



Informatica® Data Quality
10.1

Content Installation Guide

This software and documentation contain proprietary information of Informatica LLC and are provided under a license agreement containing restrictions on use and disclosure and are also protected by copyright law. Reverse engineering of the software is prohibited. 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. This Software may be protected by U.S. and/or international Patents and other Patents Pending.

Use, duplication, or disclosure of the Software by the U.S. Government is subject to the restrictions set forth in the applicable software license agreement and as provided in DFARS 227.7202-1(a) and 227.7702-3(a) (1995), DFARS 252.227-7013(1)(ii) (OCT 1988), FAR 12.212(a) (1995), FAR 52.227-19, or FAR 52.227-14 (ALT III), as applicable.

The information in this product or documentation is subject to change without notice. If you find any problems in this product or documentation, please report them to us in writing.

Informatica, Informatica Platform, Informatica Data Services, PowerCenter, PowerCenterRT, PowerCenter Connect, PowerCenter Data Analyzer, PowerExchange, PowerMart, Metadata Manager, Informatica Data Quality, Informatica Data Explorer, Informatica B2B Data Transformation, Informatica B2B Data Exchange Informatica On Demand, Informatica Identity Resolution, Informatica Application Information Lifecycle Management, Informatica Complex Event Processing, Ultra Messaging, Informatica Master Data Management, and Live Data Map are trademarks or registered trademarks of Informatica LLC in the United States and in jurisdictions throughout the world. All 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, including without limitation: Copyright DataDirect Technologies. All rights reserved. Copyright © Sun Microsystems. All rights reserved. Copyright © RSA Security Inc. All Rights Reserved. Copyright © Ordinal Technology Corp. All rights reserved. Copyright © Aandacht c.v. All rights reserved. Copyright Genivia, Inc. All rights reserved. Copyright Isomorphic Software. All rights reserved. Copyright © Meta Integration Technology, Inc. All rights reserved. Copyright © Intalio. All rights reserved. Copyright © Oracle. All rights reserved. Copyright © Adobe Systems Incorporated. All rights reserved. Copyright © DataArt, Inc. All rights reserved. Copyright © ComponentSource. All rights reserved. Copyright © Microsoft Corporation. All rights reserved. Copyright © Rogue Wave Software, Inc. All rights reserved. Copyright © Teradata Corporation. All rights reserved. Copyright © Yahoo! Inc. All rights reserved. Copyright © Glyph & Cog, LLC. All rights reserved. Copyright © Thinkmap, Inc. All rights reserved. Copyright © Clearpace Software Limited. All rights reserved. Copyright © Information Builders, Inc. All rights reserved. Copyright © OSS Nokalva, Inc. All rights reserved. Copyright Edifecs, Inc. All rights reserved. Copyright Cleo Communications, Inc. All rights reserved. Copyright © International Organization for Standardization 1986. All rights reserved. Copyright © ej-technologies GmbH. All rights reserved. Copyright © Jaspersoft Corporation. All rights reserved. Copyright © International Business Machines Corporation. All rights reserved. Copyright © yWorks GmbH. All rights reserved. Copyright © Lucent Technologies. All rights reserved. Copyright (c) University of Toronto. All rights reserved. Copyright © Daniel Veillard. All rights reserved. Copyright © Unicode, Inc. Copyright IBM Corp. All rights reserved. Copyright © MicroQuill Software Publishing, Inc. All rights reserved. Copyright © PassMark Software Pty Ltd. All rights reserved. Copyright © LogiXML, Inc. All rights reserved. Copyright © 2003-2010 Lorenzi Davide, All rights reserved. Copyright © Red Hat, Inc. All rights reserved. Copyright © The Board of Trustees of the Leland Stanford Junior University. All rights reserved. Copyright © EMC Corporation. All rights reserved. Copyright © Flexera Software. All rights reserved. Copyright © Jinfonet Software. All rights reserved. Copyright © Apple Inc. All rights reserved. Copyright © Telerik Inc. All rights reserved. Copyright © BEA Systems. All rights reserved. Copyright © PDFlib GmbH. All rights reserved. Copyright © Orientation in Objects GmbH. All rights reserved. Copyright © Tanuki Software, Ltd. All rights reserved. Copyright © Ricebridge. All rights reserved. Copyright © Sencha, Inc. All rights reserved. Copyright © Scalable Systems, Inc. All rights reserved. Copyright © jqWidgets. All rights reserved. Copyright © Tableau Software, Inc. All rights reserved. Copyright © MaxMind, Inc. All Rights Reserved. Copyright © TMat Software s.r.o. All rights reserved. Copyright © MapR Technologies Inc. All rights reserved. Copyright © Amazon Corporate LLC. All rights reserved. Copyright © Highsoft. All rights reserved. Copyright © Python Software Foundation. All rights reserved. Copyright © BeOpen.com. All rights reserved. Copyright © CNRI. All rights reserved.

