungsstellen als Identitätsanbieter und Informatica Analyst für die Verwendung von PingFederate als Identitätsanbieter konfigurieren.

Weitere Informationen finden Sie im *Informatica 10.4.0-Sicherheitshandbuch*.

## KAPITEL 2

# Neue Funktionen (10.2.2 HotFix 1)

Dieses Kapitel umfasst die folgenden Themen:

- [Befehlszeilenprogramme , 39](#)
- [Enterprise Data Catalog, 39](#)

## Befehlszeilenprogramme

In diesem Abschnitt werden neue Befehle in Version 10.2.2 HotFix 1 erläutert.

### infacmd Idm-Befehle

In der folgenden Tabelle werden neue Optionen für den Befehl infacmd Idm beschrieben:

| Neue Option   | Beschreibung   |
|---------------|--|
| -Force<br>-fr | Optional. Wenn Sie ein Backup erzwingen möchten, wenn der Backup-Modus offline ist. Führt ein erzwungenes Backup durch und überschreibt das vorhandene Backup. |
| -Force<br>-fr | Optional. Wenn Sie den vorhandenen Inhalt von HDFS und Apache Zookeeper bereinigen möchten. Erzwingt das Wiederherstellen der Backup-Daten.                    |

Weitere Informationen finden Sie im Kapitel „infacmd Idm-Befehlsreferenz“ in der *Befehlsreferenz zu Informatica 10.2.2 HotFix1*.

## Enterprise Data Catalog

In diesem Abschnitt werden die neuen Funktionen von Enterprise Data Catalog in Version 10.2.2 HotFix 1 beschrieben.

## Azure Data Lake Storage Gen2

Ab Version 10.2.2 HotFix 1 können Sie Metadaten aus einem Azure Data Lake-Speicher Gen2 als Quelle extrahieren. Sie können auf die Metadaten von Dateien und Ordnern in einem Azure Data Lake-Speicher Gen2-Storage-Konto zugreifen und diese in Enterprise Data Catalog extrahieren.

Weitere Informationen finden Sie im Kapitel „Azure Data Lake Store“ im *Konfigurationshandbuch zu Informatica 10.2.2 HotFix1 Enterprise Data Catalog Scanner*.

## Verknüpfungen ohne Unterscheidung der Groß-/Kleinschreibung

Ab Version 10.2.2 HotFix 1 können Sie Ressourcen ohne Unterscheidung der Groß-/Kleinschreibung konfigurieren. Mit dieser Option können Sie unabhängig von der Groß-/Kleinschreibung der Ressourcennamen die verknüpfte Herkunft zwischen Quellen und Zielen anzeigen.

Weitere Informationen über Verknüpfungen ohne Unterscheidung der Groß-/Kleinschreibung finden Sie im Kapitel „Verwalten von Ressourcen“ im *Administratorhandbuch zu Informatica 10.2.2 HotFix 1 Catalog*.

## Enterprise Data Catalog Tableau Extension

Ab Version 10.2.2 HotFix 1 ist Enterprise Data Catalog Tableau Extension als native Erweiterung für das Tableau-Dashboard verfügbar, das Sie in Tableau verwenden können.

Sie können Enterprise Data Catalog Tableau Extension in Tableau Desktop, Tableau Server und allen Webbrowsern verwenden, die Tableau unterstützt. Laden Sie die Erweiterung von der Enterprise Data Catalog-Anwendung herunter und fügen Sie sie zu einem Dashboard in Tableau hinzu.

Weitere Informationen über die Erweiterung finden Sie im Handbuch zu *Informatica 10.2.2 HotFix1 Enterprise Data Catalog Extension for Tableau*.

## Neue Ressourcen

Ab Version 10.2.2 HotFix 1 stehen in Enterprise Data Catalog die folgenden neuen Ressourcen zur Verfügung:

- SAP PowerDesigner Sie können Metadaten, Beziehungen und Herkunftsinformationen aus einer SAP PowerDesigner-Datenquelle extrahieren.
- SAP HANA Sie können Objekt- und Herkunftsmetadaten von einer SAP HANA-Datenbank extrahieren.

Weitere Informationen finden Sie im *Konfigurationshandbuch zu Informatica 10.2.2 HotFix1 Scanner*.

## Profiling-Warehouse-Scanner

Ab Version 10.2.2 HotFix 1 können Sie eine oder mehrere Profiling-Warehouse-Ressourcen erstellen. Sie können ein Profiling-Warehouse verknüpfen und Schemas in der Ressource auswählen. Wenn Sie die Ressource ausführen, extrahiert und migriert der Scanner die Profilergebnisse für die ausgewählten Schemas aus dem zugeordneten Profiling-Warehouse in den Katalog.

Weitere Informationen finden Sie im Kapitel „Konfigurieren von Informatica Platform-Scannern“ im *Konfigurationshandbuch zu Informatica 10.2.2 HotFix1 Enterprise Data Catalog Scanner*.



## REST-APIs

Ab Version 10.2.2 HotFix 1 können Sie die folgenden Informatica Enterprise Data Catalog REST-APIs verwenden:

- Datenbereitstellungs-REST-APIs. Sie können Verbindungen und Ressourcen zurückgeben, aktualisieren oder löschen.
- Katalogmodell-REST-APIs. Zusätzlich zu den vorhandenen REST-APIs können Sie auf die Feld-Facets, Abfrage-Facets und Suchregisterkarten zugreifen und diese aktualisieren oder löschen.
- Objekt-APIs. Zusätzlich zu den vorhandenen REST-APIs können Sie die Katalogsuche und -vorschläge auflisten.

Weitere Informationen zu den REST-APIs finden Sie im *Referenzhandbuch zu Informatica 10.2.2 HotFix 1 Enterprise Data Catalog REST-API*.

## Suche Verbesserungen

Ab Version 10.2.2 HotFix 1 können Sie in Enterprise Data Catalog aufgrund der folgenden Verbesserungen bessere Suchvorgänge durchführen:

### Suche mit doppelten Anführungszeichen

Sie können eine Suche nach Objekten mithilfe von doppelten Anführungszeichen durchführen (""), um Objekte zu finden, die exakt dem Objektnamen innerhalb der doppelten Anführungszeichen entsprechen, aber nicht den Variationen des Objektnamens im Katalog.

### Suchoperatoren

Sie können neuere Suchoperatoren verwenden, um die Suchergebnisse zu verfeinern. Die Suchoperatoren sind AND, OR, NOT, title und description.

### Suchrang

Enterprise Data Catalog verwendet einen Ranking-Algorithmus, um Datenobjekte auf der Suchergebnisseite einzustufen. Das Such-Ranking bezieht sich auf die Priorität eines Objekts im Vergleich zu anderen Objekten, die Teil bestimmter Suchergebnisse sind.

### Verwandte Suche

Sie können die Option **Verwandte Suche anzeigen** auf der Seite **Suchergebnisse** aktivieren, um verwandte Objekte anzuzeigen.

Weitere Informationen über Verbesserungen bei der Suche finden Sie im Kapitel „Suchen nach Objekten“ im *Benutzerhandbuch zu Informatica 10.2.2 HotFix 1 Enterprise Data Catalog*.

## Suche Registerkarten

Ab Version 10.2.2 HotFix 1, können Sie mit den Suchregisterkarten nach Objekten suchen, ohne dabei immer wieder die gleichen Suchkriterien festlegen zu müssen. Die Suchregisterkarten sind vordefinierte Filter im Katalog.

Weitere Informationen über Suchregisterkarten finden Sie im Kapitel „Anpassen der Suche“ im *Benutzerhandbuch zu Informatica 10.2.2 HotFix 1 Enterprise Data Catalog*.

## Unterstützte Ressourcentypen für das Standalone Scanner Utility

Ab Version 10.2.2 HotFix 1 können Sie Metadaten aus den folgenden externen Quellen extrahieren, wenn zur Laufzeit nicht auf sie zugegriffen werden kann oder sie offline sind:

- Apache Atlas
- Cloudera Navigator
- Dateisystem
- HDFS
- Hive
- Informatica Platform
- MicroStrategy
- OneDrive
- Oracle Business Intelligence
- SharePoint
- Sybase
- Tableau

Weitere Informationen finden Sie im Kapitel „Metadaten-Extraktion aus Offline- und nicht zugreifbaren Ressourcen“ im *Administratorhandbuch zu Informatica 10.2.2 HotFix 1 Enterprise Data Catalog*.

## Technical Preview

Enterprise Data Catalog Version 10.2.2 HotFix 1 enthält Funktionen, die für die technische Vorschau zur Verfügung stehen.

Funktionen für die technische Vorschau sind noch nicht produktionsreif und werden daher zwar unterstützt, aber nicht garantiert. Informatica empfiehlt, diese Funktionen nur außerhalb von Produktionsumgebungen einzusetzen. Informatica beabsichtigt, die Preview-Funktionen in kommenden GA-Versionen für die Produktion bereitzustellen. Ob dies aber tatsächlich geschieht, hängt von verschiedenen marktbezogenen oder technischen Faktoren ab. Weitere Informationen erhalten Sie beim globalen Kundensupport von Informatica.

Ab Version 10.2.2 HotFix 1 stehen die folgenden Funktionen in der technischen Vorschau zur Verfügung:

- Ab Version 10.2.2 HotFix 1 können Sie Metadaten für die Datenherkunft auf Spaltenebene extrahieren, einschließlich der Umwandlungslogik von einer Oracle Data Integrator-Datenquelle.
- Ab Version 10.2.2 HotFix 1 können Sie Metadaten für die Datenherkunft auf Spaltenebene extrahieren, einschließlich der Umwandlungslogik von einer IBM InfoSphere DataStage-Datenquelle.
- Ab Version 10.2.2 HotFix 1 können Sie die Datenherkunft auf Spaltenebene für gespeicherte Prozeduren in Oracle und SQL Server extrahieren.
- Ab Version 10.2.2 HotFix 1 können Sie die Datenbereitstellung durchführen, nachdem Sie die Datenerkennung im Katalog abgeschlossen haben. Die Datenbereitstellung hilft Ihnen dabei, Daten zur weiteren Analyse in ein Ziel zu verschieben.

Weitere Informationen zur Datenvorschau finden Sie im *Administratorhandbuch zu Informatica 10.2.2 HotFix 1 Catalog* und im *Benutzerhandbuch zu Informatica 10.2.2 Hotfix 1 Enterprise Data Catalog*.

- Ab Version 10.2.2 HotFix 1 haben Sie die Möglichkeit der Datenvorschau, um die Daten zu bewerten, bevor Sie sie in das Ziel verschieben. Die Datenvorschau ist nur für tabellarische Objekte in Oracle- und Microsoft SQL Server-Ressourcen verfügbar.

Weitere Informationen zur Datenvorschau finden Sie im *Administratorhandbuch zu Informatica 10.2.2 HotFix 1 Catalog* und im *Benutzerhandbuch zu Informatica 10.2.2 Hotfix 1 Enterprise Data Catalog*.

## KAPITEL 3

# Neue Funktionen (10.2.2 Service Pack 1)

Dieses Kapitel umfasst die folgenden Themen:

- [Big Data Management, 43](#)
- [Big Data Streaming, 44](#)
- [Enterprise Data Catalog, 45](#)
- [Enterprise Data Preparation, 45](#)
- [PowerExchange-Adapter für Informatica, 46](#)

## Big Data Management

In diesem Abschnitt werden neue Funktionen für Big Data Management in Version 10.2.2 Service Pack 1 erläutert.

### Sqoop

Ab Version 10.2.2 Service Pack 1 können Sie die folgenden neuen Sqoop-Funktionen nutzen.

#### Verbindung zu SSL-fähigen Oracle- und Microsoft SQL Server-Datenbanken

Sie können Sqoop so konfigurieren, dass eine Verbindung zu sicheren relationalen Datenbanken wie Oracle und Microsoft SQL Server hergestellt wird.

Verwenden Sie die entsprechende JDBC-Verbindungszeichenfolge und das Connect-Argument in der JDBC-Verbindung, um eine Verbindung zu einer SSL-fähigen Oracle- oder Microsoft SQL Server-Datenbank herzustellen.

Weitere Informationen finden Sie im *Informatica Big Data Management 10.2.2 Service Pack 1-Benutzerhandbuch*.

#### Unterstützung für das Sqoop-Argument „connection-param-file“

Sie können das Argument `connection-param-file` definieren, um eine Verbindung zu einer SSL-fähigen Oracle-Datenbank herzustellen. `connection-param-file` ist eine Eigenschaftsdatei, die die zusätzlichen JDBC-Parameter spezifiziert, die Sqoop verwenden muss, um eine Verbindung mit der SSL-fähigen Datenbank herzustellen.

Der Inhalt dieser Datei wird als Java-Standardeigenschaften analysiert und beim Erstellen einer Verbindung an den Treiber übergeben.

Sie können das Argument `connection-param-file` im Feld **Sqoop-Argumente** in der JDBC-Verbindung angeben.

Verwenden Sie folgende Syntax:

```
--connection-param-file <Name der Parameterdatei>
```

Weitere Informationen finden Sie im *Informatica Big Data Management 10.2.2 Service Pack 1-Benutzerhandbuch*.

## Big Data Streaming

In diesem Abschnitt werden die neuen Funktionen für Big Data Streaming in Version 10.2.2 Service Pack 1 erläutert.

### Amazon S3-Ziel

Ab Version 10.2.2 Service Pack 1 können Sie ein Streaming-Mapping erstellen, um Daten auf Amazon S3 zu schreiben.

Erstellen Sie ein Amazon S3-Datenobjekt, um Daten auf Amazon S3 zu schreiben. Sie können eine Amazon S3-Verbindung erstellen, um Amazon S3 als Ziel zu verwenden. Sie können eine Amazon S3-Verbindung im Developer Tool oder durch `infacmd` erstellen und verwalten.

Weitere Informationen finden Sie im *Informatica Big Data Streaming 10.2.2 Service Pack 1-Benutzerhandbuch*.

