



Informatica® Data Quality-Integration für
PowerCenter

10.1.1 HotFix 2

Benutzerhandbuch

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.

Den RECHTEN DER REGIERUNG DER VEREINIGTEN STAATEN unterliegende Programme, Software, Datenbanken und zugehörige Dokumentation und technische Daten, die an Kunden der Regierung der Vereinigten Staaten geliefert werden, sind "kommerzielle Computersoftware" oder "kommerzielle technische Daten" gemäß der anwendbaren Beschaffungsverordnung der Vereinigten Staaten (Federal Acquisition Regulation – FAR) und der ergänzenden Bestimmungen der spezifischen Behörde. Damit unterliegen die Nutzung, das Kopieren, die Offenlegung, das Modifizieren und die Anpassung den im anwendbaren Regierungsvertrag gemachten Einschränkungen und Lizenzbedingungen und, soweit im Rahmen der Bedingungen des Regierungsvertrags und der in FAR 52.227-19 aufgeführten Rechte anwendbar, der Lizenz für die kommerzielle Computersoftware.

Informatica, das Informatica-Logo und PowerCenter sind Marken oder eingetragene Marken der Informatica LLC in den Vereinigten Staaten von Amerika und zahlreichen anderen Ländern der Welt. Eine aktuelle Liste der Informatica-Marken finden Sie im Internet unter <https://www.informatica.com/trademarks.html>. Alle weiteren Produkt- und Firmennamen sind möglicherweise Markennamen oder Marken der jeweiligen Eigentümer.

Teile dieser Software und/oder Dokumentationen unterliegen dem Urheberrecht Dritter. Die erforderlichen Hinweise auf Drittanbieter sind im Lieferumfang des Produkts enthalten.

Weitere Informationen über die Patente finden Sie unter <https://www.informatica.com/legal/patents.html>.

HAFTUNGSAUSSCHLUSS: Informatica LLC stellt diese Dokumentation „wie besehen“ bereit, ohne ausdrückliche oder stillschweigende Gewährleistung, einschließlich, jedoch nicht beschränkt auf die Gewährleistungen der Nichtverletzung der Rechte von Dritten, der Handelsüblichkeit oder Eignung für einen bestimmten Zweck. Informatica LLC garantiert nicht die Fehlerfreiheit dieser Software oder Dokumentation. Die in dieser Software oder Dokumentation bereitgestellten Informationen können technische Ungenauigkeiten oder Druckfehler enthalten. Die in dieser Software und in dieser Dokumentation enthaltenen Informationen können jederzeit ohne vorherige Ankündigung geändert werden.

HINWEISE

Dieses Informatica-Produkt (die „Software“) umfasst bestimmte Treiber (die „DataDirect-Treiber“) von DataDirect Technologies, einem Betreiber von Progress Software Corporation („DataDirect“), die folgenden Bedingungen und Bestimmungen unterliegen:

1. DIE DATADIRECT-TREIBER WERDEN „WIE GESEHEN“ OHNE JEGLICHE GEWÄHRLEISTUNG, WEDER AUSDRÜCKLICH NOCH STILLSCHWEIGEND, BEREITGESTELLT, EINSCHLIESSLICH, JEDOCH NICHT BESCHRÄNKT AUF DIE STILLSCHWEIGENDEN GEWÄHRLEISTUNGEN DER HANDELSÜBLICHKEIT, EIGNUNG FÜR EINEN BESTIMMTEN ZWECK UND DER NICHTVERLETZUNG VON RECHTEN DRITTER.
2. IN KEINEM FALL SIND DATADIRECT ODER DRITTANBIETER DEM ENDBENUTZER GEGENÜBER HAFTBAR FÜR UNMITTELBARE, MITTELBARE, KONKRETE, NEBEN-, FOLGE- ODER ANDERE SCHÄDEN, DIE SICH AUS DER VERWENDUNG DER ODBC-TREIBER ERGEBEN, UNABHÄNGIG DAVON, OB SIE IM VORAUS ÜBER DIE MÖGLICHKEIT SOLCHER SCHÄDEN INFORMIERT WORDEN SIND ODER NICHT. DIESE BESCHRÄNKUNGEN GELTEN FÜR ALLE KLAGEGEGENSTÄNDE, EINSCHLIESSLICH, JEDOCH NICHT BESCHRÄNKT AUF VERTRAGSBRUCH, GEWÄHRLEISTUNGSBRUCH, FAHRLÄSSIGKEIT, KAUSALHAFTUNG, TÄUSCHUNG UND ANDERE UNERLAUBTE HANDLUNGEN.

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.

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.

Publikationsdatum: 2018-07-03

Inhalt

Einleitung	4
Informatica-Ressourcen.	4
Informatica-Netzwerk.	4
Informatica-Wissensdatenbank.	4
Informatica-Dokumentation.	4
Informatica-Produktverfügbarkeitsmatrizen.	5
Informatica Velocity.	5
Informatica Marketplace.	5
Globaler Kundensupport von Informatica.	5
 Kapitel 1: Informatica Data Quality-Integration.....	6
Übersicht über Data Quality-Integration.	6
Funktionen und Funktionsweise von Data Quality.	7
Integrations-Installer.	8
Export eines Objekts nach PowerCenter.	8
Installation von Referenzdaten.	8
Integrationsmatrix.	8
Regeln und Richtlinien für die Produktintegration.	10
Verhalten bei Datenqualitätsumwandlungen in PowerCenter.	10
 Kapitel 2: Installation.....	14
Installation - Übersicht.	14
Systemvoraussetzungen.	14
Vor dem Installieren.	15
Installieren des Server-Plug-Ins.	15
Installieren unter Windows.	15
Installieren unter UNIX.	16
Installieren des Client-Plug-Ins.	16
Registrierung des Plug-Ins für PowerCenter.	17
 Kapitel 3: Referenzdaten-Anforderungen.....	18
Referenzdaten - Übersicht.	18
Anforderungen für Referenzdaten-Dateien.	19
Anforderungen für Adressenreferenzdaten.	20
Anforderungen an Identitätspopulationsdateien.	21
Anforderungen an probabilistische und Klassifizierer-Modelle.	23
Speicherorte der probabilistischen und Klassifizierer-Modelldatei.	23
Speicherzuordnung des PowerCenter-Integrationsdiensts.	23
 Index.....	25

Einleitung

PowerCenter®-Benutzer können mit Mappings und Referenzdatenobjekten arbeiten, die sie aus einer Data Quality-Umgebung importieren. Das *Benutzerhandbuch zur Informatica Data Quality-Integration für PowerCenter* wurde für die PowerCenter-Benutzer geschrieben, die Datenqualitäts-Mappings ausführen und die zugehörigen Referenzdatenobjekte verwalten.

Informatica-Ressourcen

Informatica-Netzwerk

Im Informatica-Netzwerk finden Sie den globalen Kundensupport von Informatica, die Informatica-Wissensdatenbank und andere Produktressourcen. Für den Zugriff auf das Informatica-Netzwerk besuchen Sie <https://network.informatica.com>.

Als Mitglied können Sie:

- zentral auf alle Ihre Informatica-Ressourcen zugreifen.
- Durchsuchen Sie die Wissensdatenbank nach Produktressourcen, einschließlich Dokumentation, häufig gestellter Fragen und bewährter Methoden.
- Zeigen Sie Informationen zur Produktverfügbarkeit an.
- Ihre Support-Fälle prüfen.
- Ihr lokales Informatica-Netzwerk für Benutzergruppen suchen und mit anderen Benutzern zusammenarbeiten.

