

Informatica AddressDoctor Version 5.6.0 Release Notes November 2014

Copyright (c) 1998-2014 Informatica Corporation. All rights reserved.

Contents

Informatica AddressDoctor Installation.	1
Memory Requirements.	1
System Configuration.	2
Developer Support.	3
Informatica AddressDoctor Version 5.6.0.	3
Highlights of Informatica AddressDoctor Version 5.6.0.	3
New Features and Enhancements in Version 5.6.0.	4
Issues Resolved in Version 5.6.0.	8

This document contains important information about installation, new features, changed features, fixed limitations, and known limitations for Informatica AddressDoctor.

Informatica AddressDoctor Installation

Memory Requirements

Informatica AddressDoctor is designed to be highly efficient in its memory and resource usage. To ensure best possible performance, install Informatica AddressDoctor on a device that has fast input and output systems and sufficient memory.

The minimum RAM requirements for devices running Informatica AddressDoctor are as follows:

- 512 MB, if you want to use the address validation feature.
- 128 MB, if you want to use only the parsing feature.

Before you finalize the memory requirements, consider the size of the reference address databases that are required for your specific needs. Preloading databases significantly improves the performance of Informatica AddressDoctor. The device on which you install Informatica AddressDoctor must have sufficient RAM to preload all the required databases.

Currently, the worldwide postal reference database requires around 20 GB of storage space. If you need to validate addresses from across the world, RAM size of 20 GB or more helps you preload the entire database and thus improve the performance of Informatica AddressDoctor. If you need to preload

databases of size 3 GB or more, use a 64-bit operating system that offers you more flexibility with the RAM size. The maximum available RAM for a 32-bit operating system is 3 GB.

Tip: If full preloading of databases is not an option, use solid-state drives to store the reference address databases. Solid-state drives are faster than hard-disk drives and can significantly improve performance when multithreading is used.

System Configuration

When you install Informatica AddressDoctor, verify that the operating system and the processor architecture are compatible on the installation host machine. Verify also that the installation host machine runs a Java Development Kit that is compatible with the processor architecture and the operating system.

The following table lists the system configurations that Informatica supports for Informatica AddressDoctor:

Operating System	Processor Architecture	Java Development Kit
Windows XP Pro SP3 Windows Server 2008 SP2	x86 (32-bit)	Sun SE 7
Windows XP Pro x64 Edition SP2 Windows Server 2008 R2 Windows Server 2008 SP2 Windows Server 2012	x64 (64-bit)	Sun SE 7
SUSE Linux Enterprise Server 10 and 11	x86 (32-bit) x64 (64-bit)	Sun SE 7
RedHat Enterprise 5 and 6	x86 (32-bit) x64 (64-bit)	Sun SE 7
RedHat Enterprise 5 and 6	System z (64-bit)	IBM SE 7
AIX 5.3 AIX 6 AIX 7	POWER (64-bit)	IBM SE 7
Solaris 10 and 11	Intel (64-bit) SPARC (64-bit)	Sun SE 7
HP-UX 11	Intel Itanium (64-bit)	HP SE 5

Developer Support

Informatica develops Informatica AddressDoctor in the C++ programming language. The Informatica AddressDoctor software packages contain APIs in C and in Java.

The Informatica AddressDoctor user documentation contains examples for the C and Java APIs. You can use the examples to develop Informatica AddressDoctor implementations in other languages, such as C++, C#, Visual Basic.Net, PHP, Perl, Ruby, and Python.

Informatica AddressDoctor provides technical support for C-based and Java-based APIs. Informatica AddressDoctor does not provide implementation-specific technical support.

For more information about or assistance with address validation projects, contact the Informatica Professional Services team.

Informatica AddressDoctor Version 5.6.0

Highlights of Informatica AddressDoctor Version 5.6.0

The following table lists the new features, major enhancements, and key issues that Informatica AddressDoctor resolves in version 5.6.0.

New	Support for addresses from Taiwan in the Mandarin Traditional Chinese script.
New	Additional parameter to configure the output formats for Street, Building, and SubBuilding element descriptors in addresses in Australia and New Zealand.
New	Enhancements to Japan address validation, including support for Ban information, support for Gaiku codes, and address code lookup with a combination of Choumei Aza codes and Gaiku codes.
New	Support for IRIS codes in address validation for France.
New	Support for Address Key codes in address validation for the United Kingdom.
New	Support for seven-digit postal codes in address validation for Israel.
New	Support for SubBuilding and County information in Fast Completion mode in United States address validation.
New	Support for rooftop geocoding in address validation for the United Kingdom.
Improved	New and improved reference data for the United Arab Emirates.
Improved	Updated reference data and improved address validation for Spain.
Improved	Updated reference data and improved address validation for Turkey.
Improved	Recognition of keywords such as <i>Zimmer</i> and <i>App</i> in addresses from the DACH region.
Improved	Parsing and validation improvements for India address processing.
Improved	Support for newer versions of AIX, RedHat Linux, Solaris, and SUSE Linux operating systems.