This product includes software developed by the Apache Software Foundation (<http://www.apache.org/>), and/or other software which is licensed under various versions of the Apache License (the "License"). You may obtain a copy of these Licenses at <http://www.apache.org/licenses/>. Unless required by applicable law or agreed to in writing, software distributed under these Licenses is distributed on an "AS IS" BASIS, WITHOUT WARRANTIES OR CONDITIONS OF ANY KIND, either express or implied. See the Licenses for the specific language governing permissions and limitations under the Licenses.

This product includes software which was developed by Mozilla (<http://www.mozilla.org/>), software copyright The JBoss Group, LLC, all rights reserved; software copyright © 1999-2006 by Bruno Lowagie and Paulo Soares and other software which is licensed under various versions of the GNU Lesser General Public License Agreement, which may be found at <http://www.gnu.org/licenses/lgpl.html>. The materials are provided free of charge by Informatica, "as-is", without warranty of any kind, either express or implied, including but not limited to the implied warranties of merchantability and fitness for a particular purpose.

The product includes ACE(TM) and TAO(TM) software copyrighted by Douglas C. Schmidt and his research group at Washington University, University of California, Irvine, and Vanderbilt University, Copyright (©) 1993-2006, all rights reserved.

This product includes software developed by the OpenSSL Project for use in the OpenSSL Toolkit (copyright The OpenSSL Project. All Rights Reserved) and redistribution of this software is subject to terms available at <http://www.openssl.org> and <http://www.openssl.org/source/license.html>.

This product includes Curl software which is Copyright 1996-2013, Daniel Stenberg, <daniel@haxx.se>. All Rights Reserved. Permissions and limitations regarding this software are subject to terms available at <http://curl.haxx.se/docs/copyright.html>. Permission to use, copy, modify, and distribute this software for any purpose with or without fee is hereby granted, provided that the above copyright notice and this permission notice appear in all copies.

The product includes software copyright 2001-2005 (©) MetaStuff, Ltd. All Rights Reserved. Permissions and limitations regarding this software are subject to terms available at <http://www.dom4j.org/license.html>.

The product includes software copyright © 2004-2007, The Dojo Foundation. All Rights Reserved. Permissions and limitations regarding this software are subject to terms available at <http://dojotoolkit.org/license>.

This product includes ICU software which is copyright International Business Machines Corporation and others. All rights reserved. Permissions and limitations regarding this software are subject to terms available at <http://source.icu-project.org/repos/icu/icu/trunk/license.html>.

This product includes software copyright © 1996-2006 Per Bothner. All rights reserved. Your right to use such materials is set forth in the license which may be found at <http://www.gnu.org/software/kawa/Software-License.html>.

This product includes OSSP UUID software which is Copyright © 2002 Ralf S. Engelschall, Copyright © 2002 The OSSP Project Copyright © 2002 Cable & Wireless Deutschland. Permissions and limitations regarding this software are subject to terms available at <http://www.opensource.org/licenses/mit-license.php>.

This product includes software developed by Boost (<http://www.boost.org/>) or under the Boost software license. Permissions and limitations regarding this software are subject to terms available at http://www.boost.org/LICENSE_1_0.txt.

This product includes software copyright © 1997-2007 University of Cambridge. Permissions and limitations regarding this software are subject to terms available at <http://www.pcre.org/license.txt>.

This product includes software copyright © 2007 The Eclipse Foundation. All Rights Reserved. Permissions and limitations regarding this software are subject to terms available at <http://www.eclipse.org/org/documents/epl-v10.php> and at <http://www.eclipse.org/org/documents/edl-v10.php>.

