



Informatica® Data Quality
10.5.1

Content Guide

Informatica Data Quality Content Guide

10.5.1

January 2022

© Copyright Informatica LLC 1998, 2022

This software and documentation are provided only under a separate license agreement containing restrictions on use and disclosure. No part of this document may be reproduced or transmitted in any form, by any means (electronic, photocopying, recording or otherwise) without prior consent of Informatica LLC.

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, PowerCenter, and the Informatica logo 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.

The information in this documentation is subject to change without notice. If you find any problems in this documentation, report them to us at infa_documentation@informatica.com.

Informatica products are warranted according to the terms and conditions of the agreements under which they are provided. INFORMATICA PROVIDES THE INFORMATION IN THIS DOCUMENT "AS IS" WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED, INCLUDING WITHOUT ANY WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND ANY WARRANTY OR CONDITION OF NON-INFRINGEMENT.

Publication Date: 2022-01-14

Table of Contents

Preface	5
Informatica Resources.	5
Informatica Network.	5
Informatica Knowledge Base.	5
Informatica Documentation.	5
Informatica Product Availability Matrices.	6
Informatica Velocity.	6
Informatica Marketplace.	6
Informatica Global Customer Support.	6
 Chapter 1: Content Installation.....	 7
Content Overview.	7
Installation Overview.	8
Installation Prerequisites.	8
Accelerator Prerequisites.	8
Address Reference Data Prerequisites.	9
Identity Population Prerequisites.	11
Reference Table Data Prerequisites.	12
Importing Rule and Mapping Objects.	14
Importing Data Domains and Data Domain Groups.	15
Installing Reference Data Files and Other Files.	15
 Chapter 2: Configuration Steps for Address Reference Data.....	 16
Configuration Overview.	16
Configure the Address Reference Data Properties.	16
Review the Address Validator Transformation Advanced Settings.	17
Review the Address Reference Data File Status.	17
Address Reference Data Properties.	17
Rules and Guidelines for Address Reference Data Preload Options.	20
Address Validation Properties in the Preferences Window.	21
Address Validation Data Properties.	21
Address Validation License Properties.	22
Address Validation Engine Properties.	23
 Appendix A: Address Validator Transformation Advanced Properties.....	 25
Address Validator Transformation Advanced Properties.	25
Alias Locality.	25
Alias Street.	26
Casing Style.	26
Country of Origin.	27

Country Type.	27
Default Country.	28
Dual Address Priority.	28
Element Abbreviation.	28
Execution Instances.	29
Flexible Range Expansion.	30
Geocode Data Type.	30
Global Max Field Length.	31
Global Preferred Descriptor.	31
Input Format Type.	32
Input Format With Country	32
Line Separator.	32
Matching Alternatives.	33
Matching Extended Archive.	33
Matching Scope.	34
Max Result Count.	34
Mode.	34
Optimization Level.	35
Output Format Type.	36
Output Format With Country.	36
Preferred Language.	36
Preferred Script.	42
Ranges To Expand.	43
Standardize Invalid Addresses.	44
Tracing Level.	44
Index.	45

Preface

Read the *Informatica Content Guide* to learn about the prebuilt data quality rules and reference data that you can download and install in the Informatica domain. Informatica releases prebuilt rules in accelerator packages for countries, regions, and business sectors. Reference data includes address reference data files and identity population files.

Informatica Resources

Informatica provides you with a range of product resources through the Informatica Network and other online portals. Use the resources to get the most from your Informatica products and solutions and to learn from other Informatica users and subject matter experts.

Informatica Network

The Informatica Network is the gateway to many resources, including the Informatica Knowledge Base and Informatica Global Customer Support. To enter the Informatica Network, visit <https://network.informatica.com>.

As an Informatica Network member, you have the following options:

- Search the Knowledge Base for product resources.
- View product availability information.
- Create and review your support cases.
- Find your local Informatica User Group Network and collaborate with your peers.

Informatica Knowledge Base

Use the Informatica Knowledge Base to find product resources such as how-to articles, best practices, video tutorials, and answers to frequently asked questions.

To search the Knowledge Base, visit <https://search.informatica.com>. If you have questions, comments, or ideas about the Knowledge Base, contact the Informatica Knowledge Base team at KB_Feedback@informatica.com.

Informatica Documentation

Use the Informatica Documentation Portal to explore an extensive library of documentation for current and recent product releases. To explore the Documentation Portal, visit <https://docs.informatica.com>.

If you have questions, comments, or ideas about the product documentation, contact the Informatica Documentation team at infa_documentation@informatica.com.

Informatica Product Availability Matrices

Product Availability Matrices (PAMs) indicate the versions of the operating systems, databases, and types of data sources and targets that a product release supports. You can browse the Informatica PAMs at <https://network.informatica.com/community/informatica-network/product-availability-matrices>.

Informatica Velocity

Informatica Velocity is a collection of tips and best practices developed by Informatica Professional Services and based on real-world experiences from hundreds of data management projects. Informatica Velocity represents the collective knowledge of Informatica consultants who work with organizations around the world to plan, develop, deploy, and maintain successful data management solutions.

You can find Informatica Velocity resources at <http://velocity.informatica.com>. If you have questions, comments, or ideas about Informatica Velocity, contact Informatica Professional Services at ips@informatica.com.

Informatica Marketplace

The Informatica Marketplace is a forum where you can find solutions that extend and enhance your Informatica implementations. Leverage any of the hundreds of solutions from Informatica developers and partners on the Marketplace to improve your productivity and speed up time to implementation on your projects. You can find the Informatica Marketplace at <https://marketplace.informatica.com>.

Informatica Global Customer Support

You can contact a Global Support Center by telephone or through the Informatica Network.

To find your local Informatica Global Customer Support telephone number, visit the Informatica website at the following link:

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

To find online support resources on the Informatica Network, visit <https://network.informatica.com> and select the eSupport option.

CHAPTER 1

Content Installation

This chapter includes the following topics:

- [Content Overview, 7](#)
- [Installation Overview, 8](#)
- [Installation Prerequisites, 8](#)
- [Importing Rule and Mapping Objects, 14](#)
- [Importing Data Domains and Data Domain Groups, 15](#)
- [Installing Reference Data Files and Other Files, 15](#)

Content Overview

Informatica applications can use data quality rules and reference data to improve data accuracy, enhance data, and standardize data. Informatica uses the term *content* to refer to rules and reference data collectively.

Informatica distributes the following types of content:

Accelerators

Accelerators are content bundles that address common data quality problems in a country, a region, or an industry. An accelerator might contain mapplets and rules that you can use to analyze and enhance the data in an organization. An accelerator might also contain data domains that you can use to discover the types of information that the data contains. You add the mapplets, rules, and data domains to the Model repository.

Informatica Data Quality includes a Core accelerator and a Core Data Domain accelerator. You can buy and download additional accelerators from Informatica.

For more information about accelerators, see the *Data Quality Accelerator Guide*.

Address reference data files

Address reference data files contain information on all valid addresses in a country. The Address Validator transformation uses address reference data to analyze the quality of the input data that you select. The transformation compares the input data to the address reference data and fixes any error it finds in the input data.

You purchase address reference data on a subscription basis. Informatica updates address reference data files with new postal information at regular intervals. You can download the current address data files at any time during your subscription period.

Identity population files

Identity population files contain metadata for personal, household, and corporate identities. Population files also contain algorithms that apply the metadata to input data. The Match transformation and the Comparison transformation use population file data to parse potential identities from input fields.

You must copy address reference data, identity population data, and accelerator demonstration data to the installation directories. You can use Informatica Developer to import accelerator rules, demonstration mappings, and reference table metadata to the Model repository and to write reference table data to the reference data database.

Installation Overview

Use Informatica Developer to import accelerator rules, demonstration mappings, and reference tables to the Model repository and to write reference table data to the reference data database. To install address reference data, identity populations, and accelerator demonstration data, copy the files manually to the target machine.

When you install address reference data files and identity population files, verify that the Data Integration Service can access the machine to which you install the files. You install address reference data files and identity population files to an Informatica domain.

You import a set of prebuilt Informatica rules or reference data files once to a Model repository and reference data database. If more than one Developer tool or Analyst tool user imports the rules or data files, the data is either overwritten each time or installed multiple times to different folders in the same system.

Note: You must install all accelerator reference data to a single project in the Model repository.

Installation Prerequisites

Complete or verify the prerequisites for the types of content that you will install before you install the content.

You must complete the Informatica Data Quality or PowerCenter® installation before you install content.

Accelerator Prerequisites

The repository objects and data files in an accelerator operate in the same way as other objects and files in the Informatica system. Some rules and guidelines apply to the accelerator contents.

Consider the following rules and guidelines when you install an accelerator:

- Before you import or copy files, verify that you have all privileges on the Data Integration Service, the Content Management Service, and the Analyst Service.
- Import the accelerators to a single Model repository project. Create the project before you import the accelerators.
- Install the Core accelerator before you install another accelerator.
- Install the Core Data Domain accelerator before you install the Data Domain accelerator.