Fixed	Informatica AddressDoctor incorrectly identifies the building number as the house number in addresses in the United Kingdom.
Fixed	Informatica AddressDoctor misinterprets chome, ban, and go information in addresses in Japan.
Fixed	Informatica AddressDoctor incorrectly identifies a country name as a locality name when the name appears on a formatted address line in a United States address.
Fixed	Informatica AddressDoctor returns an I3 status code in Interactive mode and a V4 status code in Batch mode for the same United States address.

New Features and Enhancements in Version 5.6.0

Informatica AddressDoctor introduces the following features and enhancements in version 5.6.0:

Support for Taiwan addresses in the Mandarin Traditional Chinese script

You can use Informatica AddressDoctor to validate Taiwan addresses in the Mandarin Traditional Chinese script. You can process an address in the Mandarin Traditional Chinese script as a fielded address or as a partially-fielded address. You can also process a Taiwan address in the native script as a single-line address by using the Fielded Address Line (FAL) element.

When you validate Taiwan addresses in the native script as a single-line address entry in the FAL element, you must enter the address elements in the following order:

Locality 1, Locality 2, Locality 3, Street 1, Street 1 (Supplementary Information), Street 2, Street 3, House Number, Building, SubBuilding.

Enhancements to United States address validation

Informatica AddressDoctor adds county information and sub-building information to the address output when you perform address validation in Batch mode, Certified mode, Fast Completion mode, and Interactive mode. In earlier versions, the information was added only in Batch mode, Certified mode, and Interactive mode. Informatica AddressDoctor adds the county information to the Province 2 element in the address output. Informatica AddressDoctor also adds the sub-building information to the SubBuilding element in the address output.

When you validate a United States address that contains a valid ZIP code and a valid locality, Informatica AddressDoctor adds the county to the address output. Informatica AddressDoctor can add the county information regardless of an `IX` process status for the address. Informatica AddressDoctor adds the county information to the Province 2 element. If the Province 1 element is missing from the input address, Informatica AddressDoctor also adds the province information in the Province 1 field of the output address.

When you validate a United States address that contains hyphenated house numbers, Informatica AddressDoctor moves the second part of the house number to the SubBuilding field.

Configurable output format for element descriptors

You can configure Informatica AddressDoctor to specify the output format for street, building, and sub-building element descriptors in Australia and New Zealand addresses and the *Strasse* element

in Germany addresses. To specify the output format for the element descriptors, configure one of the following values for the `GlobalPreferredDescriptor` parameter in the `parameters.xml` file:

- `DATABASE`. Returns the element descriptor that the reference database specifies for the address. If the database does not specify a descriptor for the address, Informatica AddressDoctor copies the input value to the output address. `DATABASE` is the default value.
- `LONG`. Returns the expanded form of the element descriptor in the input address.
- `SHORT`. Returns the abbreviated form of the element descriptor in the input address.
- `PRESERVE_INPUT`. Copies the element descriptor from the input address to the output address. If the input element descriptor is not an official synonym, Informatica AddressDoctor returns the corresponding value from the reference database.

Support for Address Key codes in United Kingdom Addresses

Address Keys are eight-digit numeric codes that map to addresses in the Postcode Address File from the Royal Mail. To add an Address Key to an address, set the `AddressKey` attribute for the `SupplementaryGB` element in `Results.xml` file. To include the Address Key in validated addresses, you require the United Kingdom supplementary reference database.

Extended data support for Japan

Informatica AddressDoctor can validate *Ban* or block information in a Japan address. Informatica AddressDoctor parses the Ban data from the input address to the dependent street field or the Street 2 field. A Japanese address lists the address elements in order of size, from the largest or most general unit to the smallest or most specific unit. The *Ban* element follows the *Chome* element and precedes the *Go* element in the address.

For example, in the validated output `1620062 東京都新宿区市谷加賀町1丁目 2 - 3, 1丁目` stands for Chome, 2 for Ban, and 3 for Go.

Enhancements to Japan address validation

You can configure Informatica AddressDoctor to add the Gaiku code to a valid Japan address. You can combine the new Choumei Aza code and the Gaiku code in the `AddressCodeLookup` function to find an address.