Informatica-Wissensdatenbank

Verwenden Sie die Informatica-Wissensdatenbank, um das Informatica-Netzwerk nach Produktressourcen, wie z. B. Dokumentation, Ratgeberartikeln, bewährten Methoden und PAMs, zu durchsuchen.

Für den Zugriff auf die Wissensdatenbank besuchen Sie <https://kb.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.

Informatica-Dokumentation

Navigieren Sie zur Informatica-Wissensdatenbank unter https://kb.informatica.com/_layouts/ProductDocumentation/Page/ProductDocumentSearch.aspx, um die aktuelle Dokumentation für Ihr Produkt abzurufen.

Wenn Sie Fragen, Kommentare oder Ideen zu dieser Dokumentation haben, wenden Sie sich per E-Mail an das Informatica-Dokumentationsteam unter infa_documentation@informatica.com.

Informatica-Produktverfügbarkeitsmatrizen

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

Informatica Velocity

Bei Informatica Velocity handelt es sich um eine Sammlung von Tipps und bewährten Methoden, die von den professionellen Informatica-Diensten entwickelt wurden. Informatica Velocity basiert auf der Praxiserfahrung aus Hunderten von Datenmanagementprojekten und umfasst das kollektive Wissen unserer Berater, die mit Unternehmen aus der ganzen Welt an der Planung, Entwicklung, Bereitstellung und Wartung erfolgreicher Datenmanagementlösungen gearbeitet haben.

Als Mitglied des Informatica-Netzwerks können Sie unter <http://velocity.informatica.com> auf Informatica Velocity-Ressourcen zugreifen.

Wenn Sie Fragen, Anregungen oder Ideen zu Informatica Velocity haben, wenden Sie sich an die professionellen Informatica-Dienste unter ips@informatica.com.

Informatica Marketplace

Informatica Marketplace ist ein Forum, das Lösungen zur Erweiterung und Verbesserung Ihrer Informatica-Implementierungen bereitstellt. Indem Sie die zahlreichen Lösungen von Informatica-Entwicklern und -Partnern nutzen, können Sie Ihre Produktivität steigern und die Implementierungsdauer Ihrer Projekte verkürzen. Zugriff auf den Informatica Marketplace erhalten Sie unter <https://marketplace.informatica.com>.

Globaler Kundensupport von Informatica

Sie können sich telefonisch oder über den Online-Support mit einem globalen Support-Center im Informatica-Netzwerk in Verbindung setzen.

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

<http://www.informatica.com/us/services-and-training/support-services/global-support-centers>.

Als Mitglied des Informatica-Netzwerks können Sie den Online-Support unter <http://network.informatica.com> verwenden.

KAPITEL 1

Informatica Data Quality-Integration

Dieses Kapitel umfasst die folgenden Themen:

- [Übersicht über Data Quality-Integration, 6](#)
- [Funktionen und Funktionsweise von Data Quality, 7](#)
- [Integrations-Installer, 8](#)
- [Export eines Objekts nach PowerCenter, 8](#)
- [Installation von Referenzdaten, 8](#)
- [Integrationsmatrix, 8](#)
- [Regeln und Richtlinien für die Produktintegration, 10](#)
- [Verhalten bei Datenqualitätsumwandlungen in PowerCenter, 10](#)

Übersicht über Data Quality-Integration

Informatica-Benutzer können Mappings in Informatica Data Quality erstellen und ausführen. Die Mappings und andere Objekte werden von den Benutzern im Modellrepository gespeichert. Die Benutzer können die Objekte zur Verwendung in PowerCenter exportieren. Als PowerCenter-Benutzer können Sie die Objekte in das PowerCenter-Repository importieren und zu Sitzungen hinzufügen.

Verwenden Sie die Objekte, um den Inhalt und die Struktur Ihrer Daten zu analysieren und zu optimieren.

Sie können die Objekte, die die Benutzer in Data Quality 10.1.1 HotFix 2 erstellt haben, mit den folgenden Versionen von PowerCenter verwenden:

- PowerCenter 10.1.1 mit dem aktuellsten HotFix.
- PowerCenter 10.1.
- PowerCenter 9.6.1 mit dem aktuellsten HotFix.

Wenn Sie die aktuellen Versionen von Data Quality und PowerCenter verwenden, können Sie Objekte, die aus dem Modellrepository stammen, ohne zusätzliche Installation importieren. Wenn Sie eine frühere Version von PowerCenter verwenden, müssen Sie die aktuelle Version des Informatica Data Quality-Integrations-Plug-Ins installieren und registrieren.

Funktionen und Funktionsweise von Data Quality

Informatica Data Quality verwendet Informatica-Anwendungen, um Objekte zu erstellen, Objekte im Modellrepository zu speichern und Mappings auszuführen.

Möglicherweise müssen Sie das Informatica Data Quality-Plug-In für PowerCenter installieren, wenn Sie nicht die aktuellen Produktversionen verwenden.

Verwenden Sie Informatica Data Quality, um Mappings zu entwerfen und auszuführen, die zum Erreichen der folgenden Ziele dienen:

- **Profilieren von Daten.** Profiling macht den Inhalt und die Struktur Ihrer Daten erkennbar. Profiling ist ein wichtiger Schritt in jedem Datenprojekt, da es Stärken und Schwächen in Ihren Daten identifizieren kann und Sie bei der Definition Ihres Projektplans unterstützt.
- **Erstellen von Scorecards,** um die Datenqualität zu prüfen. Eine Scorecard ist eine grafische Darstellung der Qualitätsmessungen in einem Profil.
- **Standardisieren von Datenwerten.** Standardisieren Sie Daten, um Fehler und Inkonsistenzen zu entfernen, die Sie finden, indem Sie ein Profil ausführen. Sie können Variationen der Zeichensetzung, Formatierung und Rechtschreibung standardisieren. Sie können beispielsweise sicherstellen, dass die Werte für Stadt, Staat und Postleitzahl konsistent sind.
- **Parsen von Datensätzen.** Parsen Sie Datensätze, um die Datensatzstruktur zu verbessern und weitere Informationen aus Ihren Daten abzuleiten. Sie können ein Einzelfeld mit Freiformdaten in Bereiche teilen, die unterschiedliche Informationstypen enthalten. Sie können Ihren Datensätzen auch Informationen hinzufügen. Sie können beispielsweise Datensätze von Kunden als persönliche oder Geschäftskunden markieren.
- **Validieren von Postadressen.** Die Adressvalidierung bewertet und verbessert die Genauigkeit und Zustellbarkeit Ihrer Postadressendaten. Die Adressvalidierung korrigiert Fehler in Adressen und komplettiert unvollständige Adressen, indem die Adressdatensätze mit Referenzdaten von nationalen Zustellern verglichen werden. Die Adressvalidierung kann auch Postinformationen hinzufügen, die die Postzustellung beschleunigen und die Zustellkosten reduzieren.
- **Nach doppelten Datensätzen suchen.** Die Analyse auf doppelte Datensätze vergleicht einen Satz von Datensätzen miteinander, um ähnliche oder übereinstimmende Werte in ausgewählten Datenspalten zu finden. Das Ähnlichkeitsniveau, welches eine gute Ähnlichkeit zwischen den Bereichswerten anzeigt, wird von Ihnen vorgegeben. Sie können auch die relative Gewichtung vorgeben, die jede Spalte in den Ähnlichkeitsberechnungen erhält. Sie können beispielsweise dem Nachnamen eine höhere Priorität als dem Vornamen geben.
- **Erstellen und Durchführen von Datenqualitätsregeln.** Informatica bietet vorgefertigte Regeln, die Sie ausführen oder bearbeiten können, um Ihre Projektziele zu erreichen. Sie können im Entwicklertool Regeln erstellen.
- **Zusammenarbeiten mit Informatica-Anwendern.** Die Regeln und Referenzdatentabellen, die Sie dem Model Repository zufügen, stehen den Anwendern im Entwicklertool und im Analysetool zur Verfügung. Benutzer können in Projekten zusammenarbeiten, und in unterschiedlichen Stadien eines Projekts können unterschiedliche Benutzer den Eigentümerstatus von Objekten übernehmen.
- **Exportieren von Mappings an PowerCenter.** Sie können Mappings an PowerCenter exportieren, um die Metadaten für die physische Datenintegration oder für die Erstellung von Internetdiensten zu verwenden.