- If you import a metadata file that contains an object in common with an accelerator that you imported earlier, replace the object in the repository.

Address Reference Data Prerequisites

Install Informatica Data Quality or PowerCenter before you install address reference data to either product.

Before you install address reference data for PowerCenter, stop the PowerCenter Integration Service. Before you install address reference data for Data Quality, stop the Data Integration Service and the Content Management Service. After you install the data, restart any service that you stopped. If you do not stop and restart the services when you install address reference data, the Address Validator transformation continues to run any older data that it stores in memory.

The Address Validator transformation can read the following types of address reference data:

Types of Address Reference Data

Address code lookup data

Install address code lookup data to retrieve a partial address or full address from a code value on an input port. The completeness of the address depends on the level of address code support in the country to which the address belongs. To read the address code from an input address, select the country-specific ports in the Discrete port group.

You can select ports for the following countries:

- Austria. Returns an address to building level.
- Germany. Returns an address to locality, municipality, or street level.
- Japan. Returns an address to the unique mailbox level.
- South Africa. Returns an address to street level.
- South Korea. Returns an address to the unique mailbox level.
- Serbia. Returns an address to street level.
- United Kingdom. Returns an address to the unique mailbox level.

The Address Validator transformation reads address code lookup data when you configure the transformation to run in address code lookup mode.

Batch and interactive data

Install batch and interactive data to perform address validation on a set of address records. Use batch and interactive data to verify that the input addresses are fully deliverable and complete based on the current postal data from the national mail carrier.

When you configure the transformation to run in batch mode, the Address Validator transformation returns a single address for each input address. When you configure the transformation to run in interactive mode, the Address Validator transformation returns one or more addresses for each input address.

CAMEO data

Install CAMEO data to add customer segmentation data to residential address records. Customer segmentation data indicates the likely income level and lifestyle preferences of the residents at each address.

The Address Validator transformation reads CAMEO data when you configure the transformation to run in batch mode or certified mode.

Certified data

Install certified data to verify that address records meet the certification standards that a mail carrier defines. An address meets a certification standard if it contains data elements that can identify a unique mailbox, such as delivery point data elements. When an address meets a certification standard, the mail carrier charges a reduced delivery rate.

The following countries define certification standards:

- Australia. Certifies mail according to the Address Matching Approval System (AMAS) standard.
- Canada. Certifies mail according to the Software Evaluation And Recognition Program (SERP) standard.
- France. Certifies mail according to the National Address Management Service (SNA) standard.
- New Zealand. Certifies mail according to the SendRight standard.
- United States. Certifies mail according to the Coding Accuracy Support System (CASS) standard.

The Address Validator transformation reads certified data when you configure the transformation to run in certified mode.

Geocode data

Install geocode data to add geocodes to address records. Geocodes are latitude and longitude coordinates.

The Address Validator transformation reads geocode data when you configure the transformation to run in batch mode or certified mode.

Suggestion list data

Install suggestion list data to find alternative valid versions of a partial address record. Use suggestion list data when you configure an address validation mapping to process address records one by one in real time. The Address Validator transformation uses the data elements in the partial address to perform a duplicate check on the suggestion list data. The transformation returns any valid address that includes the information in the partial address.

The Address Validator transformation reads suggestion list data when you configure the transformation to run in suggestion list mode.

Supplementary data

Install supplementary data to add data to an address record that can assist the mail carrier in mail delivery. Use the supplementary data to add detail about the geographical or postal area that contains the address. In some countries, supplementary data can provide a unique identifier for a mailbox within the postal system.

The Address Validator transformation reads supplementary data when you configure the transformation to run in batch mode or certified mode.

Note: The transformation does not read address reference data in country recognition mode or in parse mode.

Rules and Guidelines for Address Reference Data

Informatica releases new versions of address reference data files at regular intervals. When you subscribe to address reference data for a country, you can download and install the latest data files for the country at any time.

Consider the following rules and guidelines when you work with address reference data:

- Do not run an address validation mapping or session while you install address reference data.

- Informatica works with national mail carriers to develop the address reference data. When a mail carrier updates its data records with new information, Informatica adds the information to the address reference data files for the country.
- Informatica updates address reference data files several times each year. Informatica sends you a monthly email to notify you that the latest updates are ready for download.

Address Certification Considerations

The Address Validator transformation can indicate if an address contains the data required by the certification standards of national mail carriers. The standards require that a software application validates address accuracy and prepares address records in the correct format for automated mail sorting and delivery. If you use the data in a certified validation process, update the address reference data files once a month.

If you use United States or Canadian address reference data to certify address records to the Coding Accuracy Software System (CASS) or Software Evaluation and Recognition Program (SERP) standard, you must use reference data that is no more than 60 days old.

Identity Population Prerequisites

Install the identity population files to a location that the Informatica services can access.

In a Data Quality installation, the Data Integration Service reads the population files. Install the files on the Data Integration Service host machine or to a shared directory on a machine that the Data Integration Service can access. In a PowerCenter installation, the PowerCenter Integration Service reads the population files. Install the files on the PowerCenter Integration Service host machine or to a shared directory on a machine that the Integration Service can access.

Informatica Data Quality stores the path to the population file directory in the *Reference Data Location* property on the Content Management Service. Use the Administrator tool to verify or edit the path.

Install the population files to the following directory on the Data Integration Service machine:

```
[Informatica_installation_directory]/services/DQContent/INFA_Content/identity/default
```

Before you install the population files, verify that the `/default/` directory is present. Before you create a mapping that reads the population files, verify that the *Reference Data Location* property on the Content Management Service specifies the parent directory for the `/default/` directory.

PowerCenter stores the path to the population file directory in the *IdentityReferenceDataLocation* property in the IDQTx.cfg configuration file. Open the file and verify or edit the path.

Consider the following rules and guidelines before you install the identity population files to a PowerCenter machine:

- The Content installer writes the population files to the following directory on the PowerCenter Integration Service machine:

```
[Informatica_installation_directory]/services/DQContent/INFA_Content/identity/default
```

Before you run the Content installer, verify that the `/default/` directory is present. Before you run a session that reads the population files, verify that the *IdentityReferenceDataLocation* property in the IDQTx.cfg file specifies the parent directory for the `/default/` directory.

The PowerCenter installer writes the IDQTx.cfg file to the following path:

```
[Informatica_Installation_directory]/server/bin
```

- Earlier versions of PowerCenter read the path to the population files from the *SSAPR* environment variable. The PowerCenter Integration Service can read the location of the population files from the IDQTx.cfg file or from the *SSAPR* environment variable. By default, the PowerCenter Integration Service

reads the location from the IDQTx.cfg file. If the IDQTx.cfg file does not specify the location, or if the file is not present, the PowerCenter Integration Service reads the location from the SSAPR environment variable.

- The IDQTx.cfg file and the SSAPR environment variable specify the path to the parent directory of the /default/ directory. The path does not include the /default/ directory name. The path cannot contain character spaces.
- You can use the current version of the population files with the current versions of Informatica Data Quality and PowerCenter. To use the current population files with an earlier version of PowerCenter, install the current version of the Data Quality Integration plug-in to PowerCenter.

Note: When you install the current plug-in on a PowerCenter machine, you cannot import objects from an older Model repository to the PowerCenter repository. You can continue to use any data quality object that you imported to the PowerCenter repository before you installed the current plug-in.

Reference Table Data Prerequisites

Before you import the reference data, verify that the Data Integration Service, Model Repository Service, and Content Management Service are running.

The import operation writes the metadata for the reference tables to the Model repository and writes the reference data values to tables in the reference data database. The Content Management Service stores the reference data database connection name. You associate a reference data database with a single Model repository. You can specify the same reference data database for multiple Content Management Services if the Content Management Services identify the same Model repository.

You can create the reference data database in the following relational database systems:

- IBM DB2
- Microsoft Azure SQL Server
- Microsoft SQL Server
- Oracle
- PostgreSQL

Verify that the reference data database supports mixed-case column names.

Allow 200 MB of disk space for the database.

Note: Ensure that you install the database client on the machine on which you want to run the Content Management Service.

For more information about configuring the database, see the documentation for the database system.

IBM DB2 Database Requirements

Use the following guidelines when you set up the repository on IBM DB2:

- Verify that the database user account has CREATETAB and CONNECT privileges.
- Verify that the database user has SELECT privileges on the SYSCAT.DBAUTH and SYSCAT.DBTAUTH tables.
- Informatica does not support IBM DB2 table aliases for repository tables. Verify that table aliases have not been created for any tables in the database.
- Set the tablespace pageSize parameter to 32768 bytes.
- Set the NPAGES parameter to at least 5000. The NPAGES parameter determines the number of pages in the tablespace.