The Choumei Aza code is an 11-digit code that identifies a unique delivery point in Japan. The Gaiku code is a four-digit code that identifies the block to which an address belongs. If the output address contains both the new Choumei Aza code and the Gaiku code, the first 11 digits contain the Choumei Aza code and the last four digits contain the Gaiku code. To add the Gaiku code to an address, set the `GAIKU_CODE` attribute for the `JPSupplementary` parameter in the `Results.xml` file. To find an address with a Choumei Aza code and a Gaiku code, set the `AddressCode` parameter in the `Input.xml` file to `JPN_CHOUMEI_AZA_GAIKU_CODE`.

The Informatica AddressDoctor reference database for Japan contains the Gaiku code, the old Choumei Aza code, and the new Choumei Aza code. When you set the `MatchingExtendedArchive` parameter to `ON`, Informatica AddressDoctor writes all of the codes to the output address. The old Choumei Aza code and the new Choumei Aza code both contain 11 digits. If Informatica AddressDoctor returns different values for the old and new Choumei Aza codes, that indicates that the input address is outdated and invalid. You can use the new Choumei Aza code received from Informatica AddressDoctor to search for the current address.

For addresses that are current, Informatica AddressDoctor returns only the new Choumei Aza code and the Gaiku code. For example, in 13104004000 0020 , 13104004000 is the new Choumei Aza code and 0020 is the Gaiku code.

Support for seven-digit postal codes in Israel

Informatica AddressDoctor supports the seven-digit postal codes that Israel Post defines for addresses in Israel. The seven-digit postal codes replace the five-digit postal codes that Israel post previously defined. For example, the seven-digit postal code for Nazareth in Israel is 1623726. Previously, the postal code for Nazareth was 16237.

Enhancement to address validation in the DACH region

Informatica AddressDoctor recognizes keywords such as *Zimmer* and *App* in the House Number field for addresses from Germany, Austria, and Switzerland. Informatica AddressDoctor parses the Zimmer and App information in the House Number field as sub-building information.

For example, Informatica AddressDoctor can now parse inputs such as 10 Zimmer 13 and 10 App 10203 correctly.

Support for the INSEE-9 code in France

You can configure Informatica AddressDoctor to add the INSEE-9 code to a France address. The INSEE-9 code is also called the IRIS code. INSEE, or the National Institute for Statistics and Economic Research in France, defines IRIS codes to divide France into geographical units. France has approximately 16,000 IRIS units.

For example, Informatica AddressDoctor returns INSEE code 47001 and INSEE-9 code 470010115 in the validated output for the following address.

```
<InputData>
  <AddressElements>
    <Country Item="1" Type="NAME">FRA</Country>
    <Locality Item="1" Type="COMPLETE">AGEN</Locality>
    <PostalCode Item="1" Type="UNFORMATTED">47000</PostalCode>
    <Street Item="1" Type="COMPLETE">RUE DU PUIITS DU SAUMON</Street>
    <Number Item="1" Type="COMPLETE">6</Number>
  </AddressElements>
</InputData>
```

To add the INSEE-9 code to a France address, set the `EnrichmentSupplementaryFR` parameter in the `parameter.xml` file to `On`.

Support for rooftop geocoding in the United Kingdom

You can configure Informatica AddressDoctor to return rooftop-level geocodes for addresses in the United Kingdom. Rooftop geo-coordinates are the measured coordinates for the center of the roof for the primary building on a parcel of land, and help you pinpoint an address with the highest level of precision. Rooftop geocoding uses the `GBR5GCAP.MD` database and a separate unlock code.

To include the rooftop geo-coordinates for the U.K. addresses, set the `EnrichmentGeoCodingType` attribute of the `Process` element in `parameters.xml` to `ARRIVAL_POINT`.

Enhancements to India address parsing and validation

Informatica AddressDoctor continues to fine tune the parsing and validation processes for India addresses. Informatica AddressDoctor also refines the address validation codes for India addresses to return more accurate process status and validation codes.

Address element correction of India address elements has been standardized to use the same logic that Informatica AddressDoctor uses for correcting addresses from other countries. Informatica AddressDoctor now considers inputs that follow a different casing than what is used in the reference database as exact match and validates the addresses accordingly.

Informatica AddressDoctor does not correct delivery addressline elements such as street, building, and organization if the reference address database does not contain these details for the address. In such situations, Informatica AddressDoctor copies the input to the output and downgrades the status codes and mailability values. Such addresses typically receive a C2 or V2 process status code. However, if the delivery addressline contains some other elements that Informatica AddressDoctor could validate, Informatica AddressDoctor might return higher values.