This product includes software licensed under the terms at <http://www.tcl.tk/software/tcltk/license.html>, <http://www.bosrup.com/web/overlib/?License>, <http://www.stlport.org/doc/license.html>, <http://asm.ow2.org/license.html>, <http://www.cryptix.org/LICENSE.TXT>, <http://hsqldb.org/web/hsqLicense.html>, <http://httpunit.sourceforge.net/doc/license.html>, <http://jung.sourceforge.net/license.txt>, http://www.gzip.org/zlib/zlib_license.html, <http://www.openldap.org/software/release/license.html>, <http://www.libssh2.org>, <http://slf4j.org/license.html>, <http://www.sente.ch/software/OpenSourceLicense.html>, <http://fusesource.com/downloads/license-agreements/fuse-message-broker-v-5-3-license-agreement>, <http://antlr.org/license.html>, <http://aopalliance.sourceforge.net/>, <http://www.bouncycastle.org/licence.html>, <http://www.jgraph.com/jgraphdownload.html>, <http://www.jcraft.com/jsch/LICENSE.txt>, http://jotm.objectweb.org/bsd_license.html, <http://www.w3.org/Consortium/Legal/2002/copyright-software-20021231>, <http://www.slf4j.org/license.html>, <http://nanoxml.sourceforge.net/orig/copyright.html>, <http://www.json.org/license.html>, <http://forge.ow2.org/projects/javaservice/>, <http://www.postgresql.org/about/licence.html>, <http://www.sqlite.org/copyright.html>, <http://www.tcl.tk/software/tcltk/license.html>, <http://www.jaxen.org/faq.html>, <http://www.jdom.org/docs/faq.html>, <http://www.slf4j.org/license.html>, <http://www.iodbc.org/dataspace/iodbc/wiki/IODBC/License>, <http://www.keplerproject.org/md5/license.html>, <http://www.toedter.com/en/jcalendar/license.html>, <http://www.edankert.com/bounce/index.html>, <http://www.net-snmp.org/about/license.html>, <http://www.openmdx.org/#FAQ>, http://www.php.net/license/3_01.txt, <http://srp.stanford.edu/license.txt>, <http://www.schneider.com/blowfish.html>, <http://www.jmock.org/license.html>, <http://xsom.java.net>, <http://benalman.com/about/license/>, <https://github.com/CreateJS/EaselJS/blob/master/src/easeljs/display/Bitmap.js>, <http://www.h2database.com/html/license.html#summary>, <http://jsoncpp.sourceforge.net/LICENSE>, <http://jdbc.postgresql.org/license.html>, <http://protobuf.googlecode.com/svn/trunk/src/google/protobuf/descriptor.proto>, <https://github.com/rantav/hector/blob/master/LICENSE>, <http://web.mit.edu/Kerberos/krb5-current/doc/mitK5license.html>, <http://jibx.sourceforge.net/jibx-license.html>, <https://github.com/lyokato/libgeohash/blob/master/LICENSE>, <https://github.com/hjiang/jsonxx/blob/master/LICENSE>, <https://code.google.com/p/lz4/>, <https://github.com/jedisct1/libsodium/blob/master/LICENSE>, <http://one-jar.sourceforge.net/index.php?page=documents&file=license>, <https://github.com/EsotericSoftware/kryo/blob/master/license.txt>, <http://www.scala-lang.org/license.html>, <https://github.com/tinkerpop/blueprints/blob/master/LICENSE.txt>, <http://gee.cs.oswego.edu/dl/classes/EDU/oswego/cs/dl/util/concurrent/intro.html>, <https://aws.amazon.com/ssl/>, <https://github.com/twbs/bootstrap/blob/master/LICENSE>, <https://sourceforge.net/p/xmlunit/code/HEAD/tree/trunk/LICENSE.txt>, <https://github.com/documentcloud/underscore-contrib/blob/master/LICENSE>, and <https://github.com/apache/hbase/blob/master/LICENSE.txt>.

This product includes software licensed under the Academic Free License (<http://www.opensource.org/licenses/afl-3.0.php>), the Common Development and Distribution License (<http://www.opensource.org/licenses/cddl1.php>), the Common Public License (<http://www.opensource.org/licenses/cpl1.0.php>), the Sun Binary Code License Agreement Supplemental License Terms, the BSD License (<http://www.opensource.org/licenses/bsd-license.php>), the new BSD License (<http://opensource.org/licenses/BSD-3-Clause>), the MIT License (<http://www.opensource.org/licenses/mit-license.php>), the Artistic License (<http://www.opensource.org/licenses/artistic-license-1.0>) and the Initial Developer's Public License Version 1.0 (<http://www.firebirdsql.org/en/initial-developer-s-public-license-version-1-0/>).