### TIME\_RANGE-Funktion

Ab Version 10.2.2 Service Pack 1 können Sie die `TIME_RANGE`-Funktion in einer Joiner-Umwandlung verwenden, die den Zeitbereich für die zu verbindenden Streaming-Ereignisse festlegt.

Die Funktion `TIME_RANGE` gilt nur für eine Joiner-Umwandlung in einem Streaming-Mapping.

#### Syntax

```
TIME_RANGE(EventTime1,EventTime2,Format,Interval)
```

Weitere Informationen zur `TIME_RANGE`-Funktion finden Sie im *Informatica 10.2.2 Service Pack 1 Transformation Language-Referenzhandbuch*.

### Schreiben von Daten in mehrere HDFS-Dateien

Ab Version 10.2.2 Service Pack 1 können Sie mit einem komplexen Dateidatenobjekt Quelldaten in mehrere HDFS-Dateien schreiben. Auf Basis der `FileName-Header-Portwerte` des komplexen Dateidatenobjekts erstellt der Datenintegrationsdienst zur Laufzeit mehrere HDFS-Dateien am Zielspeicherort.

Weitere Informationen finden Sie im *Informatica Big Data Streaming 10.2.2 Service Pack 1-Benutzerhandbuch*.

# Enterprise Data Catalog

In diesem Abschnitt werden die neuen Funktionen für Enterprise Data Catalog in Version 10.2.2 Service Pack 1 erläutert.

## Unterstützte Ressourcentypen für das Standalone Scanner Utility

Ab Version 10.2.2 Service Pack 1 können Sie Metadaten aus den folgenden externen Quellen extrahieren, wenn zur Laufzeit oder offline nicht auf sie zugegriffen werden kann:

- IBM DB2
- IBM DB2 for z/OS
- IBM Netezza
- JDBC
- PowerCenter
- SQL Server Integration Services

Weitere Informationen finden Sie im Kapitel „Metadaten-Extraktion aus Offline- und nicht zugreifbaren Ressourcen“ im *Informatica 10.2.2 Service Pack 1 Enterprise Data Catalog-Administratorhandbuch*.

# Enterprise Data Preparation

In diesem Abschnitt werden neue Funktionen für Enterprise Data Preparation in Version 10.2.2 Service Pack 1 erläutert.

## Zurücksetzen aller abgeleiteten Datentypen

Ab Version 10.2.2 Service Pack 1 können Sie alle abgeleiteten Typen und Datendomänen, die auf Quellenspalten in einem Arbeitsblatt angewendet werden, auf ihre ursprünglichen Typen zurücksetzen.

Weitere Informationen finden Sie im Kapitel „Daten vorbereiten“ des *Informatica 10.2.2 Service Pack 1 Enterprise Data Preparation-Benutzerhandbuchs*.

## Vorbereiten von Avro- und Parquet-Dateien

Ab Version 10.2.2 Service Pack 1 können Sie die hierarchischen Daten in Avro- und Parquet-Dateien, die Sie Ihrem Projekt als ersten Schritt der Datenvorbereitung hinzufügen, auswerten. Enterprise Data Preparation konvertiert die Avro- oder Parquet -Dateistruktur in eine flache Struktur und stellt die Daten in einem Arbeitsblatt dar, mit dem Sie die Daten vorbereiten.

Weitere Informationen finden Sie im Kapitel „Daten vorbereiten“ des *Informatica 10.2.2 Service Pack 1 Enterprise Data Preparation-Benutzerhandbuchs*.

# PowerExchange-Adapter für Informatica

In diesem Abschnitt werden neue Funktionen der Informatica-Adapter in Version 10.2.2 Service Pack 1 erläutert.

## PowerExchange for Hive

Wenn Sie ab Version 10.2.2 Service Pack 1 eine Hive-Tabelle importieren, können Sie mit der Option **Nur Standardschema anzeigen** entweder Suchergebnisse löschen und nur Tabellen anzeigen, die das Standardschema verwenden, oder nach Tabellen suchen, die ein anderes als das Standardschema verwenden.

Weitere Informationen finden Sie im *Informatica PowerExchange for Hive 10.2.2 Service Pack 1-Benutzerhandbuch*.

# KAPITEL 4

## Neue Funktionen (10.2.2)

Dieses Kapitel umfasst die folgenden Themen:

- [Anwendungsdienste, 47](#)
- [Big Data Management, 48](#)
- [Big Data Streaming , 53](#)
- [Befehlszeilenprogramme, 54](#)
- [Enterprise Data Catalog, 59](#)
- [Enterprise Data Lake, 62](#)
- [Informatica Developer, 67](#)
- [Informatica-Zuordnungen, 67](#)
- [Informatica-Umwandlungen, 69](#)
- [PowerExchange-Adapter für Informatica, 72](#)

## Anwendungsdienste

In diesem Abschnitt werden die neuen Anwendungsdienstfunktionen in Version 10.2.2 erläutert.

### Massenerfassungsdienst

Ab Version 10.2.2 können Sie sichere Kommunikation zwischen einem Massenerfassungsdienstprozess und externen Komponenten aktivieren. Sie können die HTTPS-Portnummer für das HTTPS-Protokoll und eine Schlüsselspeicherdatei für das SSL-Zertifikat angeben.

Weitere Informationen finden Sie im Kapitel „Massenerfassungsdienst“ des *Anwendungsdienst-Handbuchs zu Informatica 10.2.2*.

### Metadaten-Zugriffsdienst

Ab Version 10.2.2 können Sie den Metadaten-Zugriffsdienst so konfigurieren, dass Betriebssystemprofile verwendet werden. Verwenden Sie Betriebssystemprofile, um die Sicherheit zu erhöhen und die Entwurfszeit-Benutzerumgebung beim Importieren und Anzeigen einer Vorschau von Metadaten aus einem Hadoop-Cluster zu isolieren.

Weitere Informationen finden Sie im Kapitel „Benutzer und Gruppen“ im *Sicherheitshandbuch zu Informatica 10.2.2*.

## REST Operations Hub-Dienst

Ab Version 10.2.2 können Sie einen REST Operations Hub-Dienst für REST-Anwendungen konfigurieren. Der REST Operations Hub-Dienst ist ein REST-Systemdienst in der Informatica-Domäne. Informatica-Produktfunktionen werden damit externen Kunden gegenüber durch REST-APIs offengelegt.

Sie können den REST Operations Hub-Dienst über das Administrator-Tool oder über infacmd konfigurieren. Sie können den REST Operations Hub-Dienst verwenden, um Statistiken zur Mapping-Ausführung bei den bereitgestellten Mapping-Jobs in der Anwendung anzuzeigen.

Sie können den REST Operations Hub-Dienst verwenden, um Statistiken zur Mapping-Ausführung bei Big-Data-Mappings abzurufen, die auf dem Datenintegrationsdienst oder in der Hadoop-Umgebung ausgeführt werden.

Weitere Informationen zur REST-API finden Sie im *Big Data Management 10.2.2-Administratorhandbuch*.

## Big Data Management

In diesem Abschnitt werden neue Big Data Management-Funktionen in Version 10.2.2 erläutert.

### Integration in Azure Databricks

Ab Version 10.2.2 können Sie die Informatica-Domäne in die Azure Databricks-Umgebung integrieren.

Azure Databricks ist eine Cloud-Analyseplattform, die für Microsoft Azure-Cloud-Dienste optimiert ist. Sie umfasst die Open-Source-Technologien und -funktionen des Apache Spark-Clusters.

Die Informatica-Domäne kann auf einer Azure-VM oder vor Ort installiert werden. Das Verfahren zur Integration läuft ähnlich ab wie bei der Integration in die Hadoop-Umgebung. Sie führen Integrationsaufgaben aus, beispielsweise den Import der Clusterkonfiguration aus der Databricks-Umgebung. Für den Zugriff auf die Databricks-Umgebung verwendet die Informatica-Domäne eine Token-Authentifizierung. Die Databricks-Token-ID wird in der Databricks-Verbindung gespeichert.

#### Quellen und Ziele

Sie können Mappings mit den folgenden Quellen und Zielen innerhalb der Databricks-Umgebung ausführen:

- Microsoft Azure Data Lake Store
- Microsoft Azure Blob Storage
- Microsoft Azure SQL Data Warehouse
- Microsoft Azure Cosmos DB

#### Umwandlungen

Sie können einem Databricks-Mapping die folgenden Umwandlungen hinzufügen:

- Aggregator
- Ausdruck
- Filter
- Joiner
- Lookup
- Normalisierer



Rang  
Router  
Sortierer  
Union

Die Databricks-Spark-Engine verarbeitet die Umwandlung auf ähnliche Weise, wie die Spark-Engine Prozesse in der Hadoop-Umgebung verarbeitet.

## Datentypen

Folgende Datentypen werden unterstützt:

Array  
Bigint  
Datum/Uhrzeit  
Dezimal  
Doppelt  
Ganzzahl  
Map  
Struct  
Text  
String

## Mappings

Beim Konfigurieren eines Mappings haben Sie die Möglichkeit, das Mapping in der Databricks-Umgebung zu validieren und auszuführen. Wenn Sie das Mapping ausführen, generiert der Datenintegrationsdienst Scala-Code und leitet ihn an die Databricks-Spark-Engine weiter.

## Arbeitsabläufe

Sie können Cluster-Workflows entwickeln, um in der Databricks-Umgebung Ephemeral-Cluster zu erstellen.

Weitere Informationen finden Sie in den folgenden Handbüchern:

*Big Data Management 10.2.2-Integrationshandbuch*  
*Big Data Management 10.2.2-Administratorhandbuch*  
*Big Data Management 10.2.2-Benutzerhandbuch*

# Hierarchische Daten

In diesem Abschnitt werden neue Funktionen für hierarchische Daten in Version 10.2.2 beschrieben.

## Dynamische komplexe Ports

Ab Version 10.2.2 können Sie einem dynamischen Mapping, das auf der Spark-Engine ausgeführt wird, dynamische komplexe Ports hinzufügen. Dynamische komplexe Ports bieten Ihnen die Möglichkeit, häufige Schemaänderungen bei hierarchischen Daten in komplexen Dateien zu verwalten.

Ein dynamischer komplexer Port empfängt neue oder veränderte Elemente eines komplexen Ports basierend auf den Schemaänderungen zur Laufzeit. Die Eingaberegeln bestimmen die Elemente eines dynamischen komplexen Ports. Basierend auf den Eingaberegeln empfängt ein dynamischer komplexer Port eines oder mehrere Elemente eines komplexen Ports aus der vorgelagerten Umwandlung. Bei einigen Umwandlungen

auf der Spark-Engine können Sie dynamische komplexe Ports wie dynamisches Array, dynamisches Mapping und dynamische Struktur verwenden.

Weitere Informationen finden Sie im Kapitel zur Verarbeitung hierarchischer Daten mit Schemaänderungen im *Informatica Big Data Management 10.2.2-Benutzerhandbuch*.

## Vorschauanzeige hierarchischer Daten

Ab Version 10.2.2 können Sie hierarchische Daten mit einem Mapping im Developer Tool in der Vorschau anzeigen. Die Anzeige hierarchischer Daten in der Vorschau hilft beim Entwerfen und Debuggen von Big-Data-Mappings.

Sie können Quellen und Umwandlungen als Vorschaupunkte in einem Mapping auswählen, das die folgenden hierarchischen Typen enthält:

- Array
- Struct
- Map

Die Vorschauanzeige hierarchischer Daten ist für die technische Vorschau verfügbar. Funktionen für die technische Vorschau sind noch nicht produktionsreif und werden daher zwar unterstützt, aber nicht garantiert. Informatica empfiehlt, diese Funktionen nur in Nichtproduktionsumgebungen zu verwenden.

Weitere Informationen finden Sie im *Informatica Big Data Management 10.2.2-Benutzerhandbuch*.

## Hohe Verfügbarkeit

In diesem Abschnitt werden neue Hochverfügbarkeitsfunktionen in Version 10.2.2 beschrieben.

### Big Data Jobwiederherstellung

Ab Version 10.2.2 kann der Datenintegrationsdienst einen Big-Data-Job, der für die Ausführung auf der Spark-Engine konfiguriert wurde, bei einem unerwarteten Stopp des Datenintegrationsdienst-Knotens wiederherstellen. Wenn ein Datenintegrationsdienst-Knoten vor Abschluss eines Jobs ausfällt, sendet der Datenintegrationsdienst den Job an einen anderen Knoten. Die Verarbeitung der Job-Aufgaben wird dann von der Stelle des Knotenausfalls an fortgesetzt.

Zum Wiederherstellen von Big-Data-Mappings müssen Sie in den Eigenschaften des Datenintegrationsdiensts die Big-Data-Jobwiederherstellung aktivieren und den Job über `infacmd` ausführen.

Weitere Informationen finden Sie im Kapitel zur Datenintegrationsdienstverarbeitung im *Informatica Big Data Management 10.2.2-Administratorhandbuch*.

### Verteilte Warteschlangen beim Datenintegrationsdienst

Ab Version 10.2.2 verwendet der Datenintegrationsdienst eine verteilte Warteschlange zum Speichern von Jobinformationen, wenn für bereitgestellte Big-Data-Jobs die Wiederherstellung großer Datenmengen aktiviert ist. Die verteilte Warteschlange wird im Modellrepository gespeichert, und ein verfügbarer Datenintegrationsdienst kann Jobs aus der Warteschlange ausführen, sobald Ressourcen verfügbar sind.

Weitere Informationen finden Sie im Kapitel zur Datenintegrationsdienstverarbeitung im *Informatica Big Data Management 10.2.2-Administratorhandbuch*.

# Intelligentes Strukturmodell

In diesem Abschnitt werden neue Funktionen rund um intelligente Strukturmodelle in Version 10.2.2 beschrieben.

## Aliasse in XML-Dateien

Ab Version 10.2.2 kann die intelligente Strukturerkennung XML-Dateien verarbeiten, die verschiedene Aliasse zum Identifizieren desselben Namespace verwenden, wie es bei XML-Dateien der Fall ist, die mit einem intelligenten Strukturmodell erstellt wurden.