Integrations-Installer

Wenn Sie unterschiedliche Versionen von Data Quality und PowerCenter verwenden, um Objekte zu erstellen und zu speichern, installieren Sie das Informatica Data Quality-Plug-In für PowerCenter. Laden Sie die Plug-In-Installationsprogramme von Informatica herunter.

Installieren Sie die Plug-In-Version für die Version von PowerCenter, die Sie verwenden möchten. Für das Plug-In gibt es Client- und Server-Installer.

Führen Sie das Client-Installationsprogramm auf dem PowerCenter-Client-Computer aus, um Mapplets und Mappings im Repository anzuzeigen und diese zu Sitzungen hinzuzufügen. Führen Sie den Server-Installer auf dem PowerCenter Integration Service-Computer aus, um die Sitzungen in Arbeitsabläufen auszuführen.

Nachdem Sie das Plug-In auf dem Rechner für den Integration Service installiert haben, registrieren Sie das Plug-In beim PowerCenter Repository.

Export eines Objekts nach PowerCenter

Benutzen Sie das Developer Tool, um Objekte wie Mapplets und Mappings aus dem Informatica Modellrepository zu exportieren. Sie können Objekte vom Modellrepository in das PowerCenter-Repository exportieren und Sie können Objekte vom Modellrepository ins Dateisystem exportieren. Benutzen Sie den PowerCenter Repository Manager zum Importieren von XML-Objekten in das PowerCenter Repository.

- Weitere Informationen zum Exportieren von Objekten aus dem Modellrepository finden Sie im *Informatica Developer-Handbuch für Mapping* und im *Informatica Developer Tool-Handbuch*.
- Informationen zum Importieren von Objekten in das PowerCenter Repository finden Sie im *PowerCenter Repository-Handbuch*.

Installation von Referenzdaten

Wenn ein Mapplet oder Mapping Referenzdaten-Dateien liest, installieren Sie die Referenzdaten-Dateien auf dem PowerCenter Integration Service-Computer.

Sie können das Developer-Tool verwenden, um Referenzdaten aus dem Modellrepository zu exportieren, wenn Sie ein Mapplet oder Mapping exportieren. Der Developer exportiert die Referenzdaten als komprimierte Datei. Extrahieren Sie die Datei und kopieren Sie die Referenzdaten auf den Integration Service-Computer.

Außerdem können Sie den Data Quality Content Installer verwenden, um dem Integration Service-Computer Daten hinzuzufügen. Führen Sie den Content Installer aus, um Adressenreferenzdaten-Dateien und Identitätspopulationsdatendateien zu installieren.

Integrationsmatrix

Der Grad der Integration zwischen Data Quality und PowerCenter hängt von den jeweiligen Produktversionen ab. Sie können Objekte und Referenzdaten aus mehreren Versionen von Data Quality in mehrere Versionen

von PowerCenter exportieren. Sie können keine Objekte oder Referenzdaten von PowerCenter nach Data Quality exportieren.

In der folgenden Tabelle werden die Produktversionen aufgelistet, die Sie integrieren können:

Produktname	PowerCenter 10.1.1 HotFix 2	PowerCenter 10.1.1 HotFix 1	PowerCenter 10.1.1	PowerCenter 10.1	PowerCenter 10.0	PowerCenter 9.6.1
Data Quality 10.1.1 HotFix 2	Ja	Nein	Nein	Installieren von Plug-Ins	Nein	Installieren von Plug-Ins
Data Quality 10.1.1 HotFix 1	Nein	Ja	Nein	Installieren von Plug-Ins	Nein	Installieren von Plug-Ins
Data Quality 10.1.1	Nein	Nein	Ja	Installieren von Plug-Ins	Nein	Installieren von Plug-Ins
Data Quality 10.1	Nein	Nein	Nein	Ja	Nein	Installieren von Plug-Ins
Data Quality 10.0	Nein	Nein	Nein	Nein	Ja	Installieren von Plug-Ins
Data Quality 9.6.1	Nein	Nein	Nein	Nein	Nein	Ja

Die Tabellenwerte stellen die folgenden Integrationstypen dar:

Ja

Sie können Objekte und Dateien aus einer Data Quality-Umgebung in eine PowerCenter-Umgebung exportieren. Sie müssen kein Plug-In installieren oder registrieren. Sie können Datenqualitätsobjekte in einem PowerCenter-Repository ausführen, das Sie auf die angegebene Version aktualisieren.

Installieren von Plug-Ins

Sie können Objekte und Dateien aus einer Data Quality-Umgebung in eine PowerCenter-Umgebung exportieren. Installieren und registrieren Sie die Client- und Server-Plug-Ins für die Version von PowerCenter, die Sie verwenden. Sie können Datenqualitätsobjekte in einem PowerCenter-Repository ausführen, das Sie auf die angegebene Version aktualisieren.

Nein

Sie können die in einer Data Quality-Umgebung angegebenen Objekte und Dateien nicht in einer PowerCenter-Umgebung verwenden.

Regeln und Richtlinien für die Produktintegration

Sie können Datenqualitätsobjekte aus verschiedenen Versionen von Informatica Data Quality in verschiedene Versionen von PowerCenter exportieren. Installieren Sie das Integrations-Plug-In für die Version von Informatica Data Quality, aus der die Objekte exportiert werden.

Berücksichtigen Sie die folgenden Regeln und Richtlinien beim Integrieren von Informatica Data Quality und PowerCenter:

- Installieren Sie das Integrations-Plug-In nicht, wenn die PowerCenter-Repository-Produktversion und die Modellrepository-Produktversion übereinstimmen.
- Sie können nicht zwei Data Quality-Integrations-Plug-Ins gleichzeitig auf einem einzelnen PowerCenter-Computer ausführen. PowerCenter verwendet das Plug-In, das Sie zuletzt installiert haben. Wenn Sie zum Beispiel das Data Quality 9.6.1-Integrations-Plug-In auf einem PowerCenter 9.5.1-Computer installieren, können Sie keine Datenqualitätsobjekte mehr aus einem Modellrepository der Version 9.5.1 importieren und ausführen. Sie können weiterhin alle in das PowerCenter-Repository importierten Datenqualitätsobjekte ausführen, bevor Sie das 9.6.1-Plug-In installieren.
- Sie können keine Datenqualitäts-Mapplets oder -Mappings in einem PowerCenter-Repository bearbeiten.

