ORACLE®

EDQ Address Verification Essentials

January, 2015

Copyright © 2014 Oracle and/or its affiliates. All rights reserved. |

Contents

- Part 1 – Oracle Enterprise Data Quality Address Verification Server Overview
- Part 2 – Inputs, Configuration and Outputs
- Part 3 – Understanding the AV Accuracy Code

Copyright © 2014 Oracle and/or its affiliates. All rights reserved. |

Part 1 – Oracle Enterprise Data Quality Address Verification Server Overview

Copyright © 2014 Oracle and/or its affiliates. All rights reserved. |

What Does EDQ Address Verification Do?

- Verifies that addresses exist, and to what level:
  - For example, the street may exist, but there may be no record of the premise number within that street.
- Standardizes addresses:
  - Adds missing information.
  - Outputs a standard format.
  - Where possible, corrects typos.
- Identifies addresses' locations (geocoding):
  - Attaches latitude and longitude.
- Searches:
  - Returns addresses that are similar to the one you entered.

Copyright © 2014 Oracle and/or its affiliates. All rights reserved. |

Why Verify and Standardize Addresses?

- To reduce cost:
  - Mailing discounts and first time delivery.
- To increase customer satisfaction:
  - On time deliveries.
- To ensure accurate business and marketing intelligence.
  - Examples:
    - Locate facilities near concentrations of customers.
    - Reliable delivery targets and routes.
    - Billing efficiency.
- To improve matching results.

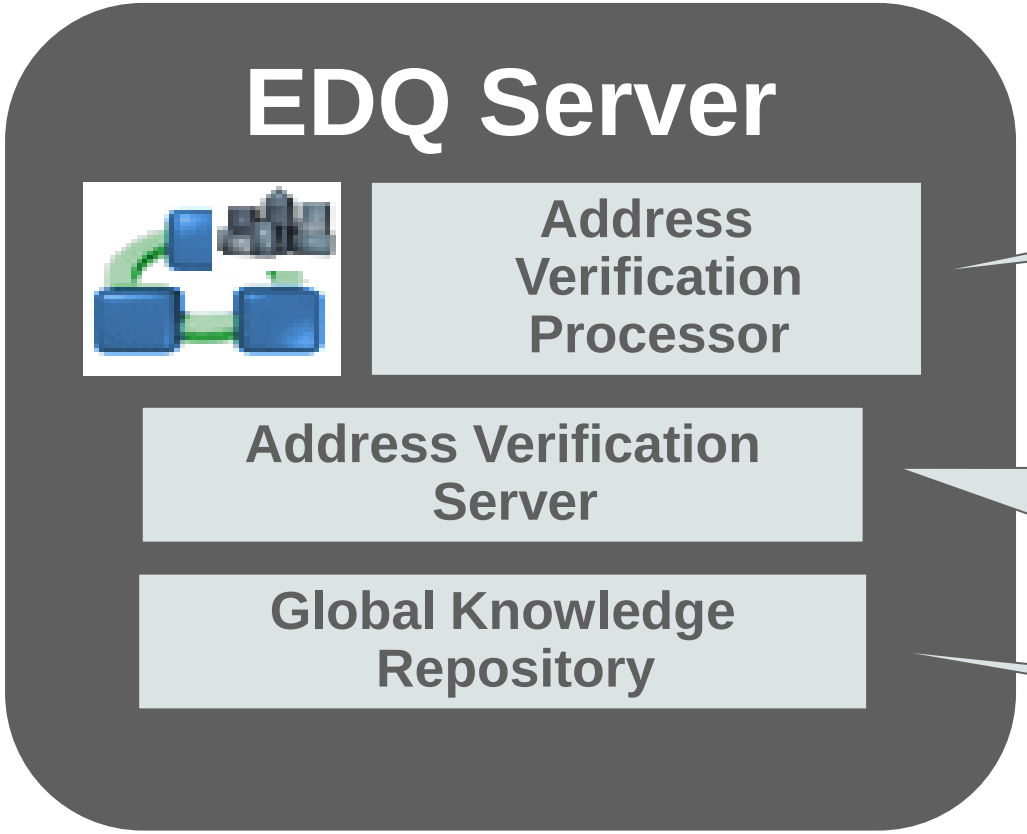
Copyright © 2014 Oracle and/or its affiliates. All rights reserved. |

Scope and Usage

- EDQ Address Verification:
  - Works worldwide.
  - Runs in real-time and batch.
  - Integrates with Enterprise Data Quality:
    - Integration is via EDQ Address Verification processor.
    - Used in EDQ Customer Data Services Pack.
    - Powerful Summarize Address Verification processor also provided.