Microsoft Azure SQL Database Requirements

Use the following guidelines when you set up the repository:

- Set the allow snapshot isolation and read committed isolation level to ALLOW_SNAPSHOT_ISOLATION and READ_COMMITTED_SNAPSHOT to minimize locking contention.

To set the isolation level for the database, run the following commands:

```
ALTER DATABASE DatabaseName SET ALLOW_SNAPSHOT_ISOLATION ON
```

```
ALTER DATABASE DatabaseName SET READ_COMMITTED_SNAPSHOT ON
```

To verify that the isolation level for the database is correct, run the following commands:

```
SELECT snapshot_isolation_state FROM sys.databases WHERE name=[DatabaseName]
```

```
SELECT is_read_committed_snapshot_on FROM sys.databases WHERE name = DatabaseName
```

- The database user account must have the CONNECT, CREATE TABLE, and CREATE VIEW privileges.

Microsoft SQL Server Database Requirements

Use the following guidelines when you set up the repository:

- Verify that the database user account has CONNECT and CREATE TABLE privileges.

Oracle Database Requirements

Use the following guidelines when you set up the repository on Oracle:

- Verify that the database user has the following privileges:

ALTER SEQUENCE

ALTER TABLE

CREATE SEQUENCE

CREATE SESSION

CREATE TABLE

CREATE VIEW

DROP SEQUENCE

DROP TABLE

- Informatica does not support Oracle public synonyms for repository tables. Verify that public synonyms have not been created for any tables in the database.

PostgreSQL Database Requirements

Use the following guidelines when you set up the repository on PostgreSQL:

- Use a JDBC connection to connect to the PostgreSQL database.

Informatica installs a DataDirect JDBC driver for PostgreSQL that you can use to connect to the database. Find the driver in the clients/DeveloperClient/infacmd installation directory, and copy the driver to the clients/externaljdbcjars directory .

- Specify the database schema name. Do not leave the schema name blank.

If the database uses the default PostgreSQL schema name of `public`, you can specify `public` as the schema name.

- Verify that the database user has the CONNECT and CREATE TABLE privileges.

Verifying the Support Status for Mixed-Case Column Names

Use the Administrator tool to verify that the reference data database supports mixed-case column names.

1. Log in to the Administrator tool.
2. Select the **Domain** tab, and select **Connections**.
3. Select the reference data database.
4. Review the **Advanced Connection Properties** for the database.
5. Verify that **Support mixed case identifiers** is set to true.
If not, edit this property.

Importing Rule and Mapping Objects

Use the Developer tool to import metadata for accelerator rules, demonstration mappings, and mapping data sources. During the import operation, select the reference data file that the rules and mappings use.

1. In the Developer tool, connect to the Model repository that contains the destination project for the metadata.
2. In the Object Explorer, select the destination project.
For example, select the *Informatica_DQ_Content* project. If required, create a project in the Model repository.
3. Select **File > Import**.
4. In the **Import** dialog box, select **Informatica > Import Object Metadata File (Advanced)**.
5. Click **Next**.
6. Browse to the XML metadata file in the accelerator directory structure, and select the file.
7. Click **Open**, and click **Next**.
8. In the **Source** pane, select the items that appear under the project node.
9. In the **Target** pane, select the destination project.
10. Click **Add to Target**.
 - If the repository project contains an object that you want to add, the Developer tool prompts you to merge the object with the current object. Click **Yes** to merge the objects.
 - If the Developer tool prompts you to rename the objects, click **No**.
 - If any object remains in the **Source** pane, use the pointer to move the object to the target project.
11. Click **Next**.
12. Browse to the compressed reference data file in the accelerator directory structure, and select the file.
13. Click **Open**.
14. Verify that the code page is UTF-8, and click **Next**.
15. In the **Target Connection** field, select the reference data database.
16. Click **Finish**.

Importing Data Domains and Data Domain Groups

Use the Preferences dialog box to import metadata for data domains and data domain groups. During the import operation, select the reference data file that the data domains use.

1. In the Developer tool, connect to the Model repository that contains the destination project for the metadata.
2. Select **Window > Preferences**.
3. In the **Preferences** dialog box, expand the Informatica node and select **Data Domain Glossary**.
4. In the repository pane, select the top-level node for the data domains or the data domain groups.
5. Click **Import**.
6. Browse to the XML metadata file in the accelerator directory structure, and select the file.
7. Click **Open**, and click **Next**.
8. In the **Source** pane, select the data domain glossary project.
9. In the **Target** pane, select the destination project.
10. Select the following option in the Resolution field:
Replace option in target
11. Click **Add Contents to Target**.
 - If the Developer tool prompts you to add the objects, click **Yes**.
 - If the Developer tool prompts you to rename the objects, click **No**.
12. Click **Next**.
13. If the import operation identifies dependencies, copy the dependent objects from the source project to the target project.
14. Click **Next**.
15. Browse to the compressed reference data file in the accelerator directory structure, and select the file.
16. Click **Open**.
17. Verify that the code page is UTF-8, and click **Next**.
18. In the **Target Connection** field, select the reference data database.
19. Click **Finish**.

Installing Reference Data Files and Other Files

Install address reference data files, identity population files, and accelerator demonstration data files on a machine that a Data Integration Service or PowerCenter Integration Service can access. Before you install any file, review the installation prerequisites.

To install the files, copy the files manually to the machine.

You can replace any older file with a file of the same name.

Each time you install address reference data, review the configuration steps. For information about the configuration steps for address reference data, see the ["Configuration Overview" on page 16](#)

CHAPTER 2

Configuration Steps for Address Reference Data

This chapter includes the following topics:

- [Configuration Overview, 16](#)
- [Address Reference Data Properties, 17](#)
- [Address Validation Properties in the Preferences Window, 21](#)

Configuration Overview

After you install address reference data for Data Quality or PowerCenter, you must configure the properties that the Data Integration Service or PowerCenter Integration Service uses when it runs an address validation mapping.

You can also verify or edit address reference data settings in the Developer tool.

Configure the Address Reference Data Properties

After you install address reference data for Data Quality or PowerCenter, configure the address reference data properties.

You provide the license key for the address reference data and the path to the address reference data files. You also determine how the Integration Service loads reference data.

If you install address reference data for Data Quality, use the Administrator tool to configure the properties on the Content Management Service. If you install address reference data for PowerCenter, configure the properties in the AD50.cfg file.

Installing Address Reference Data

After you install address reference data, you add the license key for the data to the License property on the Content Management Service or in the AD50.cfg file. If you install more than one type of address reference data, you add license keys for each type in a comma-separated string.

If you install reference data files at different times, add the license key data property with the license key for the new files. You provide the license key data as a comma-delimited string.

Updating Address Reference Data

You can update the address reference data you installed for a country without entering a new license key. You change the license key when your subscription to the data for that country expires.

Review the Address Validator Transformation Advanced Settings

After you install address reference data for Data Quality, review the Address Validator transformation advanced settings.

You can edit these settings to ensure that the address validation mapping processes the source data in the correct manner for your project. You find the advanced settings on the Advanced tab of the transformation.

Review the Address Reference Data File Status

After you install address reference data for Data Quality, review the status of the data files.

You can view a list of the address reference data files on the domain that you connect to. Verify that the files are properly licensed and that the file types match the processing mode that you configured in the Address Validator transformation. Use the Developer tool to view the file list.

Note: You can review address reference data file status at any time. Review the status at regular intervals to verify that the installed address reference data is up to date.

Address Reference Data Properties

The Data Integration Service or PowerCenter Integration Service reads address reference data properties when you run an address validation mapping or session.

If you run an address validation mapping in Data Quality, the Data Integration Service reads the address reference data properties that you set on the Content Management Service. Use the Administrator tool to configure the Content Management Service properties. If you run an address validation session in PowerCenter, the Integration Service reads the address reference data properties from the AD50.cfg file. Locate the AD50.cfg file and configure the properties.

You must enter a license key, the reference data location, and at least one data preload value before you run an address validation mapping or session. Optionally, enter values to the other properties.

Note: The AD50.cfg file and the Content Management Service use the same names for the address reference data properties. However, the property names in AD50.cfg do not contain spaces. For example, you can set the `Max Address Object Count` property on the Content Management Service. You set the `MaxAddressObjectCount` property in AD50.cfg.

The following table describes the address reference data properties:

Property	Description
License	License key to activate validation reference data. You might have more than one key, for example, if you use batch reference data and geocoding reference data. Enter keys as a comma-delimited list. The property is empty by default.
Reference Data Location	Location of the address reference data files. Enter the full path to the files. Install all address reference data files to a single location. The property is empty by default.