Verhalten bei Datenqualitätsumwandlungen in PowerCenter

Wenn Sie ein Datenqualitäts-Mapplet oder Mapping aus dem Modellrepository in das PowerCenter-Repository importieren, kann sich das Verhalten bei Datenqualitätsumwandlungen ändern.

Hinweise zur Konsolidierungsumwandlung

Informatica Data Quality hat die Standardstrategie der Konsolidierungsumwandlung in Version 9.5.1 geändert. In Version 9.5.1 und späteren Versionen wählt die Konsolidierungsumwandlung den Datensatz mit der höchsten Zeilen-ID als Survivor-Datensatz aus. In Version 9.5 und früheren Versionen verwendet die Konsolidierungsumwandlung den häufigsten nicht leeren Wert in einer Spalte, die Sie zum Identifizieren des Survivor-Datensatzes angeben.

Beim Aktualisieren eines PowerCenter-Repository mit einer Konsolidierungsumwandlung auf Version 9.5.1 oder höher behält das Upgrade den häufigsten nicht leeren Wert als Standardeinstellung bei. Wenn Sie ein Mapping mit einer Konsolidierungsumwandlung importieren, in der die Standardeinstellungen eines Modellrepository aus Version 9.5.1 oder ein höheres Repository verwendet werden, greift die Umwandlung zur Identifizierung des Survivor-Datensatzes auf die höchste Zeilen-ID zurück.

Konvertierung in Mapplets

Die folgenden Umwandlungen können beim Import in das PowerCenter-Repository in Mapplets konvertiert werden:

- Konsolidierung
- Ausnahme
- Schlüsselgenerator
- Übereinstimmung

Wenn Sie ein Mapping importieren, das eine oder mehrere der Umwandlungen enthält, konvertiert PowerCenter jede Umwandlung in ein Mapplet, das die erweiterte Konfiguration enthält.

Wenn Sie ein Mapplet importieren, das eine oder mehrere der Umwandlungen enthält, dann erweitert PowerCenter die Konfiguration jeder Umwandlung in mehrere Umwandlungen.

PowerCenter fügt möglicherweise Umwandlungen zu der Datenqualitätsumwandlung oder Mapplet-Erweiterung hinzu, um die Umwandlungslogik zu vervollständigen. Zum Beispiel erweitert PowerCenter die Match-Umwandlung mit zusätzlichen Umwandlungen.

Erwägungen bei Entscheidungsumwandlungen

Sie können eine Zuordnung oder ein Mapplet mit einer Entscheidungsumwandlung nicht in PowerCenter exportieren, wenn das Umwandlungsskript einen Systemparameter enthält. Der Systemparameter kann während des Exports nicht in einen Wert umgewandelt werden, der von PowerCenter verwendet werden kann. Bevor Sie eine Zuordnung oder ein Mapplet mit einer Entscheidungsumwandlung exportieren, die einen Systemparameter verwendet, ersetzen Sie den Parameter durch einen geeigneten Wert.

Hinweise zu Ausnahmen bei der Umwandlung

PowerCenter kann keine Ausnahmeumwandlung verwenden, die aus einem PowerCenter-Repository stammt, welches älter als Version 9.5 ist. Wenn Sie ein PowerCenter-Repository auf Version 9.5 oder höher aktualisieren und das Repository ein Mapping mit einer Ausnahmeumwandlung enthält, exportieren Sie das Mapping erneut. Erstellen Sie das Mapping bei Bedarf erneut im Developer Tool.

Beim Ausführen eines Mappings, das eine Ausnahmeumwandlung enthält, können Sie die Ausgabebetabellen der fehlerhaften oder duplizierten Datensätze im Analyst Tool aktualisieren. Wenn Sie beabsichtigen, eine Tabelle fehlerhafter Datensätze im Analyst Tool zu aktualisieren, müssen Sie für die Sitzung, in der das Mapping ausgeführt wird, die hohe Genauigkeit aktivieren.

Hinweise zur Match-Umwandlung

Sie können kein Identitätsabgleichs-Mapping in PowerCenter ausführen, wenn das Mapping Identitätsindexdaten in Datenbanktabellen liest oder schreibt. Um eine Identitätsvergleichsanalyse in PowerCenter durchzuführen, konfigurieren Sie die Match-Umwandlung im Developer Tool, um Identitätsindexdaten in temporäre Dateien zu schreiben.

Mehrfachstrategieumwandlungen

Sie können innerhalb des Developer Tools in jeder der folgenden Umwandlungen mehrere Datenumwandlungsstrategien definieren:

- Groß-/Kleinschreibung
- Klassifizierer
- Entscheidung
- Schlüsselgenerator
- Beschriftung
- Übereinstimmung
- Zusammenführung
- Parser
- Standardisierer

Wenn Sie ein Mapplet oder Mapping importieren, das eine Mehrfachstrategieumwandlung enthält, wird jede Strategie in eine separate Umwandlung konvertiert.

Partitioneinstellungen

Bevor Sie eine Sitzungsaufgabe mit einem Datenqualitäts-Mapping in einem Gitter ausführen, identifizieren Sie die PowerCenter-Version, in der Sie die Gitterknoten erstellt haben. Wenn Sie die Knoten in PowerCenter 9.5.1 HotFix 2 oder einer früheren Version von PowerCenter erstellt haben, schlägt das Mapping möglicherweise fehl.

Um sicherzustellen, dass Datenqualitäts-Mappings auf allen Knoten des Gitters ausgeführt werden können, führen Sie eine der folgenden Aufgaben durch:

- Überprüfen Sie die erweiterten Eigenschaften in allen PowerCenter-Integrationsdiensten, die Aufgaben an Knoten im Gitter verteilen. Konfigurieren Sie jeden PowerCenter-Integrationsdienst so, dass Ressourcenanforderungen beim Verteilen von Aufgaben ignoriert werden. Standardmäßig ist in den erweiterten Eigenschaften festgelegt, dass der PowerCenter-Integrationsdienst Ressourcenanforderungen ignoriert.
- Erstellen Sie alle Knoten neu, die Sie in PowerCenter 9.5.1 HotFix 2 oder einer früheren PowerCenter-Version erstellt haben. Wenn Sie den Knoten neu erstellen, wendet PowerCenter die aktuelle Knotenkonfiguration auf den Knoten an.

Wenn Sie eine Sitzung auf einem Gitter ausführen, dann verteilt der DTM-Prozess (Data Transformation Manager) die Threads auf die Knoten des Gitters. Manche Umwandlungen können nicht über ein Gitter aufgeteilt werden. Wenn eine Umwandlung nicht über ein Gitter aufgeteilt werden kann, erzeugt der DTM eine einzelne Partitionsgruppe für die Umwandlungs-Threads und führt die Threads auf einem einzelnen Knoten aus.