This product includes software copyright © 2003-2006 Joe Walnes, 2006-2007 XStream Committers. All rights reserved. Permissions and limitations regarding this software are subject to terms available at <http://xstream.codehaus.org/license.html>. This product includes software developed by the Indiana University Extreme! Lab. For further information please visit <http://www.extreme.indiana.edu/>.

This product includes software Copyright (c) 2013 Frank Balluffi and Markus Moeller. All rights reserved. Permissions and limitations regarding this software are subject to terms of the MIT license.

See patents at <https://www.informatica.com/legal/patents.html>.

DISCLAIMER: Informatica LLC provides this documentation "as is" without warranty of any kind, either express or implied, including, but not limited to, the implied warranties of noninfringement, merchantability, or use for a particular purpose. Informatica LLC does not warrant that this software or documentation is error free. The information provided in this software or documentation may include technical inaccuracies or typographical errors. The information in this software and documentation is subject to change at any time without notice.

NOTICES

This Informatica product (the "Software") includes certain drivers (the "DataDirect Drivers") from DataDirect Technologies, an operating company of Progress Software Corporation ("DataDirect") which are subject to the following terms and conditions:

1. THE DATADIRECT DRIVERS ARE PROVIDED "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESSED OR IMPLIED, INCLUDING BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NON-INFRINGEMENT.
2. IN NO EVENT WILL DATADIRECT OR ITS THIRD PARTY SUPPLIERS BE LIABLE TO THE END-USER CUSTOMER FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, CONSEQUENTIAL OR OTHER DAMAGES ARISING OUT OF THE USE OF THE ODBC DRIVERS, WHETHER OR NOT INFORMED OF THE POSSIBILITIES OF DAMAGES IN ADVANCE. THESE LIMITATIONS APPLY TO ALL CAUSES OF ACTION, INCLUDING, WITHOUT LIMITATION, BREACH OF CONTRACT, BREACH OF WARRANTY, NEGLIGENCE, STRICT LIABILITY, MISREPRESENTATION AND OTHER TORTS.

Publication Date: 2018-07-02

Table of Contents

Preface	6
Informatica Resources.	6
Informatica Network.	6
Informatica Knowledge Base.	6
Informatica Documentation.	6
Informatica Product Availability Matrixes.	7
Informatica Velocity.	7
Informatica Marketplace.	7
Informatica Global Customer Support.	7
 Chapter 1: Content Installation Overview.....	 8
Content Overview.	8
Data Quality Content Installer.	9
 Chapter 2: Installing Content.....	 11
Installation Overview.	11
Installation Prerequisites.	11
General Prerequisites.	11
Address Reference Data Prerequisites.	12
Identity Population Prerequisites.	14
Reference Table Data Prerequisites.	15
Running the Content Installer.	16
Windows Installation.	17
UNIX Installation.	17
Silent Installation.	18
Importing Rules and Mappings.	18
Updating Accelerator Content.	19
 Chapter 3: Post-Installation Steps for Address Reference Data.....	 21
Post-Installation Overview.	21
Configure the Address Reference Data Properties.	21
Review the Address Validator Transformation Advanced Settings.	22
Review the Address Reference Data File Status.	22
Address Reference Data Properties.	22
Rules and Guidelines for Address Reference Data Preload Options.	25
Address Validator Transformation Advanced Properties.	25
Alias Location.	25
Alias Street.	26
Casing Style.	26
Country of Origin.	27

Country Type.	27
Default Country.	28
Dual Address Priority.	28
Element Abbreviation.	29
Execution Instances.	29
Flexible Range Expansion.	30
Geocode Data Type.	30
Global Max Field Length.	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.	38
Ranges To Expand.	39
Standardize Invalid Addresses.	39
Tracing Level.	40
Address Reference Data File Status.	40
Index.	42

Preface

The Informatica Content Installation Guide is written for Informatica administrators who are responsible for installing prebuilt rules and reference data to Informatica Data Quality.

Informatica Resources

Informatica Network

Informatica Network hosts Informatica Global Customer Support, the Informatica Knowledge Base, and other product resources. To access Informatica Network, visit <https://network.informatica.com>.

As a member, you can:

- Access all of your Informatica resources in one place.
- Search the Knowledge Base for product resources, including documentation, FAQs, and best practices.
- View product availability information.
- 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 search Informatica Network for product resources such as documentation, how-to articles, best practices, and PAMs.

To access the Knowledge Base, visit <https://kb.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

To get the latest documentation for your product, browse the Informatica Knowledge Base at https://kb.informatica.com/_layouts/ProductDocumentation/Page/ProductDocumentSearch.aspx.