Support for current operating systems

Informatica AddressDoctor adds support for the following operating system versions:

- AIX versions 6 and 7
- RedHat Linux version 6
- Solaris version 11
- SUSE Linux version 11

Note: Informatica AddressDoctor supports the older versions of these operating systems only until the next major release of Informatica AddressDoctor.

New and improved reference data for the United Arab Emirates

Informatica AddressDoctor delivers new and improved reference data for addresses in the United Arab Emirates. Informatica AddressDoctor offers the new reference address data for the United Arab Emirates as a premium on-demand package that requires an unlock code.

To obtain the unlock code for the premium package, contact an Informatica sales representative.

New and improved reference data for Spain

Informatica AddressDoctor delivers new and improved reference data for addresses in Spain. The reference data contains the most up-to-date address information for Spain, including sub-building-level information.

Improved data coverage and address validation for Turkey

Informatica AddressDoctor delivers improved reference data for addresses in Turkey.

Informatica AddressDoctor also introduces the following improvements to address validation for addresses in Turkey:

- Identification of the building name and street name in the Delivery Address Line 1 field.
- Addition of a slash symbol (/) between a building element and a sub-building element when the sub-building element is a number.

Improved address validation for Brazil

Informatica AddressDoctor adds the following improvements to address validation for addresses in Brazil:

- Addition of third-level sub-building information to the Delivery Address Line (DAL) and Formatted Address Line (FAL) output of a valid address. The Brazil address system contains three levels of sub-building information. Informatica AddressDoctor now returns the third level of the sub-building information in the DAL and FAL elements of valid addresses.
- Validation of kilometer (KM) information in Brazil addresses as additional information in the Street Complete element. Informatica AddressDoctor can also validate the KM information in the input address against a KM range in the reference address database. For example, if the reference database contains the KM information as 20-30 range and the input contains 25 KM, Informatica AddressDoctor marks the input as valid because 25 is within the 20-30 range.

Note: Informatica AddressDoctor uses a comma, and not a decimal point, in KM information for Brazil addresses. For example, 23,5 KM, and not 23.5KM.

Issues Resolved in Version 5.6.0

The following table lists the customer-reported issues that Informatica AddressDoctor resolves in version 5.6.0.

CR Number	Description
Brazil	
401075	Informatica AddressDoctor does not return the third level of subbuilding information in the validated address outputs.
Canada	
395236	Informatica AddressDoctor fails to recognize Unit as a sub-building descriptor if it appears in the Delivery Address Line 1 element along with the street information.
France	
313013	Informatica AddressDoctor returns incomplete and inaccurate suggestions. This issue occurs when you try to validate Rue de Char addresses in Paris.
389300, 389302	Informatica AddressDoctor returns an <code>ElementResultStatus</code> value of F for the street name element when the correct value for the element is 4.
India	
387706	Informatica AddressDoctor returns inconsistent address suggestions in Interactive mode based on the character casing in use in the input address.
357469	Informatica AddressDoctor returns inconsistent results for street name information when the street name in the input address contains a comma.
Japan	

CR Number	Description
377352	<p>Informatica AddressDoctor misinterprets the chome/ban/go information. This issue occurs in the following scenarios:</p> <ul style="list-style-type: none"> - The input address does not have cho/machi information, and the chome/ban/go information uses kanji characters. In such cases, Informatica AddressDoctor interprets chome/ban/go as cho/machi information. - The input address contains chome/ban/go information in kanji characters along with the character 〇 or ノ. In such cases, Informatica AddressDoctor interprets the chome/ban/go information as a building name.
386068	<p>Informatica AddressDoctor does not return the Choumei Aza code in valid Japan addresses. This issue occurs when you set the <code>ElementAbbreviation</code> parameter to <code>On</code> and when the Delivery Address Line 1 element contains no street information.</p>
Mexico	
396728	<p>Informatica AddressDoctor omits parts of the street information that a Delivery Address Line 1 element contains. This issue occurs when the street information in the Delivery Address Line 1 element includes a number that the reference address database does not recognize.</p>
United Kingdom	
353144	<p>Informatica AddressDoctor incorrectly splits the postal code by adding a space between the characters. This issue occurs when you use the interactive mode to validate an address that contains only a part of the postal code.</p>
390619	<p>Informatica AddressDoctor returns fewer suggestions if organization information is not included in the Fast Complete input.</p>
394932	<p>Informatica AddressDoctor incorrectly identifies a building number as a house number in some addresses.</p>
United States of America	
336094	<p>Informatica AddressDoctor fails to identify information for a single delivery point and to move the additional information to a residue element. This issue occurs when the Delivery Address Line 1 element or the Delivery Address Line 2 element contains compound suite numbers.</p>
337188	<p>Informatica AddressDoctor fails to correct Park in the street name to Park Avenue despite correct matches for house number and ZIP code.</p>
369328	<p>Informatica AddressDoctor does not expand RM in a Texas address to the standard expansions for RM. The standard expansions are Ranch Road and Ranch to Market.</p>
381239	<p>When an address contains an incorrect street name, Informatica AddressDoctor incorrectly updates the ZIP code in the address and does not update the street name. This issue occurs when the input street name incorrectly omits directional information.</p>
381242	<p>Informatica AddressDoctor returns a C4 process status code instead of a V2 process status code for an address when either of the following conditions is true:</p> <ul style="list-style-type: none"> - The house number in the input address is an alphanumeric string. - The house number in the input address is out of range for the house numbers in the specified locality.
383570	<p>Informatica AddressDoctor does not correct the ZIP code information when you set the <code>MatchingScope</code> parameter to <code>Street_Level</code> or <code>Locality_Level</code>.</p>