## Datentypen

Ab Version 10.2.2 und beginnend mit der März-Version (Winter 2019) von Informatica Intelligent Cloud Services werden bei Verwendung eines intelligenten Strukturmodells durch einen komplexen Datei-Reader die Datentypen von der intelligenten Strukturerkennung an die Ausgabedatenports weitergegeben.

Erkennt die intelligente Strukturerkennung beispielsweise, dass ein Feld ein Datum enthält, gibt die intelligente Strukturerkennung die Daten als Datum und nicht als Zeichenfolge an die Ausgabedatenports weiter.

## Feldnamen

Ab Version 10.2.2 und von der März-Version (Winter 2019) von Informatica Intelligent Cloud Services an können Feldnamen in komplexen Dateidatenobjekten, die Sie aus einem intelligenten Strukturmodell importieren, mit Zahlen und reservierten Wörtern beginnen; zudem dürfen sie folgende Sonderzeichen enthalten: \. [] {} ( ) \* + - ? . ^ \$ |

Wenn ein Feld mit einer Zahl oder einem reservierten Wort beginnt, wird per Mapping von Big Data Management ein Unterstrich ( \_ ) vor dem Feldnamen eingefügt. Beginnt ein Feld in einem intelligenten Strukturmodell beispielsweise mit OR, importiert das Mapping das Feld als \_OR. Wenn der Feldname ein Sonderzeichen enthält, konvertiert das Mapping dieses Zeichen in einen Unterstrich.

## Verarbeitung großer XML-Dateien

Ab Version 10.2.2 kann die intelligente Strukturerkennung XML-Dateien streamen und Daten für sich wiederholende Elemente in Blöcken verarbeiten. Dadurch wird die Verarbeitung großer XML-Dateien effizienter.

## Datendrift

Ab Version 10.2.2 und beginnend mit der März-Version (Winter 2019) von Informatica Intelligent Cloud Services bietet die intelligente Strukturerkennung einen verbesserten Umgang mit Datendriffs.

Bei der intelligenten Strukturerkennung kommen Datendriffs vor, wenn die Eingabedaten Felder enthalten, die in der Beispieldatei nicht vorhanden waren. Die intelligente Strukturerkennung leitet die undefinierten Daten in einem solchen Fall an einen nicht zugewiesenen Datenport beim Ziel weiter, statt die Daten zu verwerfen.

# Massenerfassung

Ab Version 10.2.2 können Sie ein inkrementelles Laden ausführen, um inkrementelle Daten zu erfassen. Beim Ausführen des inkrementellen Ladens ruft die Spark-Engine inkrementelle Daten basierend auf einem Zeitstempel oder einer ID-Spalte ab und lädt die inkrementellen Daten dann in das Hive- oder HDFS-Ziel.

Wenn Sie die Daten in einem Hive-Ziel erfassen, kann die Spark-Engine auch die bei den Quelltabellen vorgenommenen Schemaänderungen verteilen.

Wenn Sie inkrementelle Daten erfassen, nutzt der Massenerfassungsdienst den inkrementellen Importmodus von Sqoop.

Weitere Informationen finden Sie im *Massenerfassungshandbuch zu Informatica Big Data Management 10.2.2*.

## Überwachung

In diesem Abschnitt werden die neuen Funktionen für die Überwachung von Big Data Management in Version 10.2.2 erläutert.

### Spark-Überwachung

Ab Version 10.2.2 können Sie Aufgaben vor und nach dem Job im Bereich „Übersichtsstatistik“ für die Spark-Überwachung anzeigen.

Weitere Informationen zu Aufgaben vor und nach dem Job finden Sie im *Informatica Big Data Management 10.2.2-Benutzerhandbuch*.

## Sicherheit

In diesem Abschnitt werden die neuen Funktionen für die Sicherheit von Big Data Management in Version 10.2.2 erläutert.

### Enterprise-Sicherheitspaket

Ab Version 10.2.2 unterstützt Informatica einen Azure HDInsight-Cluster mit Enterprise-Sicherheitspaket (Enterprise Security Package).

Das Enterprise-Sicherheitspaket verwendet Kerberos zur Authentifizierung und Apache Ranger zur Autorisierung.

Weitere Informationen zum Enterprise-Sicherheitspaket finden Sie im *Informatica Big Data Management 10.2.2-Administratorhandbuch*.

## Ziele

In diesem Abschnitt werden neue Funktionen für Ziele in Version 10.2.2 beschrieben.

### HDFS-Einfachdateiziele

Ab Version 10.2.2 können Sie Ausgabedaten an HDFS-Zieldateien und -Ablehnungsdateien anhängen. Zum Anhängen von Ausgabedaten wählen Sie die anzuhängenden Daten, wenn das HDFS-Ziel vorhanden ist.

Für eine einfachere Verwaltung der Dateien, die angehängte Daten enthalten, hängt der Datenintegrationsdienst die Ausführungskennung des Mappings an die Namen der Zieldateien und Ablehnungsdateien an.

Weitere Informationen finden Sie im Kapitel „Ziele“ im *Informatica Big Data Management 10.2.2-Benutzerhandbuch*.

# Big Data Streaming

In diesem Abschnitt werden neue Big Data Streaming-Funktionen in Version 10.2.2 erläutert.

## Azure Event Hubs-Datenobjekte

Ab Version 10.2.2 können Sie ein Streaming-Mapping mit Event Hub als Quelle in den folgenden Distributionen bereitstellen:

- Amazon EMR
- Azure HDInsight mit ADLS-Storage
- Cloudera CDH
- Hortonworks HDP

## Kontenübergreifende IAM-Rolle in Amazon Kinesis-Verbindungen

Ab Version 10.2.2 können Sie die kontenübergreifende IAM-Rolle verwenden, um eine Amazon Kinesis-Quelle zu authentifizieren.

Die kontenübergreifende IAM-Rolle gibt Ihnen die Möglichkeit, Ressourcen in einem AWS-Konto gemeinsam mit Benutzern in einem anderen AWS-Konto zu verwenden, ohne in jedem Konto die entsprechenden Benutzer anlegen zu müssen.

Weitere Informationen finden Sie im *Informatica Big Data Streaming 10.2.2-Benutzerhandbuch*.

## Intelligentes Strukturmodell

Ab Version 10.2.2 können Sie in Big Data Streaming intelligente Strukturmodelle verwenden.

Sie haben die Möglichkeit, ein intelligentes Strukturmodell in ein Kafka-, Kinesis- oder Azure Event Hubs-Datenobjekt einzubeziehen. Wenn Sie das Datenobjekt einem Mapping hinzufügen, können Sie alle Eingabetypen verarbeiten, die das Modell parsen kann.

Je nachdem, welche Datei Sie zum Erstellen des Modells verwendet haben, kann das Datenobjekt PDF-Formulare, JSON-, Microsoft Excel-, Microsoft Word-Tabellen-, CSV-, Text- oder XML-Eingabedateien akzeptieren und analysieren.

Weitere Informationen finden Sie im *Informatica Big Data Streaming 10.2.2-Benutzerhandbuch*.

## Header-Ports für Big Data Streaming-Datenobjekte

Ab Version 10.2.2 enthalten einige Datenobjekte standardmäßige Header-Ports für die mit den Ereignissen verbundenen Metadaten. Der Zeitstempelport enthält beispielsweise die Uhrzeit, zu der das Ereignis erstellt wird. Sie können die Header-Ports verwenden, um die Daten zu gruppieren und zu verarbeiten.

Weitere Informationen zu den Header-Ports finden Sie im *Informatica Big Data Streaming 10.2.2-Benutzerhandbuch*.

## AWS-Anmeldedatenprofile in Amazon Kinesis-Verbindungen

Ab Version 10.2.2 können Sie in Amazon Kinesis-Verbindungen eine auf AWS-Anmeldedatenprofilen basierende Authentifizierung verwenden.

Wenn Sie eine Amazon Kinesis-Verbindung erstellen, können Sie den Namen eines AWS-Anmeldedatenprofils eingeben. Das Mapping greift über den Profilnamen, der zur Laufzeit in der AWS-Anmeldedatendatei aufgelistet wird, auf die AWS-Anmeldedaten zu.

Weitere Informationen finden Sie im *Informatica Big Data Streaming 10.2.2-Benutzerhandbuch*.

## Strukturiertes Spark-Streaming

Ab Version 10.2.2 nutzt Big Data Streaming zum Verarbeiten von Streaming-Daten Spark Structured Streaming.

Spark Structured Streaming ist eine skalierbare und fehlertolerante Open-Source-Stream-Verarbeitungs-Engine auf Basis der Spark-Engine. Sie ermöglicht den Umgang mit spät eintretenden Streaming-Ereignissen und die Verarbeitung von Streaming-Daten basierend auf dem Zeitstempel der Quelle.

Die Spark-Engine führt das Streaming-Mapping kontinuierlich aus. Sie liest die Daten, unterteilt die Daten in Mikro-Batches, verarbeitet die Mikro-Batches, veröffentlicht die Ergebnisse und schreibt dann in ein Ziel.

Weitere Informationen finden Sie im *Informatica Big Data Streaming 10.2.2-Benutzerhandbuch*.

## Fensterumwandlung

Ab Version 10.2.2 können Sie beim Erstellen einer Fensterumwandlung die folgenden Funktionen verwenden:

### **Wasserzeichenverzögerung**

Mit der Wasserzeichenverzögerung wird festgelegt, welche Schwellenwertzeit für ein verzögertes Ereignis in einer Datengruppe angesammelt werden kann.

Die Wasserzeichenverzögerung ist ein Schwellenwert, mit dem Sie die Dauer angeben, innerhalb der spät eintreffende Daten gruppiert und verarbeitet werden können. Wenn Ereignisdaten innerhalb der Schwellenwertzeit eintreffen, werden die Daten verarbeitet und dann in der entsprechenden Datengruppe gesammelt.

### **Fensterport**

Über den Fensterport wird die Spalte für die Aufnahme der Zeitstempelwerte festgelegt, auf deren Basis Sie die Ereignisse gruppieren können. Die angesammelten Daten enthalten den Zeitstempelwert. Nutzen Sie die Fensterport-Spalte zum Gruppieren der spät eintreffenden Ereigniszeitdaten.

Weitere Informationen finden Sie im *Informatica Big Data Streaming 10.2.2-Benutzerhandbuch*.

## Befehlszeilenprogramme

In diesem Abschnitt werden neue Befehle in Version 10.2.2 erläutert.

### infacmd dis-Befehle

In diesem Abschnitt werden neue infacmd dis-Befehlsoptionen und -Ausführungsoptionen beschrieben.

## Aktualisierung von Service-Optionen

Mit dem Befehl zur Aktualisierung von Service-Optionen werden Eigenschaften des Datenintegrationsdiensts aktualisiert.

In der folgenden Tabelle werden neue Optionen für den Befehl `infacmd dis updateServiceOptions` beschrieben:

| Neue Option                                  | Beschreibung   |
|--|--|
| <code>-RecycleMode</code><br><code>rm</code> | Optional. Mit dem Wiederherstellungsmodus wird der Dienst neu gestartet, und die neuesten Dienst- und Dienstprozeßeigenschaften werden angewendet. Wählen Sie „Abbrechen“ oder „Abschließen“. <ul style="list-style-type: none"><li>- Abschließen. Stoppt alle Anwendungen und bricht alle Jobs in sämtlichen Anwendungen ab. Wartet vor der Deaktivierung des Diensts, bis alle Jobs abgebrochen wurden.</li><li>- Abbrechen. Stoppt alle Anwendungen und versucht, alle Jobs vor deren Abbruch und Deaktivieren des Diensts anzuhalten.</li></ul> Standard: Abschließen. |

In der folgenden Tabelle werden neue Optionen für die Ausführung des Befehls `infacmd dis updateServiceOptions` beschrieben:

| Neue Option                                      | Beschreibung  |
|--|---|
| <code>ExecutionOptions.BigDataJobRecovery</code> | Aktiviert die Big-Data-Jobwiederherstellung. Einstellung auf „true“ oder „false“.<br>Standard: false. |

Weitere Informationen finden Sie im Kapitel „`infacmd dis`-Befehlsreferenz“ der *Informatica 10.2.2-Befehlsreferenz*.

## infacmd ihs-Befehle

In der folgenden Tabelle werden neue Optionen für den Befehl `infacmd ihs` beschrieben:

| Neue Option                                   | Beschreibung   |
|---|--|
| <code>-PrimaryNode</code><br><code>-nm</code> | Optional. Primärer Knoten, auf dem der Dienst ausgeführt wird.   |
| <code>-BackupNodes</code><br><code>-bn</code> | Optional. Knoten, auf denen der Dienst ausgeführt werden kann, wenn der primäre Knoten nicht verfügbar ist. Sie können Backup-Knoten konfigurieren, wenn Sie hohe Verfügbarkeit haben. |

In der folgenden Tabelle werden neue `infacmd ihs`-Befehle beschrieben:

| Befehl                    | Beschreibung                              |
|---------------------------|---|
| <code>cleanCluster</code> | Bereinigt den Informatica-Cluster-Dienst. |

Weitere Informationen finden Sie im Kapitel „infacmd ihs-Befehlsreferenz“ der *Informatica 10.2.2-Befehlsreferenz*.