Die folgende Tabelle beschreibt die Partitionseinstellungen für Datenqualitätsumwandlungen:

Umwandlung	Partitionseinstellung
Adressvalidierung	Über Gitter
Zuordnung	Nicht partitionierbar
AV-Bericht	Lokal partitionierbar
Groß-/ Kleinschreibungsumwandler	Über Gitter
Klassifizierer	Lokal partitionierbar
Cluster	Nicht partitionierbar
Vergleich	Über Gitter
Konsolidierung	Nicht partitionierbar
Entscheidung	Über Gitter
Ausnahme	Nicht partitionierbar
Ausführungspunkt	Über Gitter
Schlüsselgenerator	Nicht partitionierbar
Schlüsselspeicher	Nicht partitionierbar
Beschriftung	Über Gitter
Zusammenführung	Über Gitter
Paargenerator	Nicht partitionierbar
Parser (Probabilistische Analyse)	Lokal partitionierbar

Umwandlung	Partitionseinstellung
Parser	Über Gitter
Token-Standardisierer	Über Gitter
Gewichtungsbasierter Analyzer	Über Gitter

KAPITEL 2

Installation

Dieses Kapitel umfasst die folgenden Themen:

- [Installation - Übersicht, 14](#)
- [Systemvoraussetzungen, 14](#)
- [Vor dem Installieren, 15](#)
- [Installieren des Server-Plug-Ins, 15](#)
- [Installieren des Client-Plug-Ins, 16](#)
- [Registrierung des Plug-Ins für PowerCenter, 17](#)

Installation - Übersicht

Das Plug-In der Data Quality-Integration verfügt über eine Server- und eine Clientkomponente.

Führen Sie die folgenden Schritte zum Installieren und Registrieren des Plug-Ins aus:

1. Führen Sie das Server-Installationsprogramm auf dem PowerCenter-Integrationsdienst-Computer aus.
2. Führen Sie das Clientinstallationsprogramm auf dem PowerCenter-Clientcomputer aus.
3. Registrieren Sie die Datei `IDQZIntegration.xml` beim PowerCenter-Repository.

Systemvoraussetzungen

Das Informatica Data Quality Integrations-Plugin hat die gleichen Systemvoraussetzungen wie die PowerCenter-Anwendungen und Services, für die es installiert wird.

Informationen über Voraussetzungen und Systemanforderungen für PowerCenter finden Sie in der PowerCenter Installationsdokumentation.

Vor dem Installieren

Berücksichtigen Sie das folgende Installationsprogramm-Verhalten bei der Verwendung von Referenzdaten im PowerCenter.

Das Installationsprogramm des Integration-Plug-in speichert die Eigenschaftsdateien auf dem Computer, auf dem der PowerCenter-Integrationsdienst ausgeführt wird. Die Eigenschaftsdateien geben den Speicherort der Referenzdatendateien an, die von Data Quality-Mappings gelesen werden können.

Das Installationsprogramm schreibt die folgenden Dateien:

- AD50.cfg. Speichert die Konfigurationseigenschaften für Adressenreferenzdaten.
- CLASSIFIER.properties. Speichert die Konfigurationseigenschaften für Klassifizierer-Modelldaten.
- IDQTx.cfg. Speichert die Konfigurationseigenschaften für Daten von Identitätspopulationen.
- NER.properties. Speichert die Konfigurationseigenschaften für probabilistische Modelldaten.

Das Installationsprogramm fügt die Dateien in folgendes Verzeichnis ein:

```
[PowerCenter_installation]/server/bin
```

Wenn das Installationsprogramm eine Eigenschaftsdatei im Verzeichnis `bin` findet, erstellt es davon eine Sicherungskopie und installiert eine neue Datei. Um die aktuelle Referenzdatenkonfiguration nach der Installation beizubehalten, führen Sie die Inhalte der alten und neuen Dateien zusammen.

Bei der Installation von PowerCenter 9.5.1 oder einer neueren Version werden im Laufe der Serverinstallation die Eigenschaftsdateien standardmäßig im Verzeichnis `bin` gespeichert.

Erforderliche Vorgangsweise bei Adressreferenzdaten

Vor dem Installieren der Adressreferenzdaten muss der PowerCenter-Integrationsdienst angehalten werden. Starten Sie den Service nach der Installation der Daten neu.

Installieren des Server-Plug-Ins

Der Informatica Data Quality Integration Server Plug-In-Installer fügt Anwendungsdateien für die Datenqualitätsumwandlungen zur PowerCenter Integration Service-Verzeichnisstruktur hinzu.

Installieren unter Windows

Führen Sie den Windows Server Installer aus, um das Plug-In auf einem PowerCenter Integration Service-Rechner mit Windows-Plattform hinzuzufügen.

Schließen Sie vor der Installation alle Anwendungen.

1. Melden Sie sich auf dem Computer mit einem System-Benutzerkonto an.
2. Entpacken Sie die Integration Server-Datei für Windows.
3. Gehen Sie zu dem Verzeichnis, in das Sie die Dateien entpackt haben, und starten Sie `install.exe`.
4. Wählen Sie eine Sprache aus und klicken Sie auf **OK**.
5. Auf dem **Begrüßungsbildschirm** klicken Sie auf **Weiter**.
6. Stellen Sie sicher, dass die PowerCenter-Dienste nicht laufen, und klicken Sie auf **OK**.
7. Geben Sie den Pfad für das PowerCenter-Installationsverzeichnis ein. Der Installer gibt für dieses Verzeichnis einen Standardpfad vor. Um einen anderen Pfad zu wählen, klicken Sie auf **Wählen**.

8. Klicken Sie auf **Weiter**.
9. Prüfen Sie die Vorinstallations-Zusammenfassung. Klicken Sie auf **Zurück**, um eventuelle Fehler zu korrigieren.
10. Klicken Sie auf **Installieren**.
11. Prüfen Sie die Nachinstallations-Zusammenfassung.
Für weitere Informationen zu Installationsaufgaben und zum Ansehen der Konfigurationseigenschaften für die installierten Komponenten sehen Sie sich die Installations-Logdateien an.
12. Klicken Sie auf **Fertig**.

Installieren unter UNIX

Führen Sie den UNIX Server Installer aus, um das Plug-In auf einem PowerCenter Integration Service-Rechner mit UNIX-Plattform hinzuzufügen.

1. Melden Sie sich auf dem Computer mit einem System-Benutzerkonto an.
2. Schließen Sie alle anderen Anwendungen.
3. Entpacken Sie die Integration Server-Datei für UNIX.
4. Öffnen Sie eine UNIX-Shell und begeben Sie sich in das Verzeichnis, in das Sie die Installationsdateien entpackt haben.
5. Geben Sie `sh install.bin` ein.
6. Um Englisch als Installationssprache zu akzeptieren, drücken Sie **Eingabe**.
7. Lesen Sie den Begrüßungstext und drücken Sie **Eingabe**.
8. Geben Sie einen absoluten Pfad zum PowerCenter Installationsverzeichnis ein, oder drücken Sie **Eingabe**, um den Standardpfad zu akzeptieren.
9. Prüfen Sie die Vorinstallations-Zusammenfassung. Um eventuelle Fehler zu korrigieren, geben Sie `zurück` ein und drücken Sie **Eingabe**.
10. Drücken Sie **Eingabe**, um mit der Installation zu beginnen.
11. Prüfen Sie die Nachinstallations-Zusammenfassung.
Für weitere Informationen zu Installationsaufgaben und zum Ansehen der Konfigurationseigenschaften für die installierten Komponenten sehen Sie sich die Installations-Logdateien an.
12. Drücken Sie **Eingabe**, um den Installationsprozess fertigzustellen.

Installieren des Client-Plug-Ins

Führen Sie den Data Quality Integration Client-Installer auf PowerCenter Client-Rechnern aus.