CR Number	Description
387529	Informatica AddressDoctor fails to recognize a house number that you entered in alphabetic characters with the street name in the Delivery Address Line 1 element. This issue occurs when you validate addresses in Certified mode.
389971	Informatica AddressDoctor fails to correct a typo in a street name and returns a mailability score of 0 for an address. This issue occurs when the street name begins with a directional term, such as North, and contains an unnecessary letter between the directional term and the rest of the street name. For example, Informatica AddressDoctor fails to correct Northhaven to Northaven.
391682	Informatica AddressDoctor returns low mailability scores for some valid addresses. This issue occurs when Informatica AddressDoctor copies the sub-building information in the input address to the output address.
392199	Informatica AddressDoctor incorrectly identifies some addresses as not valid. This issue occurs when the input address contains an incorrect ZIP code and when validation replaces street name and locality name aliases with preferred names.
393062	Informatica AddressDoctor does not change the street type from AVE to ST if an ordinal number, such as 1st or 2nd, precedes the descriptor.
393399	Informatica AddressDoctor incorrectly identifies street information as residue. This issue occurs in any of the following scenarios: <ul style="list-style-type: none"> - The Delivery Address Line 1 field in the input address contains multiple street names that are separated by the slash symbol (/). - The Delivery Address Line 1 field in the input address contains a street name that uses an abbreviated descriptor that Informatica AddressDoctor does not recognize. For example, the street name uses BO as a descriptor for Barrio. - The input address does not contain a house number.
394218	Informatica AddressDoctor adds residue information to the sub-building field. This issue occurs when the Delivery Address Line 1 field contains information that is not relevant to the field.
394744	Informatica AddressDoctor identifies the country name in a fielded address as a locality name.
394982	Informatica AddressDoctor incorrectly changes the character case of address elements. For example, Informatica AddressDoctor changes McDonald to Mcdonald and LaSalle St. to Lasalle St. This issue occurs when you set the GlobalCasing parameter to Mixed.
395470	Informatica AddressDoctor removes additional information, such as the post box number, from a Rural Route address. This issue occurs when Informatica AddressDoctor validates a Rural Route address that does not contain a ZIP code.
395557	Informatica AddressDoctor returns a mailability score of 5 for an address even when the chances of delivery are remote. This issue occurs when the address has a house number that fails delivery point validation or when the United States Postal Service identifies the ZIP code as non-deliverable.
398927	Informatica AddressDoctor fails to identify BLDV as BLVD, which is the standard abbreviation for boulevard.
398928	Informatica AddressDoctor fails to correct a street name if it can abbreviate more than one word in the street name. For example, Informatica AddressDoctor fails to correct East St North to East Street N.

CR Number	Description
400939	Informatica AddressDoctor marks a valid address as not valid if the address does not demarcate the directional information in the street name in the correct manner.
401151	Informatica AddressDoctor fails to recognize street information if the input address contains a street descriptor that the official name of the street does not contain.
401941	Informatica AddressDoctor fails to correct a unique ZIP+4 code and returns an incorrect address. This issue occurs when you validate an address that includes a ZIP code and a ZIP+4 code.
402289	<p>Informatica AddressDoctor incorrectly returns an I3 status when you validate an address and the following conditions are true:</p> <ul style="list-style-type: none"> - You set the <code>MatchingScope</code> parameter to <code>Locality_Level</code>. - You perform the address validation in Interactive mode. <p>Informatica AddressDoctor returns a V4 status code for the same address when you perform the address validation in Batch mode.</p>