## infacmd ipc-Befehle

In der folgenden Tabelle werden veraltete infacmd ipc-Befehle beschrieben:

| Befehl     | Beschreibung   |
|------------|--|
| ExportToPC | Exportiert Objekte aus dem Modellrepository oder einer Exportdatei und konvertiert sie in PowerCenter-Objekte. |

## infacmd Idm-Befehle

In der folgenden Tabelle werden neue Optionen für den Befehl infacmd Idm beschrieben:

| Neue Option                         | Beschreibung  |
|-------------------------------------|---|
| -PrimaryNode<br>-nm                 | Optional. Wenn Sie Hochverfügbarkeit für Enterprise Data Catalog konfigurieren möchten, geben Sie den Namen des primären Knotens an.  |
| -BackupNodes<br>-bn                 | Optional: Wenn Sie Hochverfügbarkeit für Enterprise Data Catalog konfigurieren möchten, geben Sie eine Liste von durch Kommas getrennten Backup-Knoten-Namen an.  |
| -isNotifyChangeEmailEnabled<br>-cne | Optional. Leben Sie „True“ fest, wenn Sie Änderungsbenachrichtigungen zu Objekten aktivieren möchten. Standardwert ist „False“.   |
| -ExtraJarsPath<br>-ejp              | Optional. Pfad zum Verzeichnis auf dem Computer, auf dem Sie die Informatica-Domäne installiert haben. Das Verzeichnis muss die für das Bereitstellen von Enterprise Data Catalog auf einem vorhandenen Cluster mit WANDisco Fusion erforderlichen JAR-Dateien enthalten. |
| -ExtraJarsPath<br>-ejp              | Optional. Pfad zum Verzeichnis auf dem Computer, auf dem Sie die Informatica-Domäne installiert haben. Das Verzeichnis muss die für das Bereitstellen von Enterprise Data Catalog auf einem vorhandenen Cluster mit WANDisco Fusion erforderlichen JAR-Dateien enthalten. |

In der folgenden Tabelle werden neue infacmd Idm-Befehle beschrieben:

| Befehl         | Beschreibung  |
|----------------|---|
| collectAppLogs | Erfasst Protokolldateien für YARN-Anwendungen, die zum Aktivieren des Katalogdiensts ausgeführt werden. |
| publishArchive | Erstellt eine Ressource im Offline-Modus und führt den Scan aus.  |

Weitere Informationen finden Sie im Kapitel „infacmd Idm-Befehlsreferenz“ der *Informatica 10.2.2-Befehlsreferenz*.



## infacmd mi-Befehle

In der folgenden Tabelle werden Änderungen an infacmd mi-Befehlen beschrieben:

| Befehl           | Beschreibung der Änderung   |
|------------------|---|
| createService    | Ab Version 10.2.2 können Sie über die Optionen -HttpsPort, -KeystoreFile und -KeystorePassword festlegen, ob bei den Massenerfassungsdienstprozessen eine sichere Verbindung für die Kommunikation mit externen Komponenten verwendet wird.   |
| extendedRunStats | <p>Ab Version 10.2.2 müssen Sie mit der Option -RunID verwenden die RunID der Massenerfassungsspezifikation festlegen und mit der Option -SourceName den Namen einer Quelltable angeben, um die erweiterten Ausführungsstatistiken für die Quelltable anzuzeigen. Wenn die Quelltable unter Verwendung des inkrementellen Ladens erfasst wurde, zeigen die Ausführungsstatistiken den inkrementellen Schlüssel und den Startwert.</p> <p>Bisher haben Sie die JobID für den Mapping-Job zum Erfassen der Quelltable angegeben.</p> <p>Wenn Sie ein Upgrade auf Version 10.2.2 ausführen, müssen Sie alle Skripts, die infacmd mi extendedRunStats ausführen, dahingehend aktualisieren, dass die neuen Optionen verwendet werden.</p> |
| listSpecRuns     | Ab Version 10.2.2 gibt der Befehl außerdem den Ladetyp zurück, den die Spark-Engine zum Ausführen einer Massenerfassungsspezifikation verwendet.  |
| runSpec          | Ab Version 10.2.2 können Sie die Option -LoadType verwenden, um den Ladetyp für das Ausführen einer Massenerfassungsspezifikation festzulegen. Beim Ladetyp kann es sich um vollständiges oder inkrementelles Laden handeln.  |

Weitere Informationen finden Sie im Kapitel „infacmd mi-Befehlsreferenz“ der *Informatica 10.2.2-Befehlsreferenz*.

## infacmd ms-Befehle

In der folgenden Tabelle werden die neuen infacmd-Befehle „ms“ beschrieben:

| Befehl                        | Beschreibung   |
|-------------------------------|--|
| abortAllJobs                  | Bricht alle bereitgestellten Mapping-Jobs ab, die für die Ausführung auf der Spark-Engine konfiguriert sind. Sie können wahlweise nur Jobs in der Warteschlange, nur aktuell ausgeführte Jobs oder beide Arten von Jobs abbrechen. |
| createConfigurationWithParams | Erstellt eine Clusterkonfiguration über Clusterparameter, die Sie in der Befehlszeile angeben.   |
| listMappingOptions            | Listet Mapping-Optionen in einer Anwendung auf.  |
| purgeDatabaseWorkTables       | Bereinigt alle Jobinformationen aus der Warteschlange, wenn Sie beim Datenintegrationsdienst die Wiederherstellung großer Datenmengen aktivieren.  |
| updateMappingOptions          | Aktualisiert Mapping-Optionen in einer Anwendung.  |
| updateOptimizationLevel       | Aktualisiert die Optimierungsebene für mehrere Mappings in einer Anwendung.  |

Weitere Informationen finden Sie im Kapitel „infacmd ms-Befehlsreferenz“ der *Informatica -Befehlsreferenz*.

## infacmd oie-Befehle

Ab Version 10.2.2 ist das oie-Plugin veraltet; die Unterstützung für das Plugin wird in einer zukünftigen Version entfallen. Die infacmd oie-Befehle wurden auf das Tools-Plugin migriert. Einzelheiten hierzu finden Sie im Abschnitt zu ["infacmd tools-Befehle" auf Seite 58](#).

## infacmd tools-Befehle

Ab Version 10.2.2 ersetzt das infacmd-Tools-Plugin das oie-Plugin. Das Tools-Plugin führt die bisher vom oie-Plugin ausgeführten Vorgänge zum Importieren und Exportieren von Objekten aus; Sie können damit außerdem einige zusätzliche Vorgänge ausführen.

In der folgenden Tabelle sind die infacmd oie-Befehle aufgeführt, die auf das Tools-Plugin migriert wurden:

| Früherer Befehl               | Aktueller Befehl                |
|-------------------------------|---------------------------------|
| infacmd oie deployApplication | infacmd tools deployApplication |
| infacmd oie exportObjects     | infacmd tools exportObjects     |
| infacmd oie exportResources   | infacmd tools exportResources   |
| infacmd oie ImportObjects     | infacmd tools importObjects     |

In der folgenden Tabelle werden neue infacmd-Tools-Befehle beschrieben:

| Befehl           | Beschreibung   |
|------------------|--|
| patchApplication | Stellt mithilfe einer .piar-Datei einen Anwendungs-Patch für einen Datenintegrationsdienst bereit. |

Weitere Informationen finden Sie im Kapitel „infacmd tools-Befehlsreferenz“ in der *Informatica 10.2.2-Befehlsreferenz*.

## infasetup-Befehle

In der folgenden Tabelle werden veränderte infasetup-Befehle beschrieben:

| Befehl                 | Beschreibung   |
|------------------------|--|
| DefineDomain           | Ab Version 10.2.2 wird dem Befehl DefineDomain die Option -spid hinzugefügt.           |
| updateDomainSamlConfig | Ab Version 10.2.2 wird dem Befehl updateDomainSamlConfig die Option -spid hinzugefügt. |

Weitere Informationen finden Sie im Kapitel „infasetup-Befehlsreferenz“ der *Informatica 10.2.2-Befehlsreferenz*.

# Enterprise Data Catalog

In diesem Abschnitt werden die neuen Funktionen von Enterprise Data Catalog in Version 10.2.2 beschrieben.

## Automatische Zuweisung der Geschäftsbezeichnung zu einer Spalte

Ab Version 10.2.2 leitet Enterprise Data Catalog aus dem Geschäftsbegriff die Geschäftsbezeichnung für Spaltenobjekte ab. Ein Geschäftsbegriff mit einem Zuverlässigkeitsfaktor von 80 Prozent oder mehr wird als Geschäftsbezeichnung einer Spalte übernommen. Der Zuverlässigkeitsfaktor wird durch einen internen Algorithmus berechnet. Die Berechnungsgrundlage bilden dabei akzeptierte Geschäftsbegriffe bei Datendomänen, die Ähnlichkeit von Spaltennamen sowie eine Namensähnlichkeit zwischen der Spalte und dem Geschäftsbegriff.

Weitere Informationen finden Sie im Kapitel zum Durchführen von Objektaufgaben im *Informatica 10.2.2 Enterprise Data Catalog-Benutzerhandbuch*.

## Benutzerzusammenarbeit bei Objekten

Ab Version 10.2.2 können Sie gemeinsam mit anderen Benutzern von Enterprise Data Catalog an Objekten arbeiten. Die Zusammenarbeit bei Objekten eröffnet Ihnen die Möglichkeit zur Interaktion mit den anderen Benutzern. Sie können Erkenntnisse zu den Objekten weitergeben, Fragen zu den Objekten stellen, alle Objektänderungen nachverfolgen und die Objekte zertifizieren.

Die Benutzerzusammenarbeit bei Objekten ist mit folgenden Möglichkeiten verbunden:

### **Folgen von Objekten**

Sie können Objekten folgen, um Objektänderungen im Katalog zu überwachen. Wenn Sie einem Objekt folgen, werden Sie über Änderungen informiert, die andere Benutzer an dem Objekt vornehmen. Durch diese Überwachung können Sie gegebenenfalls notwendige Maßnahmen einleiten.

### **Bewerten und Begutachten von Objekten**

Sie können Objekte anhand einer Fünf-Sterne-Skala im Katalog bewerten und begutachten. Beim Bewerten und Begutachten von Objekten geben Sie Feedback ausgehend von unterschiedlichen Aspekten des Objekts, etwa Qualität, Verwendbarkeit, Gebrauchstauglichkeit und Verfügbarkeit.

### **Nachfragen zu Objekten**

Wenn Sie mehr über ein Objekt im Katalog erfahren möchten, können Sie Fragen zum Objekt stellen. Beschreibungsbezogene Fragen sind ebenso möglich wie erkundende, vorausschauende und ursachenbezogene Fragen.

### **Zertifizieren von Objekten**

Durch das Zertifizieren eines Objekts empfehlen Sie es gewissermaßen. Für andere Benutzer ist das Objekt damit im Vergleich zu nicht zertifizierten Objekten vertrauenswürdiger.

Weitere Informationen finden Sie im Kapitel zur Benutzerzusammenarbeit bei Objekten im *Informatica 10.2.2 Enterprise Data Catalog-Benutzerhandbuch*.

## Erstellung der Anwendungsdienste von Enterprise Data Catalog unter Verwendung des Installationsprogramms

Ab Version 10.2.2 können Sie das Installationsprogramm verwenden, um nach der Installation von Enterprise Data Catalog die Anwendungsdienste von Enterprise Data Catalog zu erstellen. Sollten Sie bei der Installation

von Enterprise Data Catalog keine Anwendungsdienste erstellt haben, steht Ihnen das Installationsprogramm dafür zur Verfügung.

Weitere Informationen zum Erstellen der Anwendungsdienste unter Verwendung des Installationsprogramms finden Sie im *Installations- und Konfigurationshandbuch zu Informatica Enterprise Data Catalog 10.2.2*.

## Dienstprogramm zur Validierung benutzerdefinierter Metadaten

Ab Version 10.2.2 können Sie ein eigenständiges Java-basiertes Befehlszeilendienstprogramm zur Validierung der Syntax und Semantik von benutzerdefinierten Metadaten einsetzen, die in den Katalog aufgenommen werden sollen. Benutzerdefinierte Metadaten sind Metadaten, die Sie aus benutzerdefinierten Datenquellen einbinden möchten, für die Enterprise Data Catalog keine Ressource bietet.

Weitere Informationen zur Verwendung des Dienstprogramms finden Sie im Knowledgebase-Artikel zum Validieren benutzerdefinierter Metadaten vor der Aufnahme in den Katalog. Eine Anleitung zum Herunterladen des Dienstprogramms erhalten Sie beim globalen Kundensupport von Informatica.

## Änderungsbenachrichtigungen

Ab Version 10.2.2 werden in Enterprise Data Catalog Benachrichtigungen angezeigt, wenn an Objekten, denen man folgt, Änderungen vorgenommen werden. Zu den verschiedenen Benachrichtigungstypen gehören Anwendungsbenachrichtigungen, E-Mail-Benachrichtigungen zu Änderungen sowie E-Mail-Sammelbenachrichtigungen.

Weitere Informationen finden Sie im Kapitel zur Benutzerzusammenarbeit bei Objekten im *Informatica 10.2.2 Enterprise Data Catalog-Benutzerhandbuch*.

## Business Glossary Zuweisungsbericht

Ab Version 10.2.2 können Sie mehrere empfohlene Geschäftsbegriffe für eine Ressource übernehmen oder ablehnen.

Weitere Informationen finden Sie im Kapitel zum Durchführen von Objektaufgaben im *Informatica 10.2.2 Enterprise Data Catalog-Benutzerhandbuch*.

## Betriebssystemprofile

Ab Version 10.2.2 können Sie ein Betriebssystemprofil auswählen, wenn bei Ihnen kein standardmäßiges Betriebssystemprofil festgelegt ist. Der Datenintegrationsdienst verwendet für Data Discovery-Vorgänge die Benutzeranmeldedaten des Betriebssystemprofils. Bei der Datenerkennung werden Spaltenprofile und Datendomänenerkennungsprofile einbezogen.

Weitere Informationen zur Verwendung der Betriebssystemprofile in Enterprise Data Catalog finden Sie im Kapitel zu Enterprise Data Catalog-Konzepten im *Informatica 10.2.2 Catalog Administrator-Handbuch*.

## REST-APIs

Ab Version 10.2.2 können Sie die folgenden Informatica Enterprise Data Catalog REST-APIs verwenden:

- REST-APIs für Geschäftsbegriffe. Sie können einen akzeptierten, übernommenen oder abgelehnten Geschäftsbegriff zurückgeben, aktualisieren oder löschen.
- REST-APIs für Catalog-Ereignisse. Sie können die Benutzerkonfiguration, die E-Mail-Konfiguration sowie Benutzerabonnements abrufen, aktualisieren und löschen.