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

Informatica Product Availability Matrixes

Product Availability Matrixes (PAMs) indicate the versions of operating systems, databases, and other types of data sources and targets that a product release supports. If you are an Informatica Network member, you can access 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. Developed from the real-world experience of hundreds of data management projects, Informatica Velocity represents the collective knowledge of our consultants who have worked with organizations from around the world to plan, develop, deploy, and maintain successful data management solutions.

If you are an Informatica Network member, you can access 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 augment, extend, or enhance your Informatica implementations. By leveraging any of the hundreds of solutions from Informatica developers and partners, you can improve your productivity and speed up time to implementation on your projects. You can access Informatica Marketplace at <https://marketplace.informatica.com>.

Informatica Global Customer Support

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

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

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

If you are an Informatica Network member, you can use Online Support at <http://network.informatica.com>.

CHAPTER 1

Content Installation Overview

This chapter includes the following topics:

- [Content Overview, 8](#)
- [Data Quality Content Installer, 9](#)

Content Overview

Informatica Data Quality and PowerCenter applications can use rules and reference data to improve data accuracy and to standardize the appearance of data. Informatica uses the term *content* to collectively refer to rules and reference data.

Informatica distributes the following types of content:

Accelerators

Accelerators are content bundles that contain rules, reference tables, content sets, demonstration mappings, and demonstration data objects. Each accelerator provides solutions to common data quality issues in a country, region, or industry. The Data Quality Content installer includes the Core accelerator, which contains general data quality rules. You can purchase additional accelerators separately. 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 an 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 this data to parse potential identities from input fields.

The Content installer does not include address reference data files or identity population files. You purchase this content separately. For address reference data, you purchase an annual subscription for a specific country.

Use the Content installer executable files to install address reference data, identity population, and accelerator demonstration data. 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.

Data Quality Content Installer

The Data Quality Content installer contains installation files and the Core accelerator.

The Content installer contains the following directories:

- Accelerator_Content
- Accelerator_Sources
- Installer

Accelerator_Content Directory

The Accelerator_Content directory contains the following Core accelerator components:

Accelerator XML file

The accelerator XML file contains metadata for Model repository objects such as rules, demonstration mappings, reference data, and data objects. When you use the Developer tool to import the XML file, the Developer tool adds the objects to the Model repository.

Reference data file

The reference data file .zip file contains multiple reference data files in comma-separated DIC format. You use the Developer tool to import this .zip file as part of the accelerator XML import process. The import process converts the reference data files to database tables in the reference data database and writes metadata for the reference tables to the Model repository.

To use reference data or prebuilt rules in PowerCenter, export them as PowerCenter objects from the Informatica Data Quality Model repository.

Accelerator_Sources Directory

The Accelerator_Sources directory contains the following Core accelerator component:

Demonstration data file

The demonstration data .zip file contains comma-separated data files that demonstration mappings use as source data. You use the Content installer to install this .zip file.

Installer Directory

The Installer directory contains the following items:

Content installation files

Content installation files write reference data and data sources in the server directories on Windows and UNIX platforms. There are GUI, console, and silent installers for each supported operating system. Each content installer can also write address reference data and identity population files to the file system.

The following table lists the Windows file names:

File Name	Description
Content_installer_server.exe	Use to install content through the user interface.
SilentInstall.bat	Use to run the content installer in silent mode, for example as part of a scheduled process.
SilentInput.properties	Use to store the installation properties that <code>SilentInstall.bat</code> provides to the installer in silent mode.

The following table lists the UNIX file names:

File Name	Description
Content_installer_server.bin	Use to install content in console mode.
SilentInstall.sh	Use to run the content installer in silent mode, for example as part of a scheduled process.
SilentInput.properties	Use to store the installation properties that <code>SilentInstall.sh</code> provides to the installer in silent mode.

Installer properties file

The `SilentInput.properties` file contains the installation parameters that the silent installation process requires. Edit this file before running the silent installer.

CHAPTER 2

Installing Content

This chapter includes the following topics:

- [Installation Overview, 11](#)
- [Installation Prerequisites, 11](#)
- [Running the Content Installer, 16](#)
- [Importing Rules and Mappings, 18](#)
- [Updating Accelerator Content, 19](#)

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. Use the Content installation files to install address reference data, identity populations, and accelerator demonstration data.

When you install address reference data files and identity population files, verify that the 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. Rerun the installer to add files or update existing files.