Property	Description
Full Pre-Load Countries	<p>List of countries for which all batch, CAMEO, certified, interactive, or supplementary reference data is loaded into memory before address validation begins. Enter the three-character ISO country codes in a comma-separated list. For example, enter DEU,FRA,USA. Enter ALL to load all data sets. The property is empty by default.</p> <p>Load the full reference database to increase performance. Some countries, such as the United States, have large databases that require significant amounts of memory.</p>
Partial Pre-Load Countries	<p>List of countries for which batch, CAMEO, certified, interactive, or supplementary reference metadata and indexing structures are loaded into memory before address validation begins. Enter the three-character ISO country codes in a comma-separated list. For example, enter DEU,FRA,USA. Enter ALL to partially load all data sets. The property is empty by default.</p> <p>Partial preloading increases performance when not enough memory is available to load the complete databases into memory.</p>
No Pre-Load Countries	<p>List of countries for which no batch, CAMEO, certified, interactive, or supplementary reference data is loaded into memory before address validation begins. Enter the three-character ISO country codes in a comma-separated list. For example, enter DEU,FRA,USA. Default is ALL.</p>
Full Pre-Load Geocoding Countries	<p>List of countries for which all geocoding reference data is loaded into memory before address validation begins. Enter the three-character ISO country codes in a comma-separated list. For example, enter DEU,FRA,USA. Enter ALL to load all data sets. The property is empty by default.</p> <p>Load all reference data for a country to increase performance when processing addresses from that country. Some countries, such as the United States, have large data sets that require significant amounts of memory.</p>
Partial Pre-Load Geocoding Countries	<p>List of countries for which geocoding reference metadata and indexing structures are loaded into memory before address validation begins. Enter the three-character ISO country codes in a comma-separated list. For example, enter DEU,FRA,USA. Enter ALL to partially load all data sets. The property is empty by default.</p> <p>Partial preloading increases performance when not enough memory is available to load the complete databases into memory.</p>
No Pre-Load Geocoding Countries	<p>List of countries for which no geocoding reference data is loaded into memory before address validation begins. Enter the three-character ISO country codes in a comma-separated list. For example, enter DEU,FRA,USA. Default is ALL.</p>
Full Pre-Load Suggestion List Countries	<p>List of countries for which all suggestion list reference data is loaded into memory before address validation begins. Enter the three-character ISO country codes in a comma-separated list. For example, enter DEU,FRA,USA. Enter ALL to load all data sets. The property is empty by default.</p> <p>Load the full reference database to increase performance. Some countries, such as the United States, have large databases that require significant amounts of memory.</p>
Partial Pre-Load Suggestion List Countries	<p>List of countries for which the suggestion list reference metadata and indexing structures are loaded into memory before address validation begins. Enter the three-character ISO country codes in a comma-separated list. For example, enter DEU,FRA,USA. Enter ALL to partially load all data sets. The property is empty by default.</p> <p>Partial preloading increases performance when not enough memory is available to load the complete databases into memory.</p>
No Pre-Load Suggestion List Countries	<p>List of countries for which no suggestion list reference data is loaded into memory before address validation begins. Enter the three-character ISO country codes in a comma-separated list. For example, enter DEU,FRA,USA. Default is ALL.</p>

Property	Description
Full Pre-Load Address Code Countries	<p>List of countries for which all address code lookup reference data is loaded into memory before address validation begins. Enter the three-character ISO country codes in a comma-separated list. For example, enter DEU,FRA,USA. Enter ALL to load all data sets. The property is empty by default.</p> <p>Load the full reference database to increase performance. Some countries, such as the United States, have large databases that require significant amounts of memory.</p>
Partial Pre-Load Address Code Countries	<p>List of countries for which the address code lookup reference metadata and indexing structures are loaded into memory before address validation begins. Enter the three-character ISO country codes in a comma-separated list. For example, enter DEU,FRA,USA. Enter ALL to partially load all data sets. The property is empty by default.</p> <p>Partial preloading increases performance when not enough memory is available to load the complete databases into memory.</p>
No Pre-Load Address Code Countries	<p>List of countries for which no address code lookup reference data is loaded into memory before address validation begins. Enter the three-character ISO country codes in a comma-separated list. For example, enter DEU,FRA,USA. Default is ALL.</p>
Preloading Method	<p>Determines how the Data Integration Service preloads address reference data into memory. The MAP method and the LOAD method both allocate a block of memory and then read reference data into this block. However, the MAP method can share reference data between multiple processes. Default is MAP.</p>
Max Result Count	<p>Maximum number of addresses that address validation can return in suggestion list mode. Set a maximum number in the range 1 through 100. Default is 20.</p>
Memory Usage	<p>Number of megabytes of memory that the address validation library files can allocate. Default is 4096.</p>
Max Address Object Count	<p>Maximum number of address validation instances that can run at the same time. Default is 3. Set a value that is greater than or equal to the Maximum Parallelism value on the Data Integration Service.</p> <p>If the Data Integration Service will run mappings with Address Validator transformations that you deploy as web service applications, increase the Max Address Object Count value to at least 10. For more information on configuring address validation properties for web services, see the <i>Informatica Performance Tuning Guide</i>.</p>
Max Thread Count	<p>Maximum number of threads that address validation can use. Set to the total number of cores or threads available on a machine. Default is 2.</p>

Property	Description
Cache Size	<p>Size of cache for databases that are not preloaded. Caching reserves memory to increase lookup performance in reference data that has not been preloaded.</p> <p>Set the cache size to LARGE unless all the reference data is preloaded or you need to reduce the amount of memory usage.</p> <p>Enter one of the following options for the cache size in uppercase letters:</p> <ul style="list-style-type: none"> - NONE. No cache. Enter NONE if all reference databases are preloaded. - SMALL. Reduced cache size. - LARGE. Standard cache size. <p>Default is LARGE.</p>
SendRight Report Location	<p>Location to which an address validation mapping writes a SendRight report and any log file that relates to the report. You generate a SendRight report to verify that a set of New Zealand address records meets the certification standards of New Zealand Post. Enter a local path on the machine that hosts the Data Integration Service that runs the mapping.</p> <p>By default, address validation writes the report file to the <code>bin</code> directory of the Informatica installation. If you enter a relative path, the Content Management Service appends the path to the <code>bin</code> directory.</p>

Rules and Guidelines for Address Reference Data Preload Options

If you run a mapping that reads address reference data, verify the policy that the Data Integration Service uses to load the data into memory. To configure the policy, use the preload options on the address validation process properties. The Data Integration Service reads the preload options from the Content Management Service when an address validation mapping runs.

Consider the following rules and guidelines when you configure the preload options on the Content Management Service:

- By default, the Content Management Service applies the ALL value to the options that indicate no data preload. If you accept the default options, the Data Integration Service reads the address reference data from files in the directory structure when the mapping runs.
- The address validation process properties must indicate a preload method for each type of address reference data that a mapping specifies. If the Data Integration Service cannot determine a preload policy for a type of reference data, it ignores the reference data when the mapping runs.
- The Data Integration Service can use a different method to load data for each country. For example, you can specify full preload for United States suggestion list data and partial preload for United Kingdom suggestion list data.
- The Data Integration Service can use a different preload method for each type of data. For example, you can specify full preload for United States batch data and partial preload for United States address code data.
- Full preload settings supersede partial preload settings, and partial preload settings supersede settings that indicate no data preload.

For example, you might configure the following options:

Full Pre-Load Geocoding Countries: DEU

No Pre-Load Geocoding Countries: ALL

The options specify that the Data Integration Service loads German geocoding data into memory and does not load geocoding data for any other country.

- The Data Integration Service loads the types of address reference data that you specify in the address validation process properties. The Data Integration Service does not read the mapping metadata to identify the address reference data that the mapping specifies.

Address Validation Properties in the Preferences Window

You can view the properties of the address validation engine and the address reference data files that the engine reads in the Developer tool. The Developer tool exposes the properties of the engine that the Data Integration Service uses to run address validation mappings. The Developer tool lists the properties for the Content Management Service that governs the address validation operations.

Use the **Preferences** window in the Developer tool to review the properties. Select the **Content Status** option on the **Preferences** window to identify the Content Management Service that the current Data Integration Service uses. To view the properties, select the local Content Management Service.

You can view the following properties:

Address Validation Data

The address validation data properties list the types of reference data that the current Content Management Service can provide to the Data Integration Service. The properties also indicate the countries to which the reference data applies.

Address Validation Engine

The address validation engine properties include the current engine version, the engine in which the certification components were most recently updated, and the data preloading method.

Address Validation License

The address validation license properties include license information for the reference data that the current Content Management Service can provide to the Data Integration Service.

Address Validation Data Properties

The address validation data properties list the types of reference data that the current Content Management Service can provide to the Data Integration Service. The properties also include the countries to which the reference data applies.

The following table describes the data properties that display when you select the Content Management Service in the **Content Status** view:

Property	Description
Country ISO	The country to which the address reference data file applies. The property shows the ISO three-character code for the country.
Expiry Date	The date on which the current file expires. Informatica releases a newer file on the expiry date. You can use the current address reference data file after the expiry date, but the data in the file may no longer be accurate.
Country Type	The type of address validation that you can perform with the data. You select the processing type in the Mode option on the General Settings tab. If the mode that you select does not correspond to an address data file on the domain, the address validation mapping will fail.