- APIs für die Objektzertifizierung. Sie können die Zertifizierungseigenschaften für ein Objekt auflisten, aktualisieren und löschen.
- APIs für Objektkommentare. Sie können Kommentare, Er widerungen und Stimmen für ein Datenobjekt auflisten, erstellen, aktualisieren und löschen.
- APIs für Objektüberprüfungen. Sie können Rezensionen, Bewertungen und Stimmen zur Überprüfung auflisten, erstellen, aktualisieren und löschen.

Weitere Informationen zu den REST-APIs finden Sie im *Informatica 10.2.2 Enterprise Data Catalog REST-API-Benutzerhandbuch*.

## Quellmetadatenfilter und Datenprofilfilter

Ab Version 10.2.2 können Sie den Quellmetadatenfilter und Datenprofilfilter verwenden, um Quelltabellen und Ansichten in einer Ressourcenausführung anzugeben. Wenn Sie diese Filter einsetzen, extrahiert Enterprise Data Catalog Quellmetadaten und Profilm etadaten aus spezifischen Tabellen und Ansichten.

Weitere Informationen zum Quellmetadatenfilter und Datenprofilfilter finden Sie im Kapitel zum Ressourcenmanagement im *Informatica 10.2.2 Catalog Administrator-Handbuch*.

## Scanner-Dienstprogramm

Ab Version 10.2.2 bietet Informatica ein eigenständiges Scanner-Dienstprogramm, mit dem Sie Metadaten aus Offline-Ressourcen und nicht zugänglichen Ressourcen extrahieren können. Das Dienstprogramm enthält ein Skript, das Sie mit den zugehörigen Befehlen in einer Sequenz ausführen müssen.

Weitere Informationen zum eigenständigen Scanner-Dienstprogramm finden Sie im Anhang zur Metadatenextraktion aus Offline-Ressourcen und nicht zugänglichen Ressourcen im *Informatica 10.2.2 Catalog Administrator-Handbuch*.

## Ressourcentypen

Ab Version 10.2.2 können Sie Ressourcen für die folgenden Datenquellentypen erstellen:

### Google BigQuery

Sie können Metadaten sowie Beziehungs- und Herkunftsinformationen aus den folgenden Objekten in einer Google BigQuery-Datenquelle extrahieren:

- Projekt
- Datensatz
- Tabelle
- Ansicht

Weitere Informationen zur Konfiguration einer Google BigQuery-Datenquelle finden Sie im *Informatica 10.2.2 Catalog Administrator-Handbuch*.

### Workday

Sie können Metadaten sowie Beziehungs- und Herkunftsinformationen aus den folgenden Objekten in einer Workday-Datenquelle extrahieren:

- Dienst
- Entität
- Bericht

- Betrieb
- Datenquelle
- Eigenschaft
- Business-Objekte

Weitere Informationen zur Konfiguration einer Workday-Datenquelle finden Sie im *Informatica 10.2.2 Catalog Administrator-Handbuch* .

## Enterprise Data Lake

In diesem Abschnitt werden neue Enterprise Data Lake-Funktionen in Version 10.2.2 erläutert.

### Anwendung aktiver Regeln

Ab Version 10.2.2 können Sie aktive Regeln in Projekten verwenden.

Aktive Regeln sind Mapplets, die mit dem Developer Tool entwickelt werden. Sie können aktive Regeln einsetzen, um komplexe Umwandlungen anzuwenden, beispielsweise Aggregator- und Datenqualitätsumwandlungen bei Arbeitsblättern zum Zweck des Abgleichs und der Konsolidierung.

Bei einer aktiven Regel werden alle Zeilen innerhalb eines Datensatzes als Eingabe verwendet. Sie können mehrere Arbeitsblätter zur Verwendung als Eingabe für die Regel auswählen. Die Anwendung fügt ein Arbeitsblatt hinzu, das die Regelausgabe für das Projekt enthält.

Weitere Informationen finden Sie im Kapitel „Daten vorbereiten“ des *Informatica 10.2.2 Enterprise Data Lake-Benutzerhandbuchs*.

### Löschung doppelter Zeilen

Ab Version 10.2.2 können Sie Zeilen, die Duplikatwerte enthalten, aus einem Arbeitsblatt löschen.

Weitere Informationen finden Sie im Kapitel „Daten vorbereiten“ des *Informatica 10.2.2 Enterprise Data Lake-Benutzerhandbuchs*.

### Clustering und Kategorisierung von Spaltendaten

Ab Version 10.2.2 können Sie ähnliche Werte in einer Spalte clustern und anschließend basierend auf Empfehlungen von Enterprise Data Lake kategorisieren. Zum Clustern ähnlicher Werte wird in der Anwendung ein phonetischer Algorithmus verwendet. Danach wird Ihnen vorgeschlagen, die weniger häufig vorkommenden Werte durch den am häufigsten vorkommenden Wert zu ersetzen.

Weitere Informationen finden Sie im Kapitel „Daten vorbereiten“ des *Informatica 10.2.2 Enterprise Data Lake-Benutzerhandbuchs*.

### CLAIRE-basierte Empfehlungen

Ab Version 10.2.2 wird in der Anwendung die eingebettete CLAIRE-Discovery-Engine für maschinelles Lernen verwendet, um bei der Datenvorbereitung Empfehlungen zu geben.

Auf der Projektseite blendet die Anwendung wechselnde und zusätzliche Empfehlungen ein, die basierend auf dem Datenverlauf sowie auf dokumentierten Primärschlüssel-Fremdschlüssel-Beziehungen von Upstream-Datenquellen abgeleitet werden.

Wenn Sie während der Datenvorbereitung in einem Arbeitsblatt eine Spalte auswählen, zeigt die Anwendung ausgehend vom Spaltendatentyp in der Spaltenübersicht Vorschläge zum Verbessern der Daten an.

Wenn Sie bei zwei Arbeitsblättern eine Join-Operation durchführen, greift die Anwendung auf Primärschlüssel-Fremdschlüssel-Beziehungen zurück, um bei geringen Überschneidungen von gewünschten Schlüsselpaaren inkompatible Stichproben anzuzeigen.

Weitere Informationen finden Sie im Kapitel „Daten vorbereiten“ des *Informatica 10.2.2 Enterprise Data Lake-Benutzerhandbuchs*.

## Bedingte Aggregation

Ab Version 10.2.2 können Sie die AND- und OR-Logik nutzen, um mehrere Bedingungen auf IF-Berechnungen anzuwenden, die Sie beim Erstellen eines aggregierten Arbeitsblatts in einem Projekt einsetzen.

- Wenn Sie mehr als eine Spalte in eine Bedingung einbeziehen möchten, verwenden Sie AND in Kombination mit allen Operatoren.
- Wenn Sie mehr als einen Wert innerhalb einer Spalte in eine Bedingung einbeziehen möchten, verwenden Sie OR in Kombination mit den Operatoren IS, IS NOT und IS BETWEEN.

Weitere Informationen finden Sie im Kapitel „Daten vorbereiten“ des *Informatica 10.2.2 Enterprise Data Lake-Benutzerhandbuchs*.

## Datenmaskierung

Ab Version 10.2.2 ist Enterprise Data Lake in Informatica Dynamic Data Masking integriert. Das Datensicherheitsprodukt ermöglicht es, sensible Daten in Datenobjekten zu maskieren.

Zum Aktivieren der Datenmaskierung in Enterprise Data Lake konfigurieren Sie den Dynamic Data Masking-Server so, dass Maskierungsregeln auf Datenobjekte im Datensee angewendet werden. Konfigurieren Sie außerdem die Informatica-Domäne, um Enterprise Data Lake den Verbindungsaufbau zum Dynamic Data Masking-Server zu ermöglichen.

Dynamic Data Masking fängt Anfragen ab, die von Enterprise Data Lake an den Datensee gesendet werden, und wendet die Maskierungsregeln auf Spalten im angeforderten Objekt an. Wenn Benutzer von Enterprise Data Lake Vorgänge an Spalten betrachten oder ausführen, die maskierte Daten enthalten, werden die betreffenden Daten ausgehend von den angewendeten Maskierungsregeln ganz oder teilweise verschleiert.

Weitere Informationen finden Sie im Kapitel zum Maskieren sensibler Daten im *Informatica 10.2.2 Enterprise Data Lake-Administratorhandbuch*.

## Lokalisierung

Ab Version 10.2.2 unterstützt die Benutzeroberfläche die japanische Sprache. Außerdem haben Sie die Möglichkeit, in Projektnamen und -beschreibungen nichtlateinische Zeichen zu verwenden.

## Partitionierte Quellen und Ziele

Ab Version 10.2.2 kann Enterprise Data Lake bei Operationen zum Importieren, Veröffentlichen und Kopieren Daten aus partitionierten Quellen lesen. Die Anwendung kann bei Vorgängen zum Importieren, Veröffentlichen, Kopieren und Hochladen außerdem Daten an partitionierte Ziele im Datensee anhängen.

## Hinzufügung von Kommentaren zu Rezeptschritten

Ab Version 10.2.2 können Sie einem Rezeptschritt einen Kommentar hinzufügen. Verwenden Sie Kommentare, um die Zusammenarbeit zu verbessern und Detailangaben zum Erfüllen von Audit-Auflagen zu machen.

Weitere Informationen finden Sie im Kapitel „Daten vorbereiten“ des *Informatica 10.2.2 Enterprise Data Lake-Benutzerhandbuchs*.

## Speicherung eines Rezepts als Mapping

Ab Version 10.2.2 können Sie ein Rezept als Mapping speichern, statt das Rezept zu veröffentlichen und eine neue Ausgabetabelle zu erstellen.

Sie können das Mapping wahlweise in dem Modellrepository speichern, das mit dem Enterprise Data Lake Service verbunden ist, oder es als XML-Datei speichern. Entwickler haben die Möglichkeit, das Mapping mit dem Developer Tool zu überprüfen und zu verändern; anschließend können sie das Mapping je nach Verfügbarkeit der Systemressourcen gegebenenfalls ausführen.

Weitere Informationen finden Sie im Kapitel „Daten vorbereiten“ des *Informatica 10.2.2 Enterprise Data Lake-Benutzerhandbuchs*.

## Amazon S3, ADLS, WASB und MapR-FS als Datenquellen

Ab Version 10.2.2 können Sie Daten in Dateien vorbereiten, die in den folgenden Datenquellen gespeichert sind:

- Amazon S3
- MapR-FS
- Microsoft Azure Data Lake Storage
- Windows Azure Storage Blob

Für jede Datenquelle, die vorzubereitende Daten enthält, müssen Sie eine Ressource in Enterprise Data Catalog erstellen. Eine Ressource ist ein Repository-Objekt, das eine externe Datenquelle oder ein Metadaten-Repository darstellt. Mit einer Ressource verknüpfte Scanner extrahieren Metadaten aus der Ressource und speichern die Metadaten in Enterprise Data Catalog.

Weitere Informationen zur Erstellung von Ressourcen in Enterprise Data Catalog finden Sie im Kapitel zum Ressourcenmanagement im *Informatica 10.2.2 Catalog Administrator-Handbuch*.

## Statistische Funktionen

Ab Version 10.2.2 können Sie beim Vorbereiten von Daten die folgenden statistischen Funktionen auf Spalten in einem Arbeitsblatt anwenden:

- AVG
- AVGIF
- COUNT
- COUNTIF
- COUNTDISTINCT
- COUNTDISTINCTIF
- MAX
- MAXIF



- MIN
- MINIF
- STDDEV
- STDDEVIF
- SUM
- SUMIF
- VARIANCE
- VARIANCEIF

Weitere Informationen finden Sie im Kapitel „Daten vorbereiten“ des *Informatica 10.2.2 Enterprise Data Lake-Benutzerhandbuchs*.

## Funktionen für Datum und Uhrzeit

Ab Version 10.2.2 können Sie beim Vorbereiten von Daten die folgenden Funktionen für Datum und Uhrzeit auf Spalten in einem Arbeitsblatt anwenden:

- ADD\_TO\_DATE
- CURRENT\_DATETIME
- DATETIME
- DATE\_DIFF
- DATE\_TO\_UNIXTIME
- EXTRACT\_MONTH\_NAME
- UNIXTIME\_TO\_DATE
- Datum in Text umwandeln
- Text in Datum umwandeln

Weitere Informationen finden Sie im Kapitel „Daten vorbereiten“ des *Informatica 10.2.2 Enterprise Data Lake-Benutzerhandbuchs*.

## Mathematische Funktionen

Ab Version 10.2.2 können Sie beim Vorbereiten von Daten die folgenden mathematischen Funktionen auf Spalten anwenden:

- EXP
- LN
- LOG
- PI
- POWER
- SQRT

Weitere Informationen finden Sie im Kapitel „Daten vorbereiten“ des *Informatica 10.2.2 Enterprise Data Lake-Benutzerhandbuchs*.

## Textfunktionen

Ab Version 10.2.2 können Sie beim Vorbereiten von Daten die folgenden Textfunktionen auf Spalten anwenden:

- ENDSWITH
- ENDSWITH\_IGNORE\_CASE
- FIND\_IGNORE\_CASE
- FIND\_REGEX
- FIRST\_CHARACTER\_TO\_NUMBER
- NUMBER\_TO\_CHARACTER
- PROPER\_CASE
- REMOVE\_NON\_ALPHANUMERIC\_CHARACTERS
- STARTSWITH
- STARTSWITH\_IGNORE\_CASE
- SUBSTITUTE\_REGEX
- TRIM\_ALL
- Datum in Text umwandeln
- Zahl in Text umwandeln
- Text in Datum umwandeln
- Text in Zahl umwandeln

Weitere Informationen finden Sie im Kapitel „Daten vorbereiten“ des *Informatica 10.2.2 Enterprise Data Lake-Benutzerhandbuchs*.