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 following prerequisites before you install content.

General Prerequisites

You must install Informatica Data Quality or PowerCenter before you install content for each product.

You must know the paths to the files that you will install. You provide paths to compressed files and to directories that contain compressed files.

To install address reference data, identity population data, or additional accelerators, purchase this content separately.

Note: Do not select uncompressed files when you run the Content installer.

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.

Address Validator Modes and Address Reference Data

When you configure the Address Validator transformation, you select the type of address validation that the transformation performs. The validation type determines whether the transformation compares the input address to the address reference data. The validation type also specifies the types of address reference data that the transformation reads.

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

Address Code Lookup

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:

- 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.
- 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 data

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

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

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 batch 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.

Note: Informatica provides different types of geocode data. If you need arrival point or parcel centroid geocodes for addresses, you must purchase additional geocode data sets.

Interactive data

Install interactive data to find the complete valid address when an input address is incomplete or when you are uncertain about the validity of the input address.

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

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 releases address reference data through its Address Doctor division. Address Doctor works with national mail carriers to develop the address reference data. When a mail carrier updates its data records with new information, Address Doctor adds the information to the address reference data files for the country.
- Address Doctor 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 PowerCenter 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.

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:

- The Content installer writes the population files to the following directory on the Data 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 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.

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

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 SQL Server
- Oracle

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.DBTABAUTH` 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 SQL Server Database Requirements

Use the following guidelines when you set up the repository on Microsoft SQL Server:

- 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 account has CONNECT and RESOURCE privileges.
- Informatica does not support Oracle public synonyms for repository tables. Verify that public synonyms have not been created for any tables in the database.

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.

Running the Content Installer

Run the installer file to install address reference data files, identity population files, or accelerator demonstration data files. You can install the files to Data Quality or PowerCenter. Install the files on a machine that an Integration Service can access.

Run the installer whenever you download new content from Informatica. You do not need to uninstall older files before you run the installer. When you run the installer, the newer files overwrite older files that have the same names.

For example, you download United States address reference data in batch/interactive format. You run the Content installer and select the files that you downloaded. At a later date you download geocoding data for United States addresses. You run the Content installer again and select the geocoding files that you downloaded.

Each time you install address reference data, review the post-installation steps. For information about post-installation steps for address reference data, see the [“Post-Installation Overview” on page 21](#).

Windows Installation

Follow these steps to install address reference data, identity population data, or demonstration data files on a Windows machine.

1. Extract the Content installer .zip file.
2. Open the `Installer` directory and run `Content_installer_server.exe`.
The install wizard starts.
3. Enter the path to the root directory of the Informatica server installation. This may be a remote directory.
Browse to this directory if required.
4. If you are installing address reference data files, enter the path to the server directory where the installer will write these files.
Browse to this directory if required.
5. Click **Next**.
6. Browse to a compressed reference data file, or browse to a directory that contains reference data files, and click **Next**.
You can specify multiple file and directory paths.
7. Verify the pre-installation summary information and click **Install**.
The installer adds the data to your system.
8. Verify the post-installation information and click **Done**.

UNIX Installation

Follow these steps to install address reference data, identity population data, or demonstration data files on a UNIX machine.

1. Extract the Content installer .zip file.
Copy the `Installer` directory to the UNIX machine if necessary.
2. Open the `Installer` directory and run `Content_installer_server.bin`.
3. Specify the type of reference data you will install.
Enter 1 to install reference data from the Content CD.
Enter 2 for address reference or identity population data.
Enter 3 for both types of data.
4. Enter the path to the root directory of the Informatica server installation.
5. If you are installing address reference data files, enter the path to the server directory where the installer will write these files.
6. Enter the path to a compressed reference data file, or to a directory that contains reference data files.
You can enter multiple file and directory paths in a comma-separated list. Do not include spaces in the list.
7. Verify the pre-installation summary information.
The installer adds the data to your system.
8. Verify the post-installation information and exit the installer.

Silent Installation

You can run the Content installer in silent mode. You define the installation parameters in the `SilentInput.properties` file in the `Installer` directory. You distribute the directory to any user who will run the silent installer.

Users run the silent installer file from the `Installer` directory. On Windows systems, the silent installer file is `SilentInstall.bat`. On UNIX systems, the silent installer file is `SilentInstall.sh`.

