

Address Verification Best Practices for India

Abstract

Address Verification provides you with a highly-effective and efficient address validation solution for India addresses. However, India addresses are very complex and pose many challenges to users who are not familiar with address formats and the address system in India. This document contains information about best practices that you can implement to optimize India address validation and to obtain the best possible results.

Supported Versions

- Informatica Address Verification 3.0 and later
- Informatica Address Verification (On-Premises) 5.x

Table of Contents

Introduction.	2
Address Structure in India Addresses.	2
Understanding the Element Priority in India Addresses.	3
Choosing the Correct Input Format.	3
Preloading India Address Databases.	4
Setting the Parameters and Configuration Options.	4
Additional Documentation Resources.	4

Introduction

India has a complex address system. Problems with India addresses vary from the lack of a proper structure to too much information, including address elements that conflict with each other. These inconsistencies in India addresses might confuse people who are not familiar with the address system in India.

In some India addresses, you might find as many as six levels of locality information. You might also find additional information, such as landmark information. All of these information types contribute to a lack of standard structure and make the addresses inconsistent.

Typically, Locality 1 in an India address is the main city, whereas the remaining levels represent suburbs, sectors, districts, and other demarcations. Some India addresses contain delimiters, such as commas, between address elements, and some addresses do not contain any delimiters.

Address Structure in India Addresses

India addresses are often unstructured and often do not follow any particular standard. A typical India address contains many or all of the following elements:

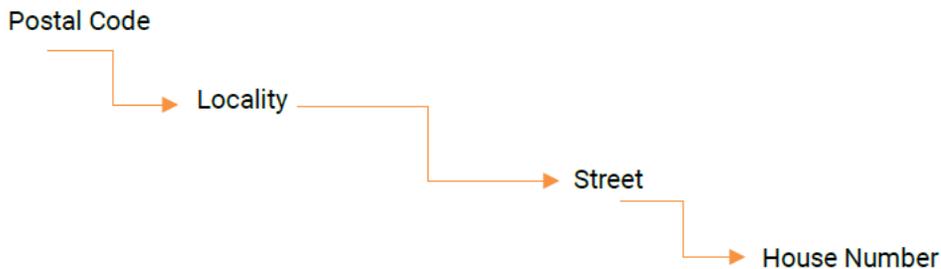
- Recipient Information, such as Contact and Organization
- House Number
- Building Number or Name, or both
- Sub-building Number or Name, or both
- Main Street
- Dependent Street

- Locality 6
- Locality 5
- Locality 4
- Locality 3
- Locality 2
- Locality 1 + Postal Code
- Province
- Country

Note: The order of the localities is from the smallest to the largest area.

Understanding the Element Priority in India Addresses

Address Verification assigns different priority levels to various elements in an India address. The following diagram shows the priorities for various elements in an address:



The address validation algorithm used by Address Verification relies on the element priority to identify the best possible matches. You receive the best results when the input address contains the maximum number of high-priority elements.

Note: To optimize India address validation results, include the high-priority elements such as postal code, locality, street name, and house or building number or name in the input address. The postal code – also known as the Postal Index Number, or PIN – is a critical input that can substantially improve the validation of India addresses.

Choosing the Correct Input Format

The best input format for India address validation is the hybrid or partially-fielded address format. You can input the address using the fielded address elements, such as Contact, Province, Locality, Country, and Postal Code, along with the AddressLines input format from input.xml.

For example, you can enter information such as the house or building number or name, street name, and PO Box in the DeliveryAddressLine (DAL) field. You can input the remaining address information, such as the locality and province information, contact, and organization, in the respective address fields.

The following examples show two India addresses in the partially-fielded address format:

```

<InputData>
  <AddressElements>
    <Country Item="1" Type="NAME">INDIA</Country>
    <Locality Item="1" Type="COMPLETE">Chennai</Locality>
    <Locality Item="2" Type="COMPLETE">Thyagaraya Nagar</Locality>
    <PostalCode Item="1" Type="UNFORMATTED">600017</PostalCode>
    <Province Item="1" Type="COUNTRY_STANDARD">Tamil Nadu</Province>
    <Organization Item="1" Type="COMPLETE">Bradur Ltd.</Organization>
    <Organization Item="1" Type="DEPARTMENT">HR Department</Organization>
  
```

```

</AddressElements>
<AddressLines>
  <DeliveryAddressLine Line="1">56 G N Chetty Road</DeliveryAddressLine>
</AddressLines>
</InputData>

<InputData>
  <AddressElements>
    <Country Item="1" Type="NAME">INDIA</Country>
    <Locality Item="1" Type="COMPLETE">Mumbai</Locality>
    <Locality Item="2" Type="COMPLETE">Worli</Locality>
    <PostalCode Item="1" Type="UNFORMATTED">400018</PostalCode>
    <Province Item="1" Type="COUNTRY_STANDARD">Maharashtra</Province>
  </AddressElements>
  <AddressLines>
    <DeliveryAddressLine Line="1">13 Mahindra Towers 2nd Floor</DeliveryAddressLine>
    <DeliveryAddressLine Line="2">Worli Road</DeliveryAddressLine>
  </AddressLines>
</InputData>

```

Consider the following rules and guidelines when you format an India address for verification:

- If you enter more than one element in a single line, add a delimiter (other than a space) between the elements. For example:

House Number, Building name, Sub-building number, Main Street

- Do not include locality, postal code, province, or country information in delivery address line input. For best results, enter the information in country specific locality line input or in discrete input fields.

For example, the following delivery address line contains information that you can move to locality and province fields:

```

<DeliveryAddressLine Line="1">13 Mahindra Towers 2nd Floor Worli Road Mumbai Maharashtra</
DeliveryAddressLine>

```

- If you cannot avoid adding this information to the delivery address line, add delimiters between the address elements. The delimiters can assist in validation when the process status scores are low.

For example, the following address appears on a single delivery address line and lacks province information:

```

<DeliveryAddressLine Line="1">112 Tcballi Varikuntapadu 524227</DeliveryAddressLine>

```

Address Verification returns an I3 process status score for the address in this format. If you add comma delimiters between the address elements, Address Verification adds the province Andhra Pradesh and returns a process status score of C2.

Preloading India Address Databases

Always do a FULL preload of the India address databases. For best results, set the value of **MaxMemoryUsageMB** in setconfig.xml to a value higher than the total memory required for the India address databases. Currently, the India address databases require about 300 MB of memory. However, if you are preloading other databases along with India address databases, add the size of those databases before setting the value of **MaxMemoryUsageMB**.

Setting the Parameters and Configuration Options

When you configure Address Verification to validate India addresses, keep the default settings for the configuration parameters. For example, keep the default values of ALL for the MatchingScope parameter and STANDARD for the OptimizationLevel parameter.

Additional Documentation Resources

To view the latest documentation for Informatica Address Verification, click the following links:

https://network.informatica.com/community/informatica-network/products/data_quality/data-as-a-service/address-doctor

https://network.informatica.com/community/informatica-network/products/data_quality/data-as-a-service/address-doctor-cloud

Author

Shahani Natalia Mendonca

Acknowledgements

The author would like to thank Amarpal Kaur Sohi and David Handy for their assistance.