Property	Description
Unlock Expiry Date	The date on which the license expires. You cannot use any version of the file after the unlock expiry date. The Unlock Expiry Date property and Expiration Date property on the Address Validation License Properties view represent the same information.
Unlock Start Date	The date on which the license becomes effective for the mode that the Country Type property identifies and the country that the Country ISO property identifies. You cannot use any version of the file before the unlock start date.

Address Validation License Properties

The address validation license properties include license information for the reference data that the current Content Management Service can provide to the Data Integration Service.

The following table describes the license properties that display when you select the Content Management Service in the **Content Status** view:

Property	Description
Unlock Code	The license code that unlocks the reference data for the mode that the Code Type property identifies. The Developer tool displays the first four characters of the code and masks the other characters.
Code Type	The mode of address validation that you can perform with the data that the license specifies. Informatica issues a single license code for each mode. The license code can apply to one or more countries. You select the processing type in the Mode option on the General Settings tab. If the mode that you select does not correspond to an address data file on the domain, the address validation mapping will fail.
Country List	The countries for which the unlock code unlocks the reference data. The Country List property contains one or more ISO three-character codes for each country.
Status	The status of the license code. The property returns OK when the license file is valid.
Expiration Date	The date on which the license expires. The Expiration Date property and the Unlock Expiry Date property on the Address Validation Data Properties view represent the same information.

Address Validation Engine Properties

The address validation engine properties include the current engine version, the engine in which the certification components were most recently updated, and the data preloading method.

The following table describes the engine properties that display when you select the Content Management Service in the **Content Status** view:

Property	Value
Engine Version	The version of the address validation engine that the Data Integration Service runs.
CASS Version	<p>The version of the address validation engine in which Informatica most recently updated the CASS certification components. Use the property to identify the engine version in a CASS certification report.</p> <p>The property also includes the CASS certification cycle that the engine supports. For example, the engine might support certification cycle N.</p>
AMAS Version	The version of the address validation engine in which Informatica most recently updated the AMAS certification components. Use the property to identify the engine version in a AMAS certification report.
SendRight Version	The version of the address validation engine in which Informatica most recently updated the SendRight certification components. Use the property to identify the engine version in a SendRight certification report.
SERP Version	The version of the address validation engine in which Informatica most recently updated the SERP certification components. Use the property to identify the engine version in a SERP certification report.
SNA Version	The version of the address validation engine in which Informatica most recently updated the SNA certification components. Use the property to identify the engine version in a SNA certification report.
Preloading Method	<p>The method that the Data Integration Service uses to preload reference database into memory. The Content Management Service properties specify the countries for which the Data Integration Service preloads reference data. The possible values are MAP and LOAD. The default value is MAP.</p> <p>The MAP method and the LOAD method both allocate a block of memory and then read the reference data into the block. However, the MAP method can share reference data between multiple processes.</p>
Cache Size	The size of the data cache that the Data Integration Service uses for reference data that the service does not preload. The possible values are NONE, SMALL, and LARGE. The default value is LARGE.
Maximum Memory Usage	The number of megabytes of memory that the address validation engine can allocate. The default value is 4096.
Maximum Address Object Count	The maximum number of address validation instances that the Data Integration Service can run at the same time. The default value is 3.
Maximum Thread Count	The maximum number of threads that address validation can use. The default value is 2.
Maximum Result Count	The maximum number of addresses that address validation can return when you run a mapping in suggestion list mode. The default value is 20. The upper limit on the property is 100.

Property	Value
Current Date	The current date. The Developer tool returns the property values that apply on the current date.
Write XML BOM	Indicates whether the Data Integration Service writes a byte order mark in the GetConfig.xml file. The possible values are ALWAYS, IF_NECESSARY, and NEVER. The default value is IF_NECESSARY.
XML Encoding	Identifies the XML encoding that the address validation engine uses to read and write data.

APPENDIX A

Address Validator Transformation Advanced Properties

This appendix includes the following topic:

- [Address Validator Transformation Advanced Properties, 25](#)

Address Validator Transformation Advanced Properties

The advanced properties on the Address Validator transformation include properties that determine how the transformation uses address reference data. Open the transformation in the Developer tool to review the advanced properties. Verify that the advanced properties define the required behavior for the address reference data that you install.

Alias Locality

Determines whether address validation replaces a valid locality alias with the official locality name.

A locality alias is an alternative locality name that the USPS recognizes as an element in a deliverable address. You can use the property when you configure the Address Validator transformation to validate United States address records in Certified mode.

The following table describes the Alias Locality options:

Option	Description
Off	Disables the Alias Locality property.
Official	Replaces any alternative locality name or locality alias with the official locality name. Default option.
Preserve	Preserves a valid alternative locality name or locality alias. If the input locality name is not valid, address validation replaces the name with the official name.

Alias Street

Determines whether address validation replaces a street alias with the official street name.

A street alias is an alternative street name that the USPS recognizes as an element in a deliverable address. You can use the property when you configure the Address Validator transformation to validate United States address records in Certified mode.

The following table describes the Alias Street options:

Option	Description
Off	Does not apply the property.
Official	Replaces any alternative street name or street alias with the official street name. Default option.
Preserve	Preserves a valid alternative street name or street alias. If the input street name is not valid, address validation replaces the name with the official name.

Casing Style

Specifies the character case style that the transformation applies to the output address data.

The following table describes the Casing Style options:

Option	Description
Assign Parameter	Uses a parameter that you define to set the casing style.
Lower	Writes the output address in lowercase letters.
Mixed	Uses the casing style in use in the destination country when it is possible to do so.
Preserved	Applies the casing style that the address reference data uses. Default option.
No Change	Does not apply a casing style to the address. Note: The No Change option does not guarantee that the output address will match the case of the input address. If address validation replaces an address element with an element from the reference data, the element follows the case that the reference data uses.
Upper	Writes the output address in uppercase letters.

You can also configure the casing style on the **General Settings** tab.

Parameter Usage

You can use one of the following parameters to specify the casing style:

- LOWER. Writes the output address in lowercase letters.
- MIXED. Uses the casing style in use in the destination country when it is possible to do so.
- NATIVE. Applies the casing style that the address reference data uses. Default option. Matches the **Preserved** option.
- NOCHANGE. Does not apply a casing style to the address.

- UPPER. Writes the output address in uppercase letters.

Enter the parameter value in uppercase.

Country of Origin

Identifies the country in which the address records are mailed.

Select a country from the list. The property is empty by default.

Country Type

Determines the format of the country name or abbreviation in Complete Address or Formatted Address Line port output data. The transformation writes the country name or abbreviation in the standard format of the country you select.

The following table describes the Country Type options:

Option	Country
ISO 2	ISO two-character country code
ISO 3	ISO three-character country code
ISO #	ISO three-digit country code
Abbreviation	(Reserved for future use)
CN	Canada
DA	(Reserved for future use)
DE	Germany
EN	Great Britain (default)
ES	Spain
FI	Finland
FR	France
GR	Greece
IT	Italy
JP	Japan
HU	Hungary
KR	Korea, Republic of
NL	Netherlands
PL	Poland

Option	Country
PT	Portugal
RU	Russia
SA	Saudi Arabia
SE	Sweden

Default Country

Specifies the address reference data set that the transformation uses when an address record does not identify a destination country.

Select a country from the list. Use the default option if the address records include country information. Default is None.

You can also configure the default country on the **General Settings** tab.

Parameter Usage

You can use a parameter to specify the default country. When you create the parameter, enter the ISO 3166-1 alpha-3 code for the country as the parameter value. When you enter a parameter value, use uppercase characters. For example, if all address records include country information, enter NONE.

Dual Address Priority

Determines the type of address to validate. Set the property when input address records contain more than one type of valid address data.

For example, use the property when an address record contains both post office box elements and street elements. Address validation reads the data elements that contain the type of address data that you specify. Address validation ignores any incompatible data in the address.

The following table describes the options on the Dual Address Priority property:

Option	Description
Delivery service	Validates delivery service data elements in an address, such as post office box elements.
Postal admin	Validates the address elements required by the local mail carrier. Default option.
Street	Validates street data elements in an address, such as building number elements and street name elements.

Element Abbreviation

Determines if the transformation returns the abbreviated form of an address element. You can set the transformation to return the abbreviated form if the address reference data contains abbreviations.

For example, the United States Postal Service (USPS) maintains short and long forms of many street and locality names. The short form of HUNTSVILLE BROWNSFERRY RD is HSV BROWNS FRY RD. You can select the

Element Abbreviation property when the street or locality values exceed the maximum field length that the USPS specifies.