1. Schließen Sie alle anderen Anwendungen.
2. Entpacken Sie die Integration Client-Datei.
3. Öffnen Sie das Verzeichnis, in dem Sie die Dateien entpackt haben.
4. Starten Sie `install.exe`.
5. Auf dem **Begrüßungsbildschirm** klicken Sie auf **Weiter**.
Der Installer fordert Sie auf, PowerCenter Client-Anwendungen zu schließen, bevor Sie mit der Installation fortfahren.

6. Stellen Sie sicher, dass keine Client-Anwendungen laufen, und klicken Sie **OK**.
7. Geben Sie den Pfad für das PowerCenter Client-Verzeichnis ein. Der Installer gibt für dieses Verzeichnis einen Standardpfad vor. Um einen anderen Pfad zu wählen, klicken Sie auf **Wählen**.
8. Klicken Sie auf **Weiter**.
9. Prüfen Sie die Vorinstallations-Zusammenfassung. Klicken Sie auf **Zurück**, um eventuelle Fehler zu korrigieren.
10. Klicken Sie auf **Installieren**.
11. Prüfen Sie die Nachinstallations-Zusammenfassung.
Für weitere Informationen zu Installationsaufgaben und zum Ansehen der Konfigurationseigenschaften für die installierten Komponenten sehen Sie sich die Installations-Logdateien an.
12. Klicken Sie auf **Fertig**.

Registrierung des Plug-Ins für PowerCenter

Registrieren Sie das Plug-In für jedes PowerCenter Repository, das Datenqualitätsobjekte enthalten soll.

Der Integrations-Installer schreibt die Plug-In-Datei `IDQZIntegration.xml` in das Verzeichnis `$INFA_HOME\server\bin\native` auf dem Rechner des PowerCenter-Integrationsdiensts. Vergewissern Sie sich, dass Sie auf die Datei zugreifen können. Falls nötig, kopieren Sie die Datei oder übertragen Sie sie per FTP auf den Rechner des PowerCenter-Integrationsdiensts.

- Führen Sie zum Registrieren des Plug-Ins den Befehl `pmrep RegisterPlugin` aus. Schließen Sie beim Ausführen des Befehls die Option `-N` ein.

Weitere Informationen über den Befehl `pmrep RegisterPlugin` finden Sie in der *Informatica-Befehlsreferenz*.

Hinweis: Falls Sie nicht über die erforderlichen Berechtigungen zum Registrieren des Plug-Ins verfügen, kontaktieren Sie bitte den Administrator des PowerCenter-Repository-Diensts.

KAPITEL 3

Referenzdaten-Anforderungen

Dieses Kapitel umfasst die folgenden Themen:

- [Referenzdaten - Übersicht, 18](#)
- [Anforderungen für Referenzdaten-Dateien, 19](#)
- [Anforderungen für Adressenreferenzdaten, 20](#)
- [Anforderungen an Identitätspopulationsdateien, 21](#)
- [Anforderungen an probabilistische und Klassifizierer-Modelle, 23](#)

Referenzdaten - Übersicht

Datenqualitätsumwandlungen können Referenzdaten lesen. Wenn eine Umwandlung im PowerCenter Repository Referenzdaten liest, stellen Sie sicher, dass die Referenzdaten in der PowerCenter-Umgebung vorhanden sind.

Datenqualitätsumwandlungen in PowerCenter können die folgenden Referenzdatentypen verwenden:

Referenzdaten-Dateien

Textdateien, die Standard- und alternative Versionen der Geschäftsbedingungen enthalten. Sie können Referenzdaten-Dateien bearbeiten.

Adressenreferenzdaten-Dateien

Datendateien, die vollständige Daten für alle Postadressen in einem Land enthalten. Sie können Adressenreferenzdaten-Dateien auf dem Rechner für den PowerCenter Integration Service installieren. Sie können Referenzdaten-Dateien nicht bearbeiten.

Identitätspopulationsdateien

Datendateien, die verschiedene Identitätstypen definieren und Algorithmen enthalten, die von der Übereinstimmungsumwandlung verwendet werden. Populationsdateien werden auf dem PowerCenter Servicerechner installiert. Sie können Populationsdateien nicht bearbeiten.

Inhaltssätze

Dateien, die Referenzdaten enthalten, welche durch eine oder mehrere Umwandlungen spezifiziert werden. Wenn Sie ein Mapping exportieren, das Inhalte liest, exportiert der Exportprozess die Inhaltsdaten mit den Umwandlungsmetadaten. Im Falle von Klassifizierer- und probabilistischen Modelldateien exportiert der Prozess die Dateidaten aus der Data Quality-Verzeichnisstruktur.

Anforderungen für Referenzdaten-Dateien

Ein Data Quality-Mapping kann Referenzdatenobjekte lesen, die im Modellrepository definiert sind.

Wenn ein Informatica Data Quality-Benutzer ein Mapping exportiert, das ein Referenzdatenobjekt liest, schreibt der Exportprozess die Referenzdaten als eine oder mehrere Referenzdaten-Dateien in das Dateisystem. Wenn Sie das Mapping in PowerCenter verwenden, müssen Sie sicherstellen, dass die Dateien an einem Speicherort installiert sind, den der Integration Service lesen kann.

Legen Sie den Speicherort für die Referenzdaten-Dateien fest, wenn Sie das Mapping exportieren. Exportieren Sie die Dateien in die PowerCenter-Verzeichnisstruktur oder exportieren Sie die Dateien in das Dateisystem und kopieren Sie die Dateien auf den PowerCenter Integration Service-Computer.

Die Speicherorte für die Referenzdaten-Dateien in der PowerCenter-Verzeichnisstruktur müssen den Referenztabelle-Speicherorten in der Modellrepository-Struktur entsprechen. Die Verzeichnisstruktur muss dem folgenden Format entsprechen:

```
<PowerCenter_Root_Directory>\services\<Model_Repository_Project_Name>\<Model_Repository_Project_Folder_Name>
```

Hinweis: PowerCenter liest Referenzdaten-Dateien im Format UTF-8. Das Developer-Tool exportiert Referenzdaten für PowerCenter im UTF-8 Format.

Installieren von Referenzdaten aus dem Data Quality Content Installer

Die Data Quality Content Installer-Dateien enthalten Referenzdaten-Dateien, die es in die PowerCenter-Verzeichnisstruktur schreibt.

Legen Sie das PowerCenter-Installationsverzeichnis beim Ausführen des Content Installer fest. Standardmäßig fügt der Content Installer den folgenden Pfad dem Stammverzeichnis hinzu:

```
\services\DQContent\INFA_Content\
```

Bearbeiten Sie den Pfad zur Modellrepository-Struktur oder kopieren Sie die Referenzdaten-Dateien in das erforderliche Verzeichnis, nachdem Sie den Content Installer ausgeführt haben.

Verwenden einer Umgebungsvariable zum Definieren des Datenspeicherorts für die Referenzdaten

Sie können eine Umgebungsvariable mit dem Namen *INFA_CONTENT* auf dem PowerCenter Integration Service-Computer definieren, um den vom Integration Service benutzten Pfad zum Lesen der Referenzdaten-Dateien festzulegen.

Verwenden Sie diese Umgebungsvariable, wenn Sie die Referenzdaten-Dateien nicht im erforderlichen Verzeichnis auf dem PowerCenter Integration Service-Computer installieren können. Installieren Sie die Dateien an einem Speicherort, auf den der PowerCenter Integration Service zugreifen kann, und erzeugen Sie die *INFA_CONTENT*-Umgebungsvariable mit dem Installationspfad.