## Fensterfunktionen

Ab Version 10.2.2 können Sie Fensterfunktionen verwenden, um Operationen für Gruppen von Zeilen innerhalb eines Arbeitsblatts auszuführen. Die Gruppe der Zeilen, auf die eine Funktion wirkt, wird als Fenster bezeichnet. Für diese Gruppe legen Sie einen Partitionsschlüssel, einen Sortierschlüssel sowie optionale Offsets fest. Eine Fensterfunktion berechnet einen Rückgabewert für jede Eingabezeile innerhalb des Fensterkontexts.

Mit Fensterfunktionen können Sie die folgenden Aufgaben ausführen:

- Abrufen von Daten aus vorherigen oder nachfolgenden Zeilen
- Berechnen einer kumulativen Summe oder eines kumulativen Durchschnitts basierend auf einer Gruppe von Zeilen
- Zuweisen einer laufenden Zeilennummer für jede Zeile in einer Gruppe von Zeilen
- Ersetzen von Nullwerten in Zeilen durch den vorhergehenden Nicht-Nullwert innerhalb einer Gruppe von Zeilen
- Generieren von Sitzungskennungen, mit denen sich Zeilen basierend auf einem spezifischen Zeitraum gruppieren lassen, etwa für die in einer Protokolldatei erfassten Website-Besuche

Sie die Möglichkeit, mehrere Fensterfunktionen auf ein Arbeitsblatt anzuwenden. Beispielsweise können Sie zunächst eine Funktion anwenden, um die Summe der Werte für jede Zeile nach der aktuellen Zeile innerhalb eines Fensters zu berechnen, und danach eine andere Funktion anwenden, um den Durchschnitt dieser Werte zu berechnen.

Enterprise Data Lake fügt zu jeder Funktion, die Sie auf das Arbeitsblatt anwenden, eine Spalte mit den jeweiligen Ergebnissen hinzu.

Weitere Informationen finden Sie im Kapitel „Daten vorbereiten“ des *Informatica 10.2.2 Enterprise Data Lake-Benutzerhandbuchs*.

## Audit-Ereignisse bereinigen

Ab Version 10.2.2 können Sie den Befehl „infacmd edl purgeevents“ ausführen, um Benutzeraktivitätsereignisse aus der Datenbank für den Audit-Verlauf zu löschen. Optional können Sie den Befehl ausführen, um Projektverlaufsereignisse aus der Datenbank zu löschen.

## Spark-Ausführungs-Engine

Ab Version 10.2.2 nutzt Enterprise Data Lake die Spark-Engine für Aufgaben mit hoher Ressourcenauslastung, beispielsweise Objektveröffentlichung, sowie zur Ausführung aktiver Regel-Mapplets, bei denen die Python-Umwandlung verwendet wird. Durch Einsatz der Spark-Engine für Aufgaben mit hoher Ressourcenauslastung wird eine höhere Leistung erreicht; zudem ist damit eine Bereitstellung von Enterprise Data Lake auf Amazon Elastic MapReduce (EMR) möglich, sodass die automatische Skalierung genutzt werden kann.

# Informatica Developer

In diesem Abschnitt werden neue Developer Tool-Funktionen in Version 10.2.2 erläutert.

## Anwendungen

Ab Version 10.2.2 können Sie inkrementelle Anwendungen erstellen. Eine inkrementelle Anwendung ist eine Anwendung, bei der Sie durch Bereitstellung eines Anwendungs-Patches eine Teilmenge der Anwendungsobjekte aktualisieren können. Der Datenintegrationsdienst aktualisiert die Objekte im Patch, während andere Anwendungsobjekte weiter ausgeführt werden.

Wenn Sie ein Upgrade auf Version 10.2.2 ausführen, erhalten die vorhandenen Anwendungen eine Kennzeichnung als vollständige Anwendungen. Sie können in Version 10.2.2 weiterhin vollständige Anwendungen erstellen, nicht aber eine vollständige Anwendung in eine inkrementelle Anwendung konvertieren.

Weitere Informationen finden Sie in den Kapiteln zur Anwendungsbereitstellung und Anwendungspatchbereitstellung im *Informatica 10.2.2 Developer Tool-Handbuch*.

# Informatica-Zuordnungen

In diesem Abschnitt werden neue Informatica-Mapping-Funktionen in Version 10.2.2 erläutert.

## Datentypen

Ab Version 10.2.2 können Sie bei Stapel-Mappings, die auf der Spark-Engine ausgeführt werden, den Hochpräzisionsmodus aktivieren. Die Spark-Engine kann Dezimalwerte mit bis zu 38 Dezimalstellen verarbeiten.

Weitere Informationen finden Sie im *Informatica Big Data Management 10.2.2-Benutzerhandbuch*.

## Zuordnungsausgaben

Ab Version 10.2.2 können Sie Mapping-Ausgaben in Stapel-Mappings verwenden, die als Mapping-Aufgaben in Workflows auf der Spark-Engine ausgeführt werden. Die Mapping-Ausgaben können Sie im Modellrepository beibehalten oder an Workflow-Variablen binden.

Weitere Informationen finden Sie im Abschnitt zu Mapping-Ausgaben im *Informatica 10.2.2 Developer-Zuordnungshandbuch* sowie im Abschnitt zur Mapping-Aufgabe im *Informatica 10.2.2 Developer - Arbeitsablaufhandbuch*.

## Zuordnungsparameter

Ab Version 10.2.2 können Sie Portausdrücken in Aggregator-, Ausdrucks- und Rangumwandlungen, die in nativen und nicht nativen Umgebungen ausgeführt werden, Ausdrucksparameter zuweisen.

Weitere Informationen finden Sie im Abschnitt zum Zuweisen von Parametern sowie im Abschnitt zu dynamischen Mappings im *Informatica 10.2.2 Developer-Zuordnungshandbuch*.

## Optimierungsebenen

Ab Version 10.2.2 können Sie für Mappings und Mapping-Aufgaben die Optimierungsebene „Auto“ konfigurieren. Mit der Optimierungsebene „Auto“ wendet der Datenintegrationsdienst basierend auf dem Ausführungsmodus und dem Mapping-Inhalt Optimierungen an.

Die standardmäßige Optimierungsebene für neue Mappings ist „Auto“.

Wenn Sie das Upgrade auf Version 10.2.2 durchführen, bleiben die in Mappings konfigurierten Optimierungsebenen unverändert. Um die Optimierungsebene „Auto“ mit aktualisierten Mappings verwenden zu können, müssen Sie die Optimierungsebene manuell ändern.

Weitere Informationen finden Sie im Kapitel „Optimierungsebenen“ des *Mapping-Handbuchs zu Informatica 10.2.2 Developer*.

## Sqoop

Ab Version 10.2.2 können Sie die folgenden neuen Sqoop-Funktionen nutzen:

### **Unterstützung für inkrementelle Datenextraktion**

Sie können ein Sqoop-Mapping konfigurieren, um eine inkrementelle Datenextraktion basierend auf einer ID oder einem Zeitstempel durchzuführen. Bei der inkrementellen Datenextraktion extrahiert Sqoop nur die Daten, bei denen seit der letzten Datenextraktion Änderungen eingetreten sind. Die inkrementelle Datenextraktion steigert die Mapping-Leistung.

### **Unterstützung für Vertica-Verbindungen**

Sie können Sqoop so konfigurieren, dass Daten aus einer Vertica-Quelle gelesen oder in ein Vertica-Ziel geschrieben werden.

### Spark-Engine-Optimierung für Sqoop-Pass-Through-Mappings

Wenn Sie ein Pass-Through-Mapping mit einer Sqoop-Quelle auf der Spark-Engine ausführen, optimiert der Datenintegrationsdienst die Mapping-Leistung in den folgenden Szenarien:

- Sie schreiben Daten in ein Hive-Ziel, das mit einer benutzerdefinierten DDL-Abfrage erstellt wurde.
- Sie schreiben Daten in ein vorhandenes Hive-Ziel, das entweder mit einer benutzerdefinierten DDL-Abfrage partitioniert oder mit einer benutzerdefinierte DDL-Abfrage partitioniert und gebündelt ist.
- Sie schreiben Daten in ein vorhandenes Hive-Ziel, das sowohl partitioniert als auch gebündelt ist.

### Unterstützung für das Argument --infoownername

Sie können das Argument --infoownername konfigurieren, um anzugeben, ob Sqoop den Eigentümernamen für ein Datenobjekt beachten muss.

Weitere Informationen finden Sie im *Informatica Big Data Management 10.2.2-Benutzerhandbuch*.

## Informatica-Umwandlungen

In diesem Abschnitt werden neue Funktionen in Informatica-Umwandlungen in Version 10.2.2 erläutert.

### Adress-Validiererumwandlung

In diesem Abschnitt werden neue Funktionen für die Adress-Validiererumwandlung erläutert.

Die Adress-Validiererumwandlung enthält zusätzliche Adressfunktionen für die folgenden Länder:

#### Alle Länder

Ab Version 10.2.2 unterstützt die Adress-Validiererumwandlung in jedem Land, für das Informatica Adressreferenzdaten bereitstellt, die Überprüfung von Adressen mit einer Zeile.

In früheren Versionen unterstützte die Umwandlung eine Überprüfung von Adressen mit einer Zeile in 26 Ländern.

Um eine Adresse mit einer Zeile zu überprüfen, geben Sie die Adresse im Port „Vollständige Adresse“ ein. Wenn sich die Adresse auf ein Land bezieht, bei dem das standardmäßig bevorzugte Skript nicht Latin oder Western ist, verwenden Sie bei der Umwandlung für die Adresse die standardmäßige Eigenschaft „Bevorzugtes Skript“.

#### Australien

Ab Version 10.2.2 können Sie die Adress-Validiererumwandlung so konfigurieren, dass australischen Anschriften Adressanreicherungen hinzugefügt werden. Mit den Anreicherungen lassen sich die geografischen Sektoren und Regionen auffinden, denen das Australia Bureau of Statistics (das australische Amt für Statistik) die Adressen zuweist. Zu den Sektoren und Regionen gehören Volkszählungserhebungsbezirke, Mesh-Blocks (kleinere statistische Raumeinheiten) sowie statistische Erhebungsgebiete.

Zum Bereitstellen der Anreicherungen verwendet die Umwandlung folgende Ports:

- Volkszählungserhebung, Distrikt-Code 2006
- Geografisch codierter nationaler Bezeichner der Adressdatei
- Großstadt-Großraum, statistisches Erhebungsgebiet, 5-stellig

- Großstadt-Großraum, Name des statistischen Erhebungsgebiets
- Statistisches Erhebungsgebiet der Ebene 1, 11-stellig
- Statistisches Erhebungsgebiet der Ebene 1, 7-stellig
- Statistisches Erhebungsgebiet der Ebene 2, 9-stellig
- Statistisches Erhebungsgebiet der Ebene 2, 5-stellig
- Name des statistischen Erhebungsgebiets der Ebene 2
- Statistisches Erhebungsgebiet der Ebene 3, 5-stellig
- Name des statistischen Erhebungsgebiets der Ebene 3
- Statistisches Erhebungsgebiet der Ebene 4, 3-stellig
- Name des statistischen Erhebungsgebiets der Ebene 4
- Mesh-Block, 11-stellig, 2011
- Mesh-Block, 11-stellig, 2016
- Code des Staats oder Gebiets
- Name des Staats oder Gebiets
- Zusätzlicher AU-Status

Die Ports sind in der Portgruppe „Ergänzend für AU“ zu finden.

## Bolivien

Ab Version 10.2.2 verbessert die Adress-Validiererumwandlung die syntaktische Analyse und Validierung von Adressen in Bolivien. Außerdem aktualisiert Informatica die Referenzdaten für Bolivien.

Die Umwandlung umfasst darüber hinaus die folgenden Verbesserungen für Bolivien:

- Adressvalidierung auf Straßenebene.
- Geokoordinaten auf Straßenmitteebene bei Adressen in großen Städten.

## Kanada

Informatica führt die folgenden Funktionen und Verbesserungen für Kanada ein:

### **Unterstützung für die Eigenschaft „Global bevorzugter Deskriptor“ bei kanadischen Adressen**

Ab Version 10.2.2 können Sie die Adress-Validiererumwandlung so konfigurieren, dass die Kurz- oder Langform eines Elementdeskriptors zurückgegeben wird.

Die Umwandlung kann die Kurz- oder Langform der folgenden Deskriptoren zurückgeben:

- Straßendeskriptoren
- Richtungswerte
- Gebäudedeskriptoren
- Gebäudeteildeskriptoren

Konfigurieren Sie zur Angabe des Ausgabeformats für die Deskriptoren die Eigenschaft „Global bevorzugter Deskriptor“ in der Umwandlung. Die Eigenschaft gilt für Deskriptoren in englischer und französischer Sprache. Standardmäßig gibt die Umwandlung den Deskriptor in dem Format zurück, das durch die Referenzdaten festgelegt wird. Wenn Sie bei der Eigenschaft die Option PRESERVE INPUT auswählen, hat die Eigenschaft der bevorzugten Sprache Vorrang vor der Eigenschaft „Global bevorzugter Deskriptor“.

### **Unterstützung für CH und CHAMBER als Gebäudeteildeskriptoren**

Ab Version 10.2.2 erkennt die Adress-Validiererumwandlung CH und CHAMBER als Gebäudeteildeskriptoren in kanadischen Adressen.

### **Kolumbien**

Ab Version 10.2.2 verbessert die Adress-Validiererumwandlung die Verarbeitung von Straßendaten in kolumbianischen Adressen. Außerdem aktualisiert Informatica die Referenzdaten für Kolumbien.

### **Frankreich**

Ab Version 10.2.2 führt Informatica die folgenden Verbesserungen für Adressen in Frankreich ein:

- Informatica verbessert die zusätzlichen Referenzdaten für Frankreich.
- Die Adress-Validiererumwandlung weist Adressen mit größerer Genauigkeit IRIS-Einheiten in Frankreich zu. Die Umwandlung verwendet die Hausnummer in der Adresse zum Überprüfen der IRIS-Einheit, zu der die Adresse gehört. Die Verwendung von Hausnummern kann die Zuweisungsgenauigkeit erhöhen, wenn die Adresse nahe an der Grenze zwischen verschiedenen Einheiten liegt.