Copyright © 2014 Oracle and/or its affiliates. All rights reserved. |

Three-Part Architecture



Can be dragged onto the EDQ Canvas. Enables you to configure inputs and options. Provided in EDQ distribution.

The brain of address verification. Consists of a set of libraries and a Java API. Must be purchased separately from Oracle.

Data library of worldwide addresses. Must be purchased from OEM partner, Loqate.

Copyright © 2014 Oracle and/or its affiliates. All rights reserved. |

What AV Does: Parse and Verify

Input Data: "9070 South Rita Road Suit 1950 Tucson US"

	Parse	Verify
Premise Number	9070	✓
Thoroughfare Name	South Rita Road	✓
Sub-Building	Suit 1950	✓
Locality	Tucson	✓
Administrative Area		
Postal Code		
Country	US	✓

Step 1 – Parse (extract) elements of the address.

Step 2 – Verify the elements against information in the Global Knowledge Repository data files.

Change character set – transliterate if necessary.

Copyright © 2014 Oracle and/or its affiliates. All rights reserved. |

What AV Does: Standardize and Geocode

Input Data: "9070 South Rita Road Suit 1950 Tucson US"

Step 3 – Standardize: add missing information and use country postal regulations to format address correctly.

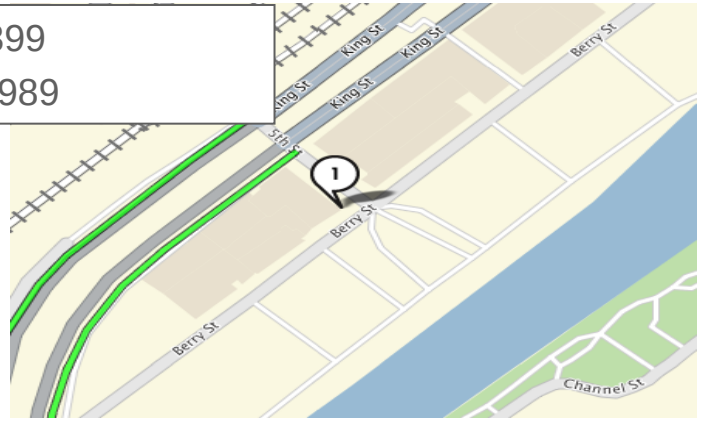
	Standardize
Premise Number	9070
Thoroughfare Name	South Rita **Rd**
Sub-Building	**Ste** 1950
Locality	Tucson
Administrative Area	**AZ**
Postal Code	**85747**
Country	US**A**