Anforderungen für Adressenreferenzdaten

Ein Mapping, das Adressvalidierungen ausführt, liest Adressenreferenzdaten. Sie können Adressenreferenzdaten von Informatica kaufen und herunterladen. Kopieren Sie die Daten auf den PowerCenter Integration Service-Computer.

Die folgende Tabelle beschreibt die Arten von Adressenreferenzdaten, die Sie installieren können:

Arten von Referenzdaten	Wann verwenden
Batch	Für allgemeine Adressvalidierungsvorgänge.
CAMEO	Zum Hinzufügen von demographischen Verbraucherdaten zu den Adressdatensätzen.
Zertifiziert	<p>Zum Sicherstellen, dass die Adressdatensätze den Zertifizierungsstandards entsprechen, die ein Postbeförderungsunternehmen definiert.</p> <p>Die folgenden Länder definieren Zertifizierungsstandards:</p> <ul style="list-style-type: none">- Australien. Zertifiziert die Post gemäß dem Standard des Address Matching Approval System (AMAS).- Kanada. Zertifiziert die Post gemäß dem Standard des Software Evaluation And Recognition Program (SERP).- Frankreich. Zertifiziert die Post gemäß dem Standard des National Address Management Service (SNA).- Neuseeland. Zertifiziert die Post gemäß dem SendRight-Standard.- Vereinigte Staaten von Amerika. Zertifiziert die Post gemäß dem Standard des Coding Accuracy Support System (CASS).
Vorschlagsliste	Zum Suchen nach alternativen gültigen Versionen eines unvollständigen Adressdatensatzes. Verwenden Sie die Vorschlagslistendaten, wenn Sie ein Adressvalidierungs-Mapping zu einzelnen Adressdatensätzen in Echtzeit konfigurieren.
Geocode	Zum Hinzufügen der Längen- und Breitengradkoordinaten zu den Adressdatensätzen.
Zusätzlich	<p>Zum Identifizieren des geografischen Gebiets oder Bevölkerungsgebiets, das eine Adresse enthält.</p> <p>Die Adressvalidator-Umwandlung kann zusätzliche Daten zu Adressdatensätzen aus den folgenden Ländern hinzufügen:</p> <ul style="list-style-type: none">- Japan. Identifiziert den städtischen Bezirk, der die Adresse enthält.- Serbien. Identifiziert die Postleitzahl auf Straßenebene für die Adresse.- Vereinigtes Königreich. Fügt eine Zustellpunkt-ID zur Adresse hinzu.- Vereinigte Staaten von Amerika. Identifiziert das Zentrum der Population, das die Adresse enthält.

Außerdem können Sie den Data Quality Content Installer verwenden, um die Adressenreferenzdaten zu installieren. Führen Sie den Server Content Installer auf dem PowerCenter Integration Service-Computer aus, um die Dateien zu installieren.

Der Data Quality-Inhalts-Installer erzeugt ein Verzeichnis `/av/` in der PowerCenter-Verzeichnisstruktur und schreibt Adressenreferenzdaten in dieses Verzeichnis. Der PowerCenter Integration Service liest den Pfad zu den Adressenreferenzdaten aus der Konfigurationsdatei `AD50.cfg`.

Das PowerCenter-Installationsprogramm schreibt die Datei `AD50.cfg` in das folgende Verzeichnis:

```
[PowerCenter_Installation]/server/bin
```

Wenn Sie Adressenreferenzdaten installieren, stellen Sie sicher, dass die folgenden Eigenschaften in AD50.cfg konfiguriert sind:

Referenzdaten-Speicherort

Geben Sie den Pfad zu den Adressenreferenzdaten bei der Eigenschaft `ReferenceDataLocation` ein.

Lizenzschlüssel

Geben Sie die Lizenzdaten bei der Eigenschaft `LicenseKey` ein. Falls Sie mehr als eine Lizenz besitzen, geben Sie die Lizenzen in Form einer kommagetrennten Zeichenketten ein.

Preload-Werte

Geben Sie mindestens ein Länderkürzel als Preload-Wert für jeden Referenzdatentyp ein, den ein Mapping einliest. Geben Sie ALL ein, um eine Preload-Einstellung auf alle Länder anzuwenden.

Der Integration Service kann für jedes Land eine andere Methode zum Laden der Daten verwenden. Sie können zum Beispiel für die Vereinigten Staaten ein volles Preload für Batch- bzw. interaktive Daten und für Großbritannien ein partielles Preload für Batch- bzw. interaktive Daten angeben. Der Integration Service kann auch für jeden Datentyp eine andere Preload-Methode verwenden. Sie können zum Beispiel für die Vereinigten Staaten ein volles Preload für Batch- bzw. interaktive Daten und ein partielles Preload für GeoCoding-Daten angeben.

Die Einstellungen für volle Preloads haben Vorrang vor den Einstellungen für partielle Preloads, und partielle Preload-Einstellungen haben Vorrang vor Einstellungen, bei denen kein Preload angegeben ist. Wenn Sie zum Beispiel ALL für Kein Daten-Preload und USA für Vollen Preload eingeben, lädt der Integration Service alle Daten für die Vereinigten Staaten in den Speicher und keine Daten für irgendein anderes Land. Wenn Sie keinen Preload benötigen, dann geben Sie für alle Referenzdatentypen, die Sie verwenden möchten, ALL für Kein Daten-Preload ein.

Für ergänzende Daten wird kein Preload-Wert eingegeben.

Sie können optional auch andere Eigenschaften in AD50.cfg einstellen.

Hinweis: Informatica Data Quality 9.x liest die Datei AD50.cfg nicht. Informatica Data Quality 9.x Benutzer müssen den Referenzdaten-Speicherort, die Lizenzschlüsselinformation sowie die Preload-Werte über das Administrator Tool eingeben.

Anforderungen an Identitätspopulationsdateien

Ein Mapping, das eine Analyse zum Identitätsvergleich durchführt, liest Referenzdaten-Dateien, genannt „Identitätspopulationen“. Sie können Dateien mit Identitätspopulationsdateien von Informatica kaufen und herunterladen. Kopieren Sie die Populationsdateien auf den PowerCenter-Integrationsdienstcomputer oder in ein gemeinsam genutztes Verzeichnis auf dem Computer, auf den der PowerCenter-Integrationsdienst zugreifen kann.

Außerdem können Sie den Data Quality Content Installer verwenden, um die Dateien mit Populationen zu installieren. Führen Sie den Server Content Installer auf dem PowerCenter-Integrationsdienst-Computer aus, um die Dateien zu installieren.

Sie müssen die Populationsdateien in ein Verzeichnis mit dem Namen `Standard` kopieren oder installieren.

Der PowerCenter-Integrationsdienst kann den Speicherort des Populationsdateiverzeichnisses aus der Konfigurationsdatei `IDQTx.cfg` oder aus der Umgebungsvariable `SSAPR` lesen. Der PowerCenter-Integrationsdienst sucht nach dem Speicherort in der Datei `IDQTx.cfg`. Wenn die Datei `IDQTx.cfg` nicht vorhanden ist oder keinen Speicherort angibt, sucht der PowerCenter-Integrationsdienst einen Speicherort in der `SSAPR`-Umgebungsvariable.