Follow these steps to prepare the properties file for silent installation:

1. Open the `Installer` directory on the Content CD image.
2. Open `SilentInput.properties`.
3. Set the following properties for the Informatica domain in which you will use the reference data:

Property	Description
<code>USER_INSTALL_DIR</code>	Path to the root directory of Informatica Data Quality or PowerCenter.
<code>USER_SELECTIONS</code>	Comma-separated list of reference data files or directories. This list must not contain spaces.
<code>UID_EXTRACTION_FLAG</code>	Determines if the installer will extract the reference data on the Content CD image. Set to 1 to extract this data.
<code>AV_EXTRACTION_FLAG</code>	Determines if the installer will extract the address reference data files at the location defined in <code>AV_INSTALL_DIR</code> . Set to 1 to extract this data.
<code>AV_INSTALL_DIR</code>	The path to the directory that contains the address reference data.

4. Save the file.

Importing Rules and Mappings

Use the Developer tool to import metadata for 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**.

Updating Accelerator Content

Use the Developer tool to import the latest rules, demonstration mappings, and reference tables in an accelerator. 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.
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 you want to update in the project. The items appear under the project node.
9. In the **Target** pane, select the destination project.
10. Click **Add to Target**.

A dialog box prompts you to merge the object that you selected with the current object in the Model repository. Click **Yes**.

A dialog box asks prompts you to rename the objects due to conflicts with the Model repository object names. Click **No**.
11. Click **Auto Match to Target**.
12. In the **Resolution** section, select **Replace object in target**.
13. Click **Next**.

The Developer tool calculates the object dependencies.
14. Click **Next**.
15. Click **Browse** to add reference data. Find 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. Click the selection arrow in the **Target Connection** field, and select the reference data database.
19. Click **Finish**.

CHAPTER 3

Post-Installation Steps for Address Reference Data

This chapter includes the following topics:

- [Post-Installation Overview, 21](#)
- [Address Reference Data Properties, 22](#)
- [Address Validator Transformation Advanced Properties, 25](#)
- [Address Reference Data File Status, 40](#)

Post-Installation Overview

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

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 Data Quality domain that you connect to. Verify that the files are properly licensed, and that the file types match the processing mode 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 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 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 to 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>
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>
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 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 Location

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

A location alias is an alternative location 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 Location options:

Option	Description
Off	Disables the Alias Location property.
Official	Replaces any alternative location name or location alias with the official location name. Default option.
Preserve	Preserves a valid alternative location name or location alias. If the input location 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

Determines the character case that the transformation uses to write output 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.
Database	Applies the casing style that the address reference data uses. Default option.
Preserved	Writes the output address in the same case as the input address.
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 parameter 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 Database option on the **General Settings** tab.
- NOCHANGE. Writes the output address in the same case as the input address. Matches the Preserved option on the **General Settings** tab.
- 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

Option	Country
IT	Italy
JP	Japan
HU	Hungary
KR	Korea, Republic of
NL	Netherlands
PL	Poland
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.

You can 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:

Austria, Canada, Denmark, Finland, Germany, Hungary, Latvia, Luxembourg, Mexico, Netherlands, Norway, Slovenia, Sweden, 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.

Parcel centroid

Returns the latitude and longitude coordinates of the geographic center of a parcel of land at ground level.

You can select the parcel centroid option for addresses in the following countries:

Austria, Canada, Denmark, Finland, Germany, Hungary, Latvia, Luxembourg, Netherlands, Norway, Slovenia, Sweden, and the United States.

If you specify parcel centroid geocodes and the Address Validator transformation cannot return the geocodes for an address, the transformation does not return any geocode data.

Rooftop

Returns the latitude and longitude coordinates that identify the physical center of the building that contains the mailbox. The Address Validator transformation returns rooftop coordinates for United Kingdom addresses.

If you specify rooftop geocodes and the Address Validator transformation cannot return the geocodes for an address, the transformation does not return any geocode data.

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.

Parameter Usage

You can use a parameter to specify the geocode type. Enter `ARRIVAL_POINT`, `PARCEL_CENTROID`, `ROOFTOP`, 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.

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
- Comma
- Line Feed/LF
- Pipe
- Semicolon
- Space
- 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

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

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.

Parameter Usage

You can use a parameter to specify the analysis mode. When you define the parameter, enter the mode as the default value. The value uses uppercase characters and does not contain spaces.