The option is cleared by default. Set the property to ON to return the abbreviated address values. The property returns the abbreviated locality name and locality code when you use the transformation in batch mode. The property returns the abbreviated street name, locality name, and locality code when you use the transformation in certified mode.

Execution Instances

Specifies the number of threads that the Data Integration Service tries to create for the current transformation at run time. The Data Integration Service considers the Execution Instances value if you override the Maximum Parallelism run-time property on the mapping that contains the transformation. The default Execution Instances value is 1.

The Data Integration Service considers multiple factors to determine the number of threads to assign to the transformation. The principal factors are the Execution Instances value and the values on the mapping and on the associated application services in the domain.

The Data Integration Service reads the following values when it calculates the number of threads to use for the transformation:

- The *Maximum Parallelism* value on the Data Integration Service. Default is 1.
- Any *Maximum Parallelism* value that you set at the mapping level. Default is Auto.
- The *Execution Instances* value on the transformation. Default is 1.

If you override the Maximum Parallelism value at the mapping level, the Data Integration Service attempts to use the lowest value across the properties to determine the number of threads.

If you use the default Maximum Parallelism value at the mapping level, the Data Integration Service ignores the Execution Instances value.

The Data Integration Service also considers the *Max Address Object Count* property on the Content Management Service when it calculates the number of threads to create. The *Max Address Object Count* property determines the maximum number of address validation instances that can run concurrently in a mapping. The *Max Address Object Count* property value must be greater than or equal to the *Maximum Parallelism* value on the Data Integration Service.

Rules and Guidelines for the Execution Instances Property

Consider the following rules and guidelines when you set the number of execution instances:

- Multiple users might run concurrent mappings on a Data Integration Service. To calculate the correct number of threads, divide the number of central processing units that the service can access by the number of concurrent mappings.
- In PowerCenter, the *AD50.cfg* configuration file specifies the maximum number of address validation instances that can run concurrently in a mapping.
- When you use the default Execution Instances value and the default Maximum Parallelism values, the transformation operations are not partitionable.
- When you set an Execution Instances value greater than 1, you change the Address Validator transformation from a passive transformation to an active transformation.

Flexible Range Expansion

Imposes a practical limit on the number of addresses that the Address Validator transformation returns when you set the Ranges to Expand property. You can set the Ranges to Expand property and the Flexible Range Expansion property when you configure the transformation to run in suggestion list mode.

The Ranges to Expand property determines how the transformation returns address suggestions when an input address does not contain house number data. If the input address does not include contextual data, such as a full post code, the Ranges to Expand property can generate a large number of very similar addresses. The Flexible Range Expansion property restricts the number of addresses that the Ranges to Expand property generates for a single address. Set the Flexible Range Expansion property to On when you set the Ranges to Expand property to All.

The following table describes the options on the Flexible Range Expansion property:

Option	Description
On	Address validation limits the number of addresses that the Ranges to Expand property adds to the suggestion list. Default option.
Off	Address validation does not limit the number of addresses that the Ranges to Expand property adds to the suggestion list.

Note: The Address Validator transformation applies the Flexible Range Expansion property in a different way to every address that it returns to the suggestion list. The transformation does not impose a fixed limit on the number of expanded addresses in the list. The transformation also considers the Max Result Count property setting when it calculates the number of expanded addresses to include in the list.

Geocode Data Type

Determines how the Address Validator transformation calculates geocode data for an address. Geocodes are latitude and longitude coordinates.

The geocoding results that the transformation returns depend on the geocoding reference data that you install. For information about geocoding reference data, contact Informatica.

Select one of the following geocode options:

Arrival point

Returns the latitude and longitude coordinates of the entrance to a building or parcel of land. Default option.

You can select the arrival point option for addresses in the following countries:

Australia, Austria, Canada, Croatia, Denmark, Estonia, Finland, France, Germany, Hungary, Italy, Latvia, Liechtenstein, Lithuania, Luxembourg, Mexico, Monaco, Netherlands, Norway, Poland, Slovakia, Slovenia, Sweden, Switzerland, and the United States.

If you specify arrival point geocodes and the Address Validator transformation cannot return the geocodes for an address, the transformation returns interpolated geocodes.

Standard

Returns the estimated latitude and longitude coordinates of the entrance to the building or parcel of land. An estimated geocode is also called an interpolated geocode.

The Address Validator transformation uses the nearest available geocodes in the reference data to estimate the geocodes for the address.

Note: Informatica no longer issues reference data for parcel centroid or rooftop geocoding.

Parameter Usage

You can use a parameter to specify the geocode type. Enter `ARRIVAL_POINT` or `NONE`. To return the standard geocodes, enter `NONE`.

Enter the parameter value in uppercase.

Global Max Field Length

Determines the maximum number of characters on any line in the address. If the Address Validator transformation writes an output address line that contains more characters than you specify, the transformation abbreviates the address elements on the line.

Use the property to control the line length in the address. For example, the SNA standards require that an address contains no more than 38 characters on any line. If you generate addresses to the SNA standard, set the Global Max Field Length to 38.

Default is 1024.

Parameter Usage

You can use a parameter to specify the maximum number of addresses. To set the parameter value, enter an integer from 0 through 1024.

Global Preferred Descriptor

Determines the format of the building descriptor, sub-building descriptor, and street descriptor that the Address Validator transformation writes to the output data. Select a descriptor when the address reference data for the destination country contains a range of descriptors for one or more of the data elements.

The following table describes the options on the property:

Option	Description
Database	Returns the descriptor that the reference database specifies for the element in the address. If the database does not specify a descriptor for the address, the transformation copies the input value to the output address. Database is the default value.
Long	Returns the complete form of the descriptor, for example <i>Street</i> .
Preserve Input	Copies the descriptor from the input address to the output address. If the input descriptor is not a valid version of the descriptor, the transformation returns an equivalent valid descriptor from the reference database.
Short	Returns an abbreviated form of the descriptor, for example <i>St</i> .

Input Format Type

Describes the most common type of information contained in unfielded input data. Use the Input Format Type property when you connect input data to the Complete Address or Formatted Address Line ports. Select the option that best describes the information in the mapping source data.

Select one of the following options:

- All
- Address
- Organization
- Contact
- Organization/Contact
The address includes organization information and contact information.
- Organization/Dept
The address includes organization information and department information.

Default is All.

Input Format With Country

Specifies whether the input contains country data. Select the property if you connect input data to the Complete Address or Formatted Address Line input ports and if the data contains country information.

The option is cleared by default.

Line Separator

Specifies the delimiter symbol that indicates line breaks in a formatted address.

Select one of the following options:

- Assign a parameter to identify the line separator
- Carriage return (CR)
- Comma
- Line Feed (LF)
- None
- Semicolon
- Tab
- Windows New Line (CRLF)

Default is semicolon.

You can also configure the line separator on the **General Settings** tab.

Parameter Usage

You can use a parameter to specify the line separator. The parameter value is case-sensitive. Enter the parameter value in uppercase characters.

Enter one of the following values:

- CR
- CRLF

- COMMA
- LF
- PIPE
- SEMICOLON
- SPACE
- TAB

Matching Alternatives

Determines whether address validation recognizes alternative place names, such as synonyms or historical names, in an input address. The property applies to street, locality, and province data.

Note: The Matching Alternatives property does not preserve alternative names in a validated address.

The following table describes the Matching Alternatives options:

Option	Description
All	Recognizes all known alternative street names and place names. Default option.
Archives only	Recognizes historical names only. For example, address validation validates "Constantinople" as a historical version of "Istanbul."
None	Does not recognize alternative street names or place names.
Synonyms only	Recognizes synonyms and exonyms only. For example, address validation validates "Londres" as an exonym of "London."

Matching Extended Archive

Determines whether address validation returns a unique delivery point code for an out-of-date Japanese address.

The address reference data files for Japan include data for out-of-date or retired addresses alongside the current addresses for the corresponding mailboxes. When you select the Matching Extended Archive property, address validation returns the delivery point code for the current version of each address. Address validation also writes a value to the Extended Element Result Status port to indicate that the input address is out of date.

To retrieve the current address from the address reference data, enter the address code as an input element.

The following table describes the Matching Extended Archive options:

Option	Description
Off	Does not apply the property.
On	Returns the address code for the current version of an out-of-date Japanese address.

The Matching Extended Archive property uses supplementary data and address code lookup data for Japan. To apply the property in address validation, configure the transformation to run in address code lookup mode.

Matching Scope

Determines the amount of data that the transformation matches against the address reference data during address validation.

The following table describes the Matching Scope options:

Option	Description
All	Validates all selected ports. Default option.
Delivery Point	Validates building and sub-building address data in addition to data that the Street option validates.
Locality	Validates province, locality, and postcode data.
Street	Validates street address data in addition to data that the Locality option validates.

Max Result Count

Determines the maximum number of addresses that address validation can return in suggestion list mode.