Hinweis: Die Datei IDQTx.cfg und die SSAPR-Umgebungsvariable geben den Pfad zum übergeordneten Verzeichnis des Standard-Verzeichnisses an. Der Pfad enthält nicht den Namen des Verzeichnisses.

Das PowerCenter-Installationsprogramm schreibt die Datei idqtx.cfg in das folgende Verzeichnis:

```
[PowerCenter_Installation]/server/bin
```

Die Datei IDQTx.cfg verwendet die Eigenschaft *IdentityReferenceDataLocation*, um den Speicherort des Populationsdateiverzeichnisses anzugeben. Aktualisieren Sie die Eigenschaft mit dem aktuellen Speicherort des Verzeichnisses.

Sie legen beispielsweise den folgenden Pfad in der Eigenschaft *IdentityReferenceDataLocation* fest:

```
[PowerCenter_Installation]/services/DQContent/INFA_Content/identity/
```

Sie können einen absoluten oder einen relativen Pfad zum Standardverzeichnis in der Eigenschaft *IdentityReferenceDataLocation* festlegen. Legen Sie einen absoluten Pfad zum Verzeichnis auf dem PowerCenter-Integrationsdienstcomputer oder zu einem gemeinsam genutzten Verzeichnis auf einem Computer fest, auf den der Dienst zugreifen kann. Legen Sie einen relativen Pfad zu einem Verzeichnis auf dem PowerCenter-Integrationsdienstcomputer fest.

Der PowerCenter-Integrationsdienst liest einen relativen Pfad aus dem folgenden Verzeichnis:

```
[PowerCenter_Installation]/server/bin
```

Um beispielsweise einen relativen Pfad zum Populationsdateiverzeichnis im Verzeichnis `bin` festzulegen, legen Sie die folgenden relativen Pfad fest:

```
./
```

Eigenschaften der Datei IDQTx.cfg

Wenn Sie die Eigenschaft *IdentityReferenceDataLocation* in der Datei idqtx.cfg festlegen, können Sie auch die folgenden Eigenschaften festlegen:

IdentityCacheDir

Pfad zu dem Verzeichnis, das temporäre Datendateien speichert, wenn das Mapping ausgeführt wird. Der Pfad identifiziert ein übergeordnetes Verzeichnis. Der PowerCenter-Integrationsdienst schreibt die temporären Dateien in Verzeichnisse unter den von Ihnen angegebenen Speicherort.

Die Eigenschaft hat die folgenden Standardwert:

```
./identityCache
```

IdentityIndexDir

Pfad zu dem Verzeichnis, das die temporären Indexdateien enthält, die während der Ausführung des Mappings erstellt wurden. Die Identitätsvergleichsanalyse verwendet den Index, um Datensätze vor der Vergleichsanalyse in Gruppen zu sortieren. Der Pfad identifiziert ein übergeordnetes Verzeichnis. Der PowerCenter-Integrationsdienst schreibt die Indexdateien in Verzeichnisse unter den von Ihnen angegebenen Speicherort.

Die Eigenschaft hat die folgenden Standardwert:

```
./identityIndex
```

Hinweis: Ein Developer-Tool-Benutzer kann den Cache-Verzeichnispfad und den Index-Verzeichnispfad in der Match-Umwandlung festlegen. Der PowerCenter-Integrationsdienst verwendet den Pfad in der Datei IDQTx.cfg, wenn der Pfad in der Match-Umwandlung nicht angegeben ist.

Anforderungen an probabilistische und Klassifizierer-Modelle

Probabilistische und Klassifizierer-Modelle sind Referenzdatendateien, die die Informationstypen in einer Datenzeichenfolge identifizieren. Bevor Sie ein Mapping ausführen, die eine probabilistische oder Klassifizierer-Modelldatei ausführt, konfigurieren Sie PowerCenter zum Lesen der Dateien.

Führen Sie die folgenden Konfigurationsaufgaben aus:

- Überprüfen Sie den Speicherort der Modelldatei in der PowerCenter-Verzeichnisstruktur.
- Aktualisieren Sie die Java-Speicherzuordnung für den PowerCenter-Integrationsdienst.

Speicherorte der probabilistischen und Klassifizierer-Modelldatei

Wenn Sie ein Mapping exportieren, das probabilistische oder Klassifizierer-Modelldateien liest, legen Sie ein Verzeichnis als Speicherort für die exportierte Modelldatei fest. Exportieren Sie die Datei in die PowerCenter-Verzeichnisstruktur oder exportieren Sie die Datei in das Dateisystem und kopieren Sie diese auf den PowerCenter-Services-Computer.

Der PowerCenter-Integrationsdienst liest die Modelldateien aus dem folgenden Verzeichnis:

```
[PowerCenter_Installation]/server/bin
```

Wenn die Dateien nicht am Standardspeicherort vorhanden sind, liest der PowerCenter-Integrationsdienst die Datenspeicherorte aus den Eigenschaftendateien. Aktualisieren Sie die Eigenschaftendateien mit den Pfaden zu den Modelldateien.

Die folgende Tabelle beschreibt die Eigenschaftendatei:

Dateiname	Beschreibung
CLASSIFIER.properties	Enthält die Eigenschaft <i>ClassifierFileLocation</i> . Aktualisieren Sie die Eigenschaft, um den Pfad zu den Klassifizierer-Modelldateien zu identifizieren.
NER.properties	Enthält die Eigenschaft <i>NERFileLocation</i> . Aktualisieren Sie die Eigenschaft, um den Pfad zu den probabilistischen Modelldateien zu identifizieren.

Der PowerCenter-Integrationsdienst liest die Eigenschaftendateien aus dem folgenden Verzeichnis:

```
[PowerCenter_Installation]/server/bin
```

Sie können Sie einen absoluten Pfad oder einen relativen Pfad festlegen, um die Speicherorte der Modelldatei in jeder Eigenschaftendatei zu identifizieren. Wenn Sie einen relativen Pfad festlegen, fügt der PowerCenter-Integrationsdienst den Pfad zum folgenden Verzeichnis hinzu:

```
[PowerCenter_Installation]/server/bin
```

Hinweis: Verwenden Sie auf allen Plattformen Schrägstriche zum Definieren der Verzeichnispfade zu den Modelldateien.

Speicherzuordnung des PowerCenter-Integrationsdiensts

Ein PowerCenter-Integrationsdienst verwendet große Mengen von Speicher, um eine Klassifizierer-Modelldatei oder eine probabilistische Modelldatei zu schreiben. Wenn der PowerCenter-Integrationsdienst

nicht auf den erforderlichen Speicher zugreifen kann, kann er die Modelldatei nicht lesen und die Daten nicht verarbeiten.

Überprüfen Sie die Speicherzuordnung für den PowerCenter-Integrationsdienst. Wenn Sie Sitzungen ausführen, die auf eine Klassifizierer-Modelldatei oder eine probalistische Modelldatei verweisen, legen Sie die minimale Speicherzuordnung auf 512 MB fest.

Verwenden Sie die Prozesseigenschaft *Java SDK Maximalspeicher*, um den Speicher zuzuordnen. Um die Eigenschaft zu überprüfen oder zu aktualisieren, melden Sie sich beim Administrator-Tool an und wählen Sie den PowerCenter-Integrationsdienst aus, der die Sitzungs-Tasks ausführt. Sie finden die Eigenschaft *Java SDK Maximalspeicher* in der Ansicht **Prozesse**.

INDEX

D

Data Quality-Integration [6](#)

I

Informatica Data Quality [6](#)