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

Mode	Description	Parameter Value
Batch	Performs address validation against address reference data. Default option.	BATCH
Certified	Performs address validation against reference data to postal service certification standards.	CERTIFIED
Country recognition	Identifies input addresses that do not contain country information. Appends country information to an address when the locality or province data identifies a single country.	COUNTRYRECOGNITION
Parse	Parses data into address fields and does not perform address validation.	PARSE
Suggestion list	Performs an address check against the address reference data and generates a list of addresses that are possible matches for an input address. Use suggestion list mode with point of entry applications.	FASTCOMPLETION
Assign parameter	Assign a parameter that identifies an address validation mode.	Enter a parameter value from the current table.

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 an address when the reference data sets contain data in more than one language.

An address reference data set specifies a default language for each valid address. In some countries, the default language can vary from one region to another. For example, the transformation returns Canadian addresses in English or in French according to the province that the address identifies. The address reference data sets for Canada store address elements in both languages. Use the Preferred Language property to return the addresses in each case in the non-default 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.

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

Option	Description
Alternative 1, Alternative 2, Alternative 3	Returns each address element in the same language regardless of the default language for the country or region.
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.

Option	Description
English	Returns the locality names and the province names 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 each address in the language that the input address uses. The transformation preserves the input language when the address reference data contains the address in the input language.

If the property specifies a language that is not available in the address reference data, the transformation returns the addresses in the default language.

Alternative Addresses in Belgium, Canada, and Switzerland

By default, address validation returns addresses in the local language of each region in Belgium, Canada, and Switzerland. You can configure the Preferred Language property to use the same language or languages for all addresses within each country. Use the Alternative 1, Alternative 2, and Alternative 3 options to specify the language.

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

Option	Language
Alternative 1	Flemish
Alternative 2	French
Alternative 3	German

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

Option	Language
Alternative 1	English
Alternative 2	French
Alternative 3	Not in use for Canadian addresses.

When you select the Alternative 2 option for Canadian addresses, address validation returns the street descriptors and the province names in French. Address validation returns all other address elements in English.

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

Option	Language
Alternative 1	German
Alternative 2	French
Alternative 3	Italian

When you select the Alternative 3 option for Swiss addresses, address validation returns the locality names in Italian. Address validation returns all other address elements in the default language for the region.

Note: If you set a preferred language and preferred script for the output data, verify that the language and the character code you select are compatible.

Preferred Script

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

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. By default, the property uses the character set that the local address reference data uses.

The following table describes the options on the property:

Option	Description
ASCII (Simplified)	Returns address data in ASCII characters.
ASCII (Extended)	Returns address data in ASCII characters and expands special characters. For example, Ö transliterates to OE.
Database	Returns address data in the character set of the address reference data that applies to the address. Default is Database.
Latin	Returns address data in Latin I characters.
Latin (Alt)	Returns address data in Latin I characters with alternative transliteration.
Postal Admin	Returns address data in Latin I or ASCII characters according to the preference of the mail carrier in the destination country.
Postal Admin (Alt)	Returns address data in Latin I or ASCII characters. Includes alternative characters, if the mail carrier in the destination country supports the characters.
Preserve Input	Returns address data in the character set that the input address uses.

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 set a preferred language and preferred script for the output data, 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.

Address Reference Data File Status

Use the Developer tool to review the status of the address reference data files on the domain. The status information includes the license expiry date for each file and the type of processing that you can perform with the file.

Use the **Preferences** window in the Developer tool to review the address reference data file status. Select the **Content Status** option on the **Preferences** window to review the status information.

The following table describes the status properties that display by default when you select **Content Status**:

Property	Description
Country ISO	The country to which the address reference data file applies. This property shows the ISO three-character abbreviation for the country.
Expiry Date	The date on which the file is due to be replaced by a newer file. You can use an address reference data file after its expiry date, but the data in the file may no longer be accurate.

Property	Description
Country Type	The type of address processing that you can perform with the data. You select the processing type in the Mode option on the General Settings tab. If the mode you select does not correspond to an address data file on the domain, the address validation mapping will fail.
Unlock Expiry Date	The date on which the file license expires. You cannot use the file after the unlock expiry date.
Unlock Start Date	The date on which the license is effective for the file. You cannot use the file before the unlock start date.

Right-click the property table to view a list of additional properties.

INDEX

I

IBM DB2 database requirements
reference data warehouse [15](#)

M

Microsoft SQL Server database requirements
reference data warehouse [16](#)

O

Oracle database requirements
reference data warehouse [16](#)

R

reference data warehouse
IBM DB2 database requirements [15](#)
Microsoft SQL Server database requirements [16](#)
Oracle database requirements [16](#)