You can set a maximum number in the range 1 through 100. Default is 20.

Note: Suggestion list mode performs an address check against address reference data and returns a list of addresses that are possible matches for the input address. When you verify an address in suggestion list mode, address validation returns the best matches first.

Parameter Usage

You can use a parameter to specify the maximum number of addresses. To set the parameter value, enter an integer from 0 through 100.

Mode

Determines the type of address analysis that the transformation performs. You can also configure the mode on the **General Settings** tab of the transformation.

The following table describes the Mode menu options and the corresponding parameter values that you can set:

Mode Type	Description
Address code lookup	Returns a partial address or a complete address from the reference data when you provide an address code as an input. Several countries support address codes that represent the locality, street, building, or unique mailbox for an address.
Batch	Performs address validation on the records in a data set. Batch validation focuses on address completeness and deliverability. Batch mode does not return suggestions for poor-quality addresses. Batch is the default mode.
Certified	Performs address validation on the records in a data set to the certification standards of the specified country. The certification standards require that each address identifies a unique mailbox. You can perform certified address validation on addresses in Australia, France, New Zealand, the United Kingdom, and the United States.

Mode Type	Description
Country recognition	Determines a destination country for the postal address. The transformation does not perform address validation in country recognition mode.
Interactive	Completes an incomplete valid address. When an incomplete input address matches more than one address in the reference data, the transformation returns all valid addresses up to the limit that the Max Result Count specifies.
Parse	Parses data into address fields. The transformation does not perform address validation in parse mode.
Suggestion list	Returns a list of valid addresses from the reference data when an input address contains fragmentary information. When an address fragment matches more than one address in the reference data, the transformation returns all valid addresses up to the limit that the Max Result Count specifies.

Optimization Level

Determines how the transformation matches input address data and address reference data. The property defines the type of match that the transformation must find between the input data and reference data before it can update the address record.

The following table describes the Optimization Level options:

Option	Description
Narrow	The transformation parses building numbers or house numbers from street information before it performs validation. Otherwise the transformation validates the input address elements strictly according to the input port structure. The narrow option performs the fastest address validation, but it can return less accurate results than other options.
Standard	The transformation parses multiple types of address information from the input data before it performs validation. When you select the standard option, the transformation updates an address if it can match multiple input values with the reference data. Default is Standard.
Wide	The transformation uses the standard parsing settings and performs additional parsing operations across the input data. When you select the wide option, the transformation updates an address if it can match at least one input value with the reference data. The wide option increases mapping run times.

Parameter Usage

You can use a parameter to specify the optimization level. Enter NARROW, STANDARD, or WIDE. Enter the parameter value in uppercase.

Output Format Type

Describes the most common type of information that the transformation writes on the Complete Address or Formatted Address Line output port. Select the option that best describes the data that you expect on the output port.

Select one of the following options:

- All
- Address
- Organization
- Contact
- Organization/Contact
The address includes organization information and contact information.
- Organization/Dept
The address includes organization information and department information.

Default is All.

Output Format With Country

Determines if the transformation writes country identification data to the Complete Address or Formatted Address Line output ports.

The option is cleared by default.

Preferred Language

Determines the languages in which the Address Validator transformation returns address elements when the reference data sets contain data in more than one language. You can set a preferred language for addresses in Belgium, Canada, China, Finland, Hong Kong, Ireland, Israel, Macau, Switzerland, and Taiwan.

The Address Validator transformation can return address data in the following languages:

- The default language for the address in the address reference data. The default language is the main spoken language in the region to which each address belongs.
- Any other language that the address reference data supports for an address. For example, the Belgium reference data contains address elements in Flemish, French, and German.

The address reference data might contain data for a single address element or for a complete address in multiple languages. For example, address validation can return all address elements for Ireland in the English language and can return street, locality, and province information in the Irish language. Additionally, the reference data might specify different default languages for addresses in different parts of a country. For example, in the Switzerland reference data, the default language varies from region to region between French, German, and Italian.

The following table summarizes the options that you can select on the Preferred Language property:

Option	Description
Database	Returns each address in the language that the address reference data specifies. The address reference data might specify different languages for addresses in different regions in a country. Database is the default option.
Alternative 1, Alternative 2, Alternative 3	Returns address elements in an alternative language from the reference data. The alternative languages depends on the country to which the address belongs.
English	Returns address elements in English when the reference data contains the data in English. Returns the other address elements in the default language of the region to which the address belongs.
Preserve Input	Returns the address information in the input language. Address validation preserves the language if the reference data contains the address information in the input language. If address validation detects more than one supported language in the input address, it returns the address in the database language. If Address Verification cannot return an element in the input language, it returns the element in the database language.

Note: An address reference data set might contain some address elements in a non-default language but not others. If the transformation cannot find an element in the language that the property specifies, the transformation returns the element in the default language.

When you set a preferred language option, verify that the character set that the Preferred Script property specifies is compatible with the output address data that you expect.

Multilanguage Support for Belgium Addresses

The following table describes the languages that you can specify for addresses in Belgium:

Option	Description
Database	Default value. Returns addresses in the main language of the region to which the address belongs. The language might be Flemish, French, or German.
English	Returns the province, locality, and street information in English if the address reference data contains the information in English. Returns the other address elements in the main language of the region to which the address belongs.
Alternative 1	Returns the province, locality, and street information in Flemish.
Alternative 2	Returns the province, locality, and street information in French.
Alternative 3	Returns the province, locality, and street information in German.
Preserve Input	Returns the address information in the input language. Address validation preserves the language if the reference data contains the address information in the input language. If address validation detects more than one supported language in the input address, it returns the address in the database language. If Address Verification cannot return an element in the input language, it returns the element in the database language.

Multilanguage Support for Canada Addresses

The following table describes the languages that you can specify for addresses in Canada:

Option	Description
Database	Default value. Returns addresses in English for all provinces except Quebec. Returns Quebec addresses in French.
English	Returns all addresses in English.
Alternative 1	Returns all addresses in English.
Alternative 2	Returns Quebec addresses in French. In provinces other than Quebec, the transformation returns the street descriptors, directional information, and province names in French and returns other address elements in English.
Preserve Input	Returns the address information in the input language. Address validation preserves the language if the reference data contains the address information in the input language. If address validation detects more than one supported language in the input address, it returns the address in the database language. If Address Verification cannot return an element in the input language, it returns the element in the database language.

Multilanguage Support for China Addresses

The following table describes the languages that you can specify for addresses in China:

Option	Description
Database	Default value. Returns all address information in Chinese.
English	Returns the English-language versions of street descriptor and street directional values. Returns all other address information in the Chinese language. The English address elements omit transliteration elements such as "shi."
Alternative 1	Returns all address information in the database language.
Alternative 2	Returns all address information in the database language.
Alternative 3	Returns all address information in the database language.
Preserve Input	Returns the address information in the input language. Address validation preserves the language if the reference data contains the address information in the input language. If address validation detects more than one supported language in the input address, it returns the address in the database language. If Address Verification cannot return an element in the input language, it returns the element in the database language.

Consider the following rules and guidelines when you select the preferred language:

- To return the address in the Chinese language, select Database, Alternative 1, Alternative 2, or Alternative 3.
To return the address in a Chinese character set, set the Preferred Script property to Database.
- To return street descriptor and street directional information in the English language, select English.

To return the address in a Latin or ASCII character set, set the Preferred Script property to a LATIN or ASCII value.

- If you select a LATIN or ASCII value as the preferred script and Database as the preferred language, address validation returns the address data in Pinyin.

Multilanguage Support for Finland Addresses

The following table describes the languages that you can specify for addresses in Finland:

Option	Description
Database	Default value. Returns all address information in Finnish.
Alternative 1	Returns all address information in the database language.
Alternative 2	Returns the street, locality, and province information in Swedish. Returns all other information in Finnish.
Preserve Input	Returns the address information in the input language. Address validation preserves the language if the reference data contains the address information in the input language. If address validation detects more than one supported language in the input address, it returns the address in the database language. If Address Verification cannot return an element in the input language, it returns the element in the database language.

Multilanguage Support for Hong Kong Addresses

The following table describes the languages that you can specify for addresses in Hong Kong:

Option	Description
Database	Default value. Returns all address information in Chinese.
English	Returns all address information in English.
Alternative 1	Returns all address information in the database language.
Alternative 2	Returns all address information in English.
Alternative 3	Returns all address information in the database language.
Preserve Input	Returns the address information in the input language. Address validation preserves the language if the reference data contains the address information in the input language. If address validation detects more than one supported language in the input address, it returns the address in the database language. If Address Verification cannot return an element in the input language, it returns the element in the database language.

Consider the following rules and guidelines when you select the preferred language for Hong Kong:

- To return the address in a Chinese character set, set the Preferred Script property to Database.
- To return the address in a Latin or ASCII character set, set the Preferred Script property to a LATIN or ASCII value.
- The language of the input data can affect the operation of the Preserve Input option on a Hong Kong address. Address validation identifies the input language as English when the input data uses 7-bit ASCII characters and includes an English-language descriptor.