### **Israel**

Mit der Version 10.2.2 führt Informatica die folgenden Funktionen und Verbesserungen für Israel ein:

#### **Mehrsprachige Unterstützung für Adressen in Israel**

Sie können die Adress-Validiererumwandlung so konfigurieren, dass eine israelische Adresse in englischer oder hebräischer Sprache zurückgegeben wird.

Verwenden Sie die Eigenschaft „Bevorzugte Sprache“, um die bevorzugte Sprache für die Adressen auszuwählen, die von der Umwandlung zurückgegeben werden.

Standardsprache für Adressen in Israel ist Hebräisch. Wenn Adressinformationen in hebräischer Sprache zurückgegeben werden sollen, setzen Sie die Eigenschaft „Bevorzugte Sprache“ auf DATABASE oder ALTERNATIVE\_1. Wenn Adressinformationen in englischer Sprache zurückgegeben werden sollen, setzen Sie die Eigenschaft „Bevorzugte Sprache“ auf ENGLISH oder ALTERNATIVE\_2.

#### **Unterstützung für mehrere Zeichensätze bei Adressen in Israel**

Bei der Adress-Validiererumwandlung können israelische Adressen in hebräischen und lateinischen Zeichensätzen gelesen und geschrieben werden.

Verwenden Sie die Eigenschaft „Bevorzugtes Skript“, um den bevorzugten Zeichensatz für die Adressdaten auszuwählen.

Standardzeichensatz für Adressen in Israel ist Hebräisch. Wenn Sie die Eigenschaft „Bevorzugtes Skript“ auf „Lateinisch“ oder „Lateinisch-1“ setzen, werden hebräische Adressen bei der Umwandlung in lateinische Zeichen transliteriert.

### **Peru**

Ab Version 10.2.2 prüft die Adress-Validiererumwandlung Adressen in Peru auf Hausnummerenebene. Außerdem aktualisiert Informatica die Referenzdaten für Peru.

### **Schweden**

Ab Version 10.2.2 ist bei der Adress-Validiererumwandlung die Überprüfung von Straßennamen in schwedischen Adressen verbessert.

Die Umwandlung verbessert die Überprüfung von Straßennamen auf folgende Weise:

- Die Umwandlung kann Straßennamen erkennen, die mit dem Buchstaben G als Platzhalter für die Endung GATAN enden.

- Die Umwandlung kann Straßennamen erkennen, die mit dem Buchstaben V als Platzhalter für die Endung VÄGEN enden.
- Die Adress-Validiererumwandlung kann Straßennamen mit nicht korrektem Deskriptor erkennen und korrigieren, wenn die Langform oder Kurzform des Deskriptors verwendet wird.

Die Umwandlung kann beispielsweise bei der folgenden Adresse RUNIUSV bzw. RUNIUSVÄGEN zu RUNIUSGATAN korrigieren:

RUNIUSGATAN 7

SE-112 55 STOCKHOLM

## USA

Ab Version 10.2 HotFix 2 können Sie die Adress-Validiererumwandlung so konfigurieren, dass US-amerikanische Adressen erkannt werden, an denen an einem oder mehreren Wochentagen keine Post zustellbar ist.

Zum Erkennen der Adressen verwenden Sie den Port *Tage ohne Lieferung*. Der Port enthält eine siebenstellige Zeichenfolge, die die Wochentage von Sonntag bis Samstag darstellt. Jede Position in der Zeichenfolge steht für einen anderen Tag.

Die Adress-Validiererumwandlung gibt den ersten Buchstaben eines Wochentags an der entsprechenden Position des Ports zurück, wenn an dem betreffenden Tag an der Adresse keine Post empfangen werden kann. Bei den Positionen für die anderen Wochentage gibt die Umwandlung ein Bindestrichsymbol zurück.

Der Wert S----FS würde beim Port „Tage ohne Lieferung“ beispielsweise anzeigen, dass an einer Adresse an Sonntagen, Freitagen und Samstagen keine Post entgegengenommen wird.

Der Port „Tage ohne Lieferung“ befindet sich in der Portgruppe „Spezifisch für USA“ im Basic-Modell. Um Daten am Port „Tage ohne Lieferung“ zu empfangen, führen Sie die Adress-Validiererumwandlung im zertifizierten Modus aus. Die Umwandlung liest die Portwerte aus den Datenbankdateien USA5C129.MD und USA5C130.MD.

Umfassende Informationen zu den Funktionen und Vorgängen der Engine für die Adressenüberprüfungssoftware in Version 10.2.2 finden Sie im *Entwicklerhandbuch für Informatica Address Verification 5.14.0*.

## Updatestrategie-Umwandlung

Ab Version 10.2.2 können Sie bei einem Mapping, das auf der Spark-Engine ausgeführt wird, eine Updatestrategie-Umwandlung verwenden, um relationale Ziele zu aktualisieren.

Bisher konnten Sie eine Updatestrategie-Umwandlung bei einem Mapping, das auf der Spark-Engine ausgeführt wird, ausschließlich zum Aktualisieren von Hive-Zielen verwenden.

Weitere Informationen finden Sie im Kapitel „Updatestrategie-Umwandlung“ im *Developer-Umwandlungshandbuch*.

# PowerExchange-Adapter für Informatica

In diesem Abschnitt werden neue Funktionen der Informatica-Adapter in Version 10.2.2 erläutert.



## PowerExchange for Amazon Redshift

Ab Version 10.2.2 enthält PowerExchange for Amazon Redshift die folgenden Funktionen:

- Für die folgenden Regionen können Lese- und Schreibvorgänge durchgeführt werden:
  - China (Ningxia)
  - EU (Paris)
- Sie können Amazon Redshift-Objekte in einem Mapping als dynamische Quellen und Ziele einsetzen.
- Sie können Oktalwerte von druckbaren und nicht druckbaren ASCII-Zeichen als DELIMITER oder QUOTE verwenden.
- Sie können Pre-SQL- und Post-SQL-Befehle eingeben, um Abfragen für Quell- und Zielobjekte in einem Mapping auszuführen.
- Sie können eine SQL-Abfrage für Lese-Datenobjekte in einem Mapping definieren, um die Standardabfrage zu überschreiben. Sie können eine SQL-Anweisung eingeben, die von der Amazon Redshift-Datenbank unterstützt wird.
- Wenn Sie große Amazon S3-Objekte in mehreren Teilen herunterladen, können Sie die maximale Größe eines Amazon S3-Objekts angeben.
- Beim Lesen von Daten aus einer Amazon Redshift-Quelle können Sie eindeutige Werte lesen.
- Wenn Sie ein Objekt in Amazon S3 hochladen, können Sie die Mindestgröße des Objekts angeben und die Zahl der Threads festlegen, um das Objekt parallel als Set von unabhängigen Teilen hochzuladen.
- Sie können wahlweise eine vorhandene Zieltabelle beibehalten, eine Zieltabelle zur Laufzeit ersetzen oder eine neue Zieltabelle erstellen, wenn die Tabelle im Ziel nicht vorhanden ist.
- Sie können die Updatestrategie-Umwandlungen für ein Amazon Redshift-Ziel in der nativen Umgebung konfigurieren.
- Beim Schreiben von Daten in Amazon Redshift können Sie das Amazon Redshift-Zieltabellenschema und den Tabellennamen zur Laufzeit überschreiben.
- Beim Verbindungstyp ODBC kann der Datenintegrationsdienst die Umwandlungslogik mit quellseitiger oder vollständiger Pushdown-Optimierung zu Amazon Redshift-Quellen und -Zielen verschieben.
- Sie können serverseitige Verschlüsselung mit AWS KMS (AWS Key Management Service) bei Amazon EMR in Version 5.16 sowie Cloudera CDH in Version 5.15 und 5.16 verwenden.
- PowerExchange for Amazon Redshift unterstützt das AWS SDK für Java in Version 1.11.354.

Weitere Informationen finden Sie im *Informatica PowerExchange for Amazon Redshift 10.2.2-Benutzerhandbuch*.

## PowerExchange for Amazon S3

Ab Version 10.2.2 enthält PowerExchange for Amazon S3 die folgenden Funktionen:

- Für die folgenden Regionen können Lese- und Schreibvorgänge durchgeführt werden:
  - China (Ningxia)
  - EU (Paris)
  - AWS GovCloud (USA)
- Sie können Amazon S3-Objekte in einem Mapping als dynamische Quellen und Ziele einsetzen.
- Wenn Sie ein Mapping in der nativen Umgebung oder auf der Spark-Engine ausführen, um Daten aus einer Avro-, Einfach-, JSON-, ORC- oder Parquet-Datei zu lesen, können Sie beim Angeben des Quellverzeichnis- bzw. Quelldateinamens Platzhalterzeichen verwenden.

- Sie können den im Amazon S3-Bucket gespeicherten Objekten ein oder mehrere Tags hinzufügen, um die Objekte zu kategorisieren. Jedes Tag enthält ein Schlüssel-Wert-Paar. Sie können entweder die Schlüssel-Wert-Paare eintragen oder den absoluten Dateipfad angeben, der die Schlüssel-Wert-Paare enthält.
- Sie können den maximalen Schwellenwert zum Herunterladen eines Amazon S3-Objekts in mehreren Teilen festlegen.
- Wenn Sie ein Objekt in Amazon S3 hochladen, können Sie die Mindestgröße des Objekts angeben und die Zahl der Threads festlegen, um das Objekt parallel als Set von unabhängigen Teilen hochzuladen.
- Wenn Sie einen Lese- oder Schreibvorgang für ein Datenobjekt erstellen, können Sie vorhandene Daten im FileName-Port lesen, der den Endpunktnamen und Quellpfad der Datei enthält.
- Beim Erstellen eines Lese- oder Schreibvorgangs für ein Datenobjekt haben Sie direkt auf der Port-Registerkarte die Möglichkeit, neue Spalten hinzuzufügen oder die Spalten zu verändern.
- Beim Erstellen eines Mappings zum Lesen oder Schreiben einer Avro-, JSON-, ORC- oder Parquet-Datei können Sie die Spalten der Quellumwandlungen, Zielumwandlungen oder anderer Umwandlungen aus der Port-Registerkarte kopieren und diese Spalten dann direkt in die Lese- oder Schreiboperation für ein Datenobjekt einfügen.
- Sie können das Amazon S3-Dateiformat aktualisieren, ohne die Spaltenmetadaten im Schema-Feld der Spaltenprojektionseigenschaften zu verlieren – selbst nach dem Konfigurieren der Spaltenprojektionseigenschaften für ein anderes Amazon S3-Dateiformat.
- Sie können serverseitige Verschlüsselung mit AWS KMS (AWS Key Management Service) bei Amazon EMR in Version 5.16 sowie Cloudera CDH in Version 5.15 und 5.16 verwenden.
- PowerExchange for Amazon S3 unterstützt das AWS SDK für Java in Version 1.11.354.

Weitere Informationen finden Sie im *Informatica PowerExchange for Amazon S3 10.2.2-Benutzerhandbuch*.

## PowerExchange for Google BigQuery

Ab Version 10.2.2 können Sie über die Kontextmenüoption **Ziel erstellen** ein Google BigQuery-Ziel erstellen.

Weitere Informationen finden Sie im *Informatica PowerExchange for Google BigQuery 10.2.2-Benutzerhandbuch*.

## PowerExchange for HBase

Ab Version 10.2.2 enthält PowerExchange for HBase die folgenden neuen Funktionen:

- Wenn Sie ein HBase-Datenobjekt erstellen, können Sie ein Betriebssystemprofil auswählen, um die Sicherheit zu erhöhen und die Entwurfszeit-Benutzerumgebung beim Importieren und Anzeigen einer Vorschau von Metadaten aus einem Hadoop-Cluster zu isolieren.

**Hinweis:** Sie können ein Betriebssystemprofil wählen, wenn der Metadaten-Zugriffsdienst für die Verwendung von Betriebssystemprofilen konfiguriert ist. Der Metadaten-Zugriffsdienst importiert die Metadaten mit dem Betriebssystemprofil, das dem Benutzer standardmäßig zugewiesen ist. Sie können das Betriebssystemprofil in der Liste der verfügbaren Betriebssystemprofile ändern.

- Sie können die HBase-Objekte als dynamische Quellen und Ziele in einem Mapping verwenden.
- Sie können ein Mapping für die Spark-Engine ausführen, um Daten in einer HBase-Ressource nachzuschlagen.

Weitere Informationen finden Sie im *Informatica PowerExchange for HBase 10.2.2-Benutzerhandbuch*.

## PowerExchange for HDFS

Ab Version 10.2.2 enthält PowerExchange for HDFS die folgenden neuen Funktionen:

- Wenn Sie ein komplexes Dateidatenobjekt erstellen, können Sie ein Betriebssystemprofil auswählen, um die Sicherheit zu erhöhen und die Entwurfszeit-Benutzerumgebung beim Importieren und Anzeigen einer Vorschau von Metadaten aus einem Hadoop-Cluster zu isolieren.

**Hinweis:** Sie können ein Betriebssystemprofil wählen, wenn der Metadaten-Zugriffsdienst für die Verwendung von Betriebssystemprofilen konfiguriert ist. Der Metadaten-Zugriffsdienst importiert die Metadaten mit dem Betriebssystemprofil, das dem Benutzer standardmäßig zugewiesen ist. Sie können das Betriebssystemprofil in der Liste der verfügbaren Betriebssystemprofile ändern.

- Wenn Sie ein Mapping in der nativen Umgebung oder auf der Spark-Engine ausführen, um Daten aus einem komplexen Dateidatenobjekt zu lesen, können Sie beim Angeben des Quellverzeichnis- bzw. Quelldateinamens Platzhalterzeichen verwenden.

Sie können die folgenden Platzhalterzeichen einsetzen:

### ? (Fragezeichen)

Das Fragezeichen (?) ermöglicht ein Vorkommen eines beliebigen Zeichens.