Multilanguage Support for Ireland Addresses

The following table describes the languages that you can specify for addresses in Ireland:

Option	Description
Database	Default value. Returns all address information in English.
English	Returns all address information in English.
Alternative 1	Returns all address information in English.
Alternative 2	Returns the street, locality, and county information in Irish. Returns all other address information in English.
Preserve Input	Returns the address information in the input language. Address validation preserves the language if the reference data contains the address information in the input language. If address validation detects more than one supported language in the input address, it returns the address in the database language. If Address Verification cannot return an element in the input language, it returns the element in the database language.

Multilanguage Support for Israel Addresses

The following table describes the languages that you can specify for addresses in Israel:

Option	Description
Database	Default value. Returns all address information in Hebrew.
English	Returns all address information in English.
Alternative 1	Returns all address information in Hebrew.
Alternative 2	Returns all address information in English.
Preserve Input	Returns the address information in the input language. Address validation preserves the language if the reference data contains the address information in the input language. If address validation detects more than one supported language in the input address, it returns the address in the database language. If Address Verification cannot return an element in the input language, it returns the element in the database language.

Consider the following rules and guidelines when you select the preferred language:

- To return the addresses in a Hebrew character set, set the Preferred Script property to Database.
- To return the addresses in a Latin or ASCII character set, set the Preferred Script property to a LATIN or ASCII value.
- If you select a Latin character set as the preferred script and you select Hebrew as the preferred language, address validation transliterates the Hebrew address into Latin characters. For optimal results in a Latin character set, select English as the preferred language.

Multilanguage Support for Macau Addresses

The following table describes the languages that you can specify for addresses in Macau:

Option	Description
Database	Default value. Returns all address information in Chinese.
Alternative 1	Returns all address information in the database language.
Alternative 2	Returns all address information in Portuguese.
Preserve Input	Returns the address information in the input language. Address validation preserves the language if the reference data contains the address information in the input language. If address validation detects more than one supported language in the input address, it returns the address in the database language. If Address Verification cannot return an element in the input language, it returns the element in the database language.

- To return the address in a Chinese character set, set the Preferred Script property to Database.
- To return the address in a Latin or ASCII character set, set the Preferred Script property to a LATIN or ASCII value.
- The language of the input data can affect the operation of the Preserve Input option on a Macau address. Address validation identifies the input language as Portuguese when the input data uses 7-bit ASCII characters and includes a Portuguese-language descriptor.

Multilanguage Support for Switzerland

The following table describes the languages that you can specify for addresses in Switzerland:

Option	Description
Database	Default value. Returns addresses in the main language of the region to which the address belongs. For example, address validation returns a Zurich address in German and a Geneva address in French.
English	Returns the locality and province information in English if the reference address database contains the information in English. Returns the other address elements in the main language of the region to which the address belongs. Address validation returns the locality information in English for some localities, for example Geneva and Zurich.
Alternative 1	Returns the province and locality information in German.
Alternative 2	Returns the province and locality information in French.
Alternative 3	Returns the province and locality information in Italian.
Preserve Input	Returns the address information in the input language. Address validation preserves the language if the reference data contains the address information in the input language. If address validation detects more than one supported language in the input address, it returns the address in the database language. If Address Verification cannot return an element in the input language, it returns the element in the database language.

Note: Address validation also returns street information for addresses in Biel/Bienne in the alternative language that you configure.

Multilanguage Support for Taiwan

The following table describes the languages that you can specify for addresses in Taiwan:

Option	Description
Database	Default value. Returns all address information in Chinese.
English	Returns all address information in English.
Preserve Input	Returns the address information in the input language. Address validation preserves the language if the reference data contains the address information in the input language. If address validation detects more than one supported language in the input address, it returns the address in the database language. If Address Verification cannot return an element in the input language, it returns the element in the database language.

Consider the following rules and guidelines when you select the preferred language:

- To return the address in a Chinese character set, set the Preferred Script parameter to Database.
- To return the address in a Latin or ASCII character set, set the Preferred Script parameter to a LATIN or ASCII value.
- The language of the input data can affect the operation of the Preserve Input option on a Taiwan address. Address validation identifies the input language as English when the input data uses 7-bit ASCII characters and includes an English-language descriptor.

Preferred Script

Determines the character set that the Address Validator transformation uses for output data.

The following table describes the options on the property:

Option	Description
ASCII (Simplified)	Returns an address in ASCII characters.
ASCII (Extended)	Returns an address in ASCII characters and expands any special character in an address. For example, Ö transliterates to OE.
Database	Returns an address in the character set that the address reference data uses for the default language. Database is the default value.
Latin	Returns an address in the Latin character set.
Latin (Alt)	Returns an address in an alternative Latin character set. For example, specify Latin to return a South Korea address in the Revised Romanization transliteration. Specify Latin (Alt) to return a South Korea address in the older, ISO/TR 11941 transliteration.
Postal Admin	Returns an address in the script that the postal service local to the address prefers.

Option	Description
Postal Admin (Alt)	Returns an address in a script that the postal service local to the address approves as an alternative script.
Preserve Input	Returns address data in the character set that the input address uses.

The transformation can process a data source that contains data in multiple languages and character sets. The transformation converts all input data to the Unicode UCS-2 character set and processes the data in the UCS-2 format. After the transformation processes the data, it converts the data in each address record to the character set that you specify in the property. The process is called transliteration.

Transliteration can use the numeric representations of each character in a character set when it converts characters for processing. Transliteration can also convert characters phonetically when there is no equivalent numeric representation of a character. If the Address Validator transformation cannot map a character to UCS-2, it converts the character to a space.

Note: If you update the preferred language or the preferred script on the transformation, verify that the language and the character code that you select are compatible.

Ranges To Expand

Determines how the Address Validator transformation returns suggested addresses for a street address that does not specify a house number. Use the property when the transformation runs in suggestion list mode.

The Address Validator transformation reads a partial or incomplete street address in suggestion list mode. The transformation compares the address to the address reference data, and it returns all similar addresses to the end user. If the input address does not contain a house number, the transformation can return one or more house number suggestions for the street. The Ranges to Expand property determines how the transformation returns the addresses.

The transformation can return the range of valid house numbers in a single address, or it can return a separate address for each valid house number. The transformation can also return an address for each number in the range from the lowest to the highest house number on the street.

The following table describes the options on the property:

Option	Description
All	Address validation returns a suggested address for every house number in the range of possible house numbers on the street.
None	Address validation returns a single address that identifies the lowest and highest house numbers in the valid range for the street.
Only with valid items	Address validation returns a suggested address for every house number that the address reference data recognizes as a deliverable address.

Note: Suggestion list mode can use other elements in the address to specify the valid range of street numbers. For example, a ZIP Code might identify the city block that contains the address mailbox. The Address Validator transformation can use the ZIP Code to identify the lowest and highest valid house numbers on the block.

If the transformation cannot determine a house number range within practical limits, the number of suggested addresses can grow to an unusable size. To restrict the number of addresses that the Ranges to Expand property generates, set the Flexible Range Expansion property to On.

Standardize Invalid Addresses

Determines if the address validation process standardizes the data values in an undeliverable address. The property applies to address records that return a Match Code status in the range I1 through I4.

When you standardize the data, you increase the likelihood that a downstream data process returns accurate results. For example, a duplicate analysis mapping might return a higher match score for two address records that present common address elements in the same format.

Address validation can standardize the following address elements:

- Street suffix elements, such as road and boulevard.
- Predirectional and postdirectional elements, such as north, south, east, and west.
- Delivery service elements, such as Post Office Box.
- Sub-building elements, such as apartment, floor, and suite.
- State or province names. Standardization returns the abbreviated forms of the names.

The following table describes the options on the property:

Option	Description
Off	Address validation does not correct data errors. Default option.
On	Address validation corrects data errors.

Parameter Usage

You can assign a parameter to specify the standardization policy for data errors. Enter OFF or ON as the parameter value. Enter the value in uppercase.

Tracing Level

Sets the amount of detail that is included in the log.

You can configure tracing levels for logs.

Configure the following property on the **Advanced** tab:

Tracing Level

Amount of detail that appears in the log for this transformation. You can choose terse, normal, verbose initialization, or verbose data. Default is normal.

INDEX

I

IBM DB2 database requirements
reference data warehouse [12](#)

M

Microsoft Azure SQL database requirements
reference data warehouse [13](#)
Microsoft SQL Server database requirements
reference data warehouse [13](#)

O

Oracle database requirements
reference data warehouse [13](#)

R

reference data warehouse
IBM DB2 database requirements [12](#)
Microsoft Azure SQL database requirements [13](#)
Microsoft SQL Server database requirements [13](#)
Oracle database requirements [13](#)