### \* (Sternchen)

Das Sternchen (\*) ermöglicht null oder mehrere Vorkommen eines beliebigen Zeichens.

- Sie können komplexe Dateiobjekte als dynamische Quellen und Ziele in einem Mapping verwenden.
- Sie können komplexe Dateiobjekte verwenden, um Daten aus einem komplexen Dateisystem zu lesen oder in ein komplexes Dateisystem zu schreiben.
- Wenn Sie ein Mapping in der nativen Umgebung oder auf der Spark-Engine ausführen, um Daten in ein komplexes Dateidatenobjekt zu schreiben, können Sie Zieldaten überschreiben. Der Datenintegrationsdienst löscht die Zieldaten vor dem Schreiben neuer Daten.
- Wenn Sie eine Lese- oder Schreiboperation für ein Datenobjekt erstellen, können Sie die vorhandenen Daten im FileName-Port lesen, der den Endpunktnamen und Quellpfad der Datei enthält.
- Sie können die Datenobjektvorgänge jetzt unmittelbar nach dem Erstellen des Lese- oder Schreibvorgangs für ein Datenobjekt anzeigen.
- Beim Erstellen eines Lese- oder Schreibvorgangs für ein Datenobjekt haben Sie die Möglichkeit, neue Spalten hinzuzufügen oder die Spalten zu verändern.
- Beim Lesen oder Schreiben in eine Avro-, JSON-, ORC- oder Parquet-Datei können Sie die Spalten der Quellumwandlungen, Zielumwandlungen oder anderer Umwandlungen kopieren und diese Spalten dann direkt in den Lese- oder Schreibvorgang für ein Datenobjekt einfügen.

Weitere Informationen finden Sie im *Informatica PowerExchange for HDFS 10.2.2-Benutzerhandbuch*.

## PowerExchange for Hive

Ab Version 10.2.2 enthält PowerExchange for Hive die folgenden neuen Funktionen:

- Für ein Hive-Ziel können Sie die folgenden Zielschemastrategie-Optionen konfigurieren:
  - RETAIN – vorhandenes Zielschema beibehalten
  - CREATE – Tabelle zur Laufzeit erstellen oder ersetzen
  - APPLYNEWCOLUMNS – Tabelle ändern und nur neue Spalten anwenden
  - APPLYNEWSHEMA – Tabelle ändern und neues Schema anwenden
  - FAIL – fehlerhaftes Mapping, wenn das Zielschema abweicht

- Parameter zuweisen
- Vor dem Laden von Daten können Sie ein internes oder externes partitioniertes Hive-Ziel trunkieren. Diese Option ist relevant, wenn Sie das Mapping in der Hadoop-Umgebung ausführen.
- Sie können eine Lese- oder Schreibumwandlung für Hive im nativen Modus erstellen, um Daten aus einer Hive-Quelle zu lesen oder Daten in ein Hive-Ziel zu schreiben.
- Wenn Sie Daten in ein Hive-Ziel schreiben, können Sie für eine Hive-Verbindung die folgenden Eigenschaften konfigurieren:
  - Hive-Staging-Verzeichnis auf HDFS. Dies steht für das HDFS-Verzeichnis für Hive-Staging-Tabellen. Die Option ist relevant und erforderlich, wenn Sie Daten in ein Hive-Ziel in der nativen Umgebung schreiben.
  - Name der Hive-Staging-Datenbank. Dies steht für Namespace für Hive-Staging-Tabellen. Diese Option ist relevant, wenn Sie Mappings in der nativen Umgebung ausführen, um Daten in ein Hive-Ziel zu schreiben. Wenn Sie Mappings auf der Blaze- oder Spark-Engine ausführen, ist es nicht erforderlich, den Namen der Hive-Staging-Datenbank in der Hive-Verbindung zu konfigurieren. Der Datenintegrationsdienst verwendet den Wert, den Sie in der Hadoop-Verbindung konfigurieren.

Weitere Informationen finden Sie im *Informatica PowerExchange for Hive 10.2.2-Benutzerhandbuch*.

## PowerExchange for MapR-DB

Ab Version 10.2.2 können Sie beim Erstellen eines HBase-Datenobjekts für MapR-DB ein Betriebssystemprofil auswählen, um die Sicherheit zu erhöhen und die Entwurfszeit-Benutzerumgebung beim Importieren und Anzeigen einer Vorschau von Metadaten aus einem Hadoop-Cluster zu isolieren.

**Hinweis:** Sie können ein Betriebssystemprofil wählen, wenn der Metadaten-Zugriffsdienst für die Verwendung von Betriebssystemprofilen konfiguriert ist. Der Metadaten-Zugriffsdienst importiert die Metadaten mit dem Betriebssystemprofil, das dem Benutzer standardmäßig zugewiesen ist. Sie können das Betriebssystemprofil in der Liste der verfügbaren Betriebssystemprofile ändern.

Weitere Informationen finden Sie im *Informatica PowerExchange for MapR-DB 10.2.2 -Benutzerhandbuch*.

## PowerExchange for Microsoft Azure Blob Storage

Ab Version 10.2.2 enthält PowerExchange for Microsoft Azure Blob Storage die folgenden Funktionen:

- Sie können Mappings in der Azure Databricks-Umgebung ausführen.
- Sie können die Microsoft Azure-Endpunkte „US Government“ in Mappings konfigurieren.
- Sie können Daten beim Lesen von Daten aus sowie beim Schreiben von Daten in Microsoft Azure Blob Storage in den folgenden Formaten komprimieren:
  - None
  - Deflate
  - Gzip
  - Bzip2
  - Lzo
  - Snappy
- Sie können Microsoft Azure Blob Storage-Objekte als dynamische Quellen und Ziele in einem Mapping verwenden.
- Sie können zur Laufzeit in der nativen Umgebung den Namen der Datei lesen, aus der der Datenintegrationsdienst die Daten liest.

- Sie können in den erweiterten Quell- und Zieleigenschaften den relativen Pfad in **Blob-Container-Überschreibung** konfigurieren.

Weitere Informationen finden Sie im *Informatica PowerExchange for Microsoft Azure Blob Storage 10.2.2-Benutzerhandbuch*.

## PowerExchange for Microsoft Azure Cosmos DB SQL API

Ab Version 10.2.2 bietet PowerExchange for Microsoft Azure Cosmos DB SQL API die folgenden Möglichkeiten:

- Sie können Mappings in der Azure Databricks-Umgebung ausführen. Databricks-Unterstützung für PowerExchange for Microsoft Azure Cosmos DB SQL API steht bei der technischen Vorschau zur Verfügung. Funktionen für die technische Vorschau sind noch nicht produktionsreif und werden daher zwar unterstützt, aber nicht garantiert. Informatica empfiehlt, diese Funktionen nur in Nichtproduktionsumgebungen zu verwenden.

Weitere Informationen finden Sie im *Informatica PowerExchange for Microsoft Azure Cosmos DB SQL API 10.2.2-Benutzerhandbuch*.

## PowerExchange for Microsoft Azure Data Lake Store

Ab Version 10.2.2 bietet PowerExchange for Microsoft Azure Data Lake Store die folgenden Möglichkeiten:

- Sie können Mappings in der Azure Databricks-Umgebung ausführen.
- Sie können komplexe Datentypen wie array, struct und map in Mappings verwenden, die in der Hadoop-Umgebung ausgeführt werden. Bei komplexen Datentypen liest, verarbeitet und schreibt die jeweilige Engine hierarchische Daten direkt in komplexe Avro-, JSON- und Parquet-Dateien. Bei einer intelligent strukturierten Quelle können Sie nur den Lesevorgang konfigurieren.
- Sie können Mappings zum Lesen und Schreiben von Avro- und Parquet-Dateien erstellen, die nur primitive Datentypen in der nativen Umgebung enthalten.
- Sie können ein Verzeichnis in einem Mapping als Quelle auswählen, um mehrere Dateien aus dem Verzeichnis zu lesen.
- Sie können Microsoft Azure Data Lake Store-Objekte als dynamische Quellen und Ziele in einem Mapping verwenden.
- Sie können mit der Option „Ziel erstellen“ ein Microsoft Azure Data Lake Store-Ziel erstellen.

Weitere Informationen finden Sie im *Informatica PowerExchange for Microsoft Azure Data Lake Store 10.2.2-Benutzerhandbuch*.

## PowerExchange for Microsoft Azure SQL Data Warehouse

Ab Version 10.2.2 bietet PowerExchange for Microsoft Azure SQL Data Warehouse die folgenden Möglichkeiten:

- Sie können Mappings in der Azure Databricks-Umgebung ausführen.
- Sie können die Microsoft Azure-Endpunkte „US Government“ in Mappings konfigurieren, die in der nativen Umgebung und auf der Spark-Engine ausgeführt werden.
- Sie können Fehlerdateien im Microsoft Azure Blob Storage-Container generieren. Die Fehlerdateien enthalten abgelehnte Zeilen und die Ursache für diese abgelehnten Zeilen.
- Sie können die Batch-Größe in den erweiterten Zieleigenschaften in der nativen Umgebung festlegen.

- Sie können vollständige Pushdown-Optimierung konfigurieren, um die Umwandlungslogik in Quell- und Zieldatenbanken zu verschieben. Verwenden Sie die Pushdown-Optimierung, um durch Nutzung der Datenbankressourcen die Aufgabenleistung zu verbessern.
- Sie können Microsoft Azure SQL Data Warehouse-Objekte als dynamische Quellen und Ziele in einem Mapping verwenden.

Die vollständige Pushdown-Optimierung steht ebenso wie die Funktion der dynamischen Mappings bei PowerExchange for Microsoft Azure SQL Data Warehouse für die technische Vorschau zur Verfügung. Funktionen für die technische Vorschau sind noch nicht produktionsreif und werden daher zwar unterstützt, aber nicht garantiert. Informatica empfiehlt, diese Funktionen nur in Nichtproduktionsumgebungen zu verwenden.

Weitere Informationen finden Sie im *Informatica PowerExchange for Microsoft Azure SQL Data Warehouse 10.2.2-Benutzerhandbuch*.

## PowerExchange for Salesforce

Ab Version 10.2.2 bietet PowerExchange for Salesforce die folgenden neuen Möglichkeiten:

- Sie können unter Verwendung von Version 43.0 und 44.0 der Salesforce-API eine Salesforce-Verbindung aufbauen und auf Salesforce-Objekte zugreifen.
- Sie können OAuth für Salesforce-Verbindungen konfigurieren.
- Sie können den nativen Ausdrucksfilter für die den Quelldatenobjektvorgang konfigurieren.
- Sie können die folgenden Lesevorgangseigenschaften für ein Salesforce-Datenobjekt parametrisieren:
  - SOQL-Filterbedingung
  - CDC-Start-Zeitstempel
  - CDC-Ende-Zeitstempel
  - PK-Segmentgröße
  - Startzeilen-ID PK-Segmentgröße

Sie können die folgenden Schreibvorgangseigenschaften für ein Salesforce-Datenobjekt parametrisieren:

- Präfix für BULK-Erfolgs- und -Fehlerdateien festlegen
- SFDC-Erfolgsdateiverzeichnis
- Speicherort der BULK-Fehlerdateien festlegen

Weitere Informationen finden Sie im *Informatica PowerExchange for Salesforce 10.2.2-Benutzerhandbuch*.

## PowerExchange for Snowflake

Ab Version 10.2.2 bietet PowerExchange for Snowflake die folgenden neuen Funktionen:

- Sie können eine Okta-SSO-Authentifizierung konfigurieren, indem Sie die Authentifizierungsdetails in den JDBC-URL-Parametern der Snowflake-Verbindung festlegen.
- Sie können ein SQL-Override konfigurieren, um die zum Extrahieren von Daten aus der Snowflake-Quelle verwendete Standard-SQL-Abfrage zu überschreiben. Die Festlegung des SQL-Overrides erfolgt in den Eigenschaften des Lesevorgangs für das Snowflake-Datenobjekt.
- Sie haben die Möglichkeit, die Dateien vor dem Schreiben in Snowflake-Tabellen zu komprimieren und die Schreibleistung zu optimieren. Dazu dienen die erweiterten Eigenschaften. In den erweiterten Eigenschaften des Schreibvorgangs für das Snowflake-Datenobjekt können Sie im Feld für die **zusätzlichen Schreibparameter zur Laufzeit** den Kompressionsparameter ein- bzw. ausschalten.

- Der Datenintegrationsdienst verwendet die Snowflake Spark Connector-APIs, um Snowflake-Mappings auf der Spark-Engine auszuführen.
- Sie können Daten aus Snowflake lesen und Daten in Snowflake schreiben, die für das Staging von Daten in Azure oder Amazon aktiviert sind.

Weitere Informationen finden Sie im *Informatica PowerExchange for Snowflake 10.2.2 -Benutzerhandbuch*.

## PowerExchange for Teradata Parallel Transporter API

Ab Version 10.2.2 umfasst PowerExchange for Teradata Parallel Transporter API die folgenden Funktionen in den erweiterten Zieleigenschaften:

- Sie können ein Ersatzzeichen angeben, das in der Teradata-Datenbank beim Laden von Daten in Ziele anstelle eines nicht unterstützten Teradata-Unicode-Zeichens verwendet wird.
- Wenn Sie für das Laden von Daten in Teradata-Ziele ein Zeichen angegeben haben, das statt eines nicht unterstützten Zeichens verwendet werden soll, können Sie für die Teradata-Zielfdatenbank wahlweise Version 8.x bis 13.x oder 14.x und höher festlegen. Verwenden Sie dieses Attribut in Kombination mit dem Ersatzzeichen-Attribut. Der Datenintegrationsdienst ignoriert dieses Attribut, wenn Sie im Zuge des Ladens von Daten in Teradata-Ziele kein Ersatzzeichen angegeben haben.
- Beim Schreiben von Daten in Teradata können Sie das Teradata-Zieltabellenschema und den Tabellennamen zur Laufzeit überschreiben.

Weitere Informationen finden Sie im *Informatica PowerExchange for Teradata Parallel Transporter API 10.2.2-Benutzerhandbuch*.