AV Whole Address
9070 South Rita Rd Ste 1950,Tucson AZ 85747

Step 4 – Find the location

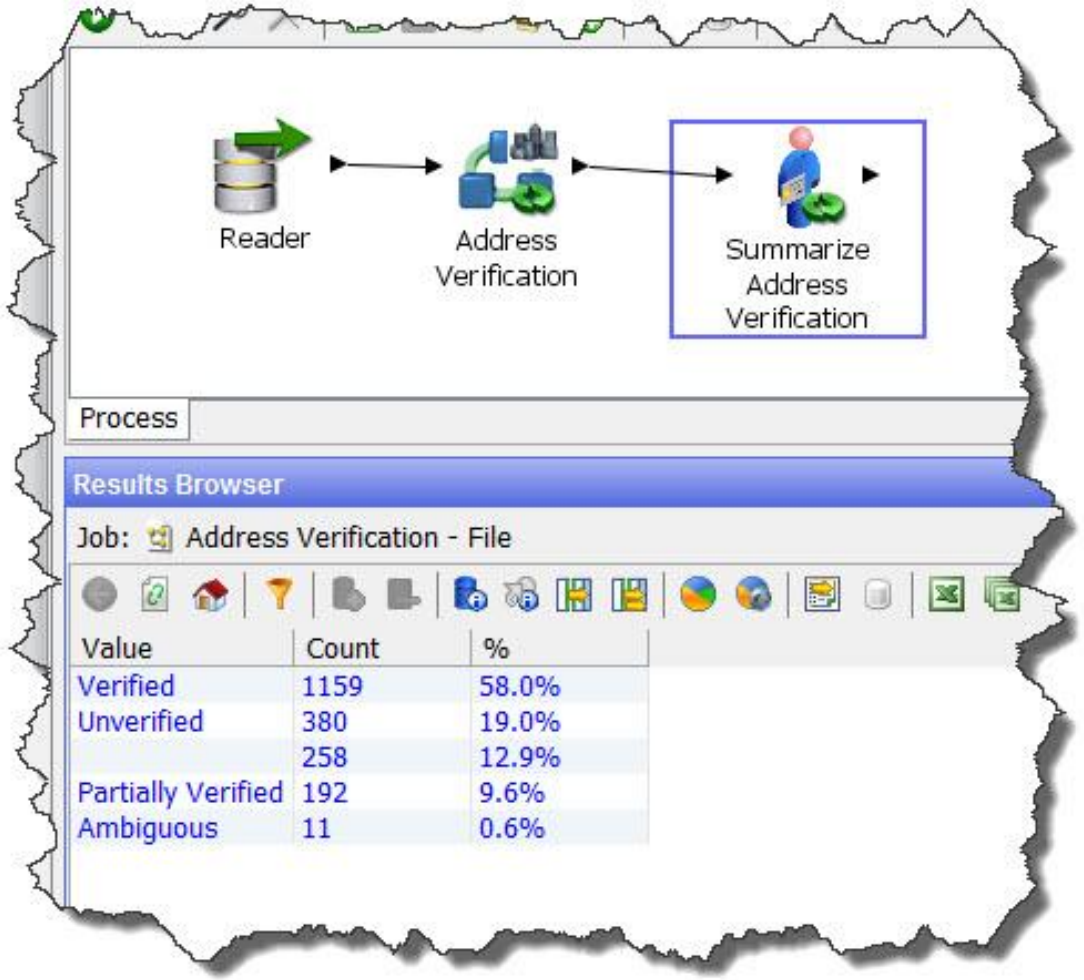
Geocode

Latitude 32.204399
Longitude -110.907989



Copyright © 2014 Oracle and/or its affiliates. All rights reserved. |

The Summarize Address Verification Processor



- Provides succinct run-down of address verification results.
  - Number of addresses verified to each level.
  - Postal code statuses.
  - And more...

Copyright © 2014 Oracle and/or its affiliates. All rights reserved. |

AV is Used in the EDQ Customer Data Services Pack (CDS)

- The Customer Data Services Pack (CDS):
  - Enables EDQ to provide the Data Quality Protection component of a customer hub or Customer Relationship Management system.
  - Includes an Address Clean service that uses Address Verification.
- The Address Clean Service:
  - Verifies input address, returning a verification code and description.
  - Geocodes the address, returning latitude and longitude co-ordinates, with additional metadata.
  - Corrects, standardizes and completes input addresses, provided the address was verified to a sufficient, configurable, level.

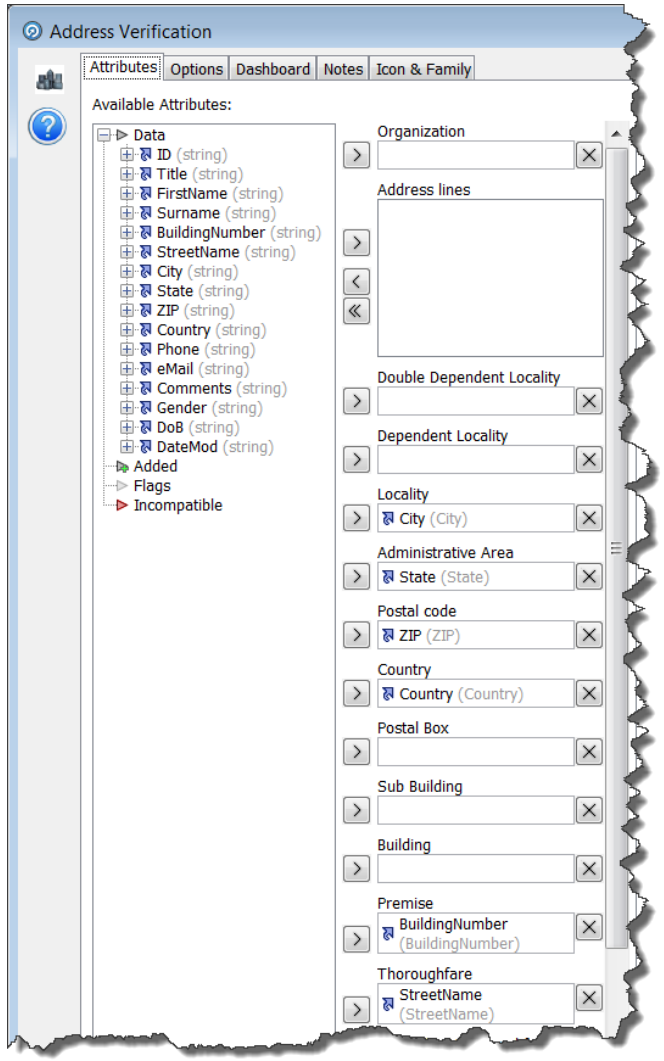
Copyright © 2014 Oracle and/or its affiliates. All rights reserved. |

Part 2 – Inputs, Configuration and Outputs

Copyright © 2014 Oracle and/or its affiliates. All rights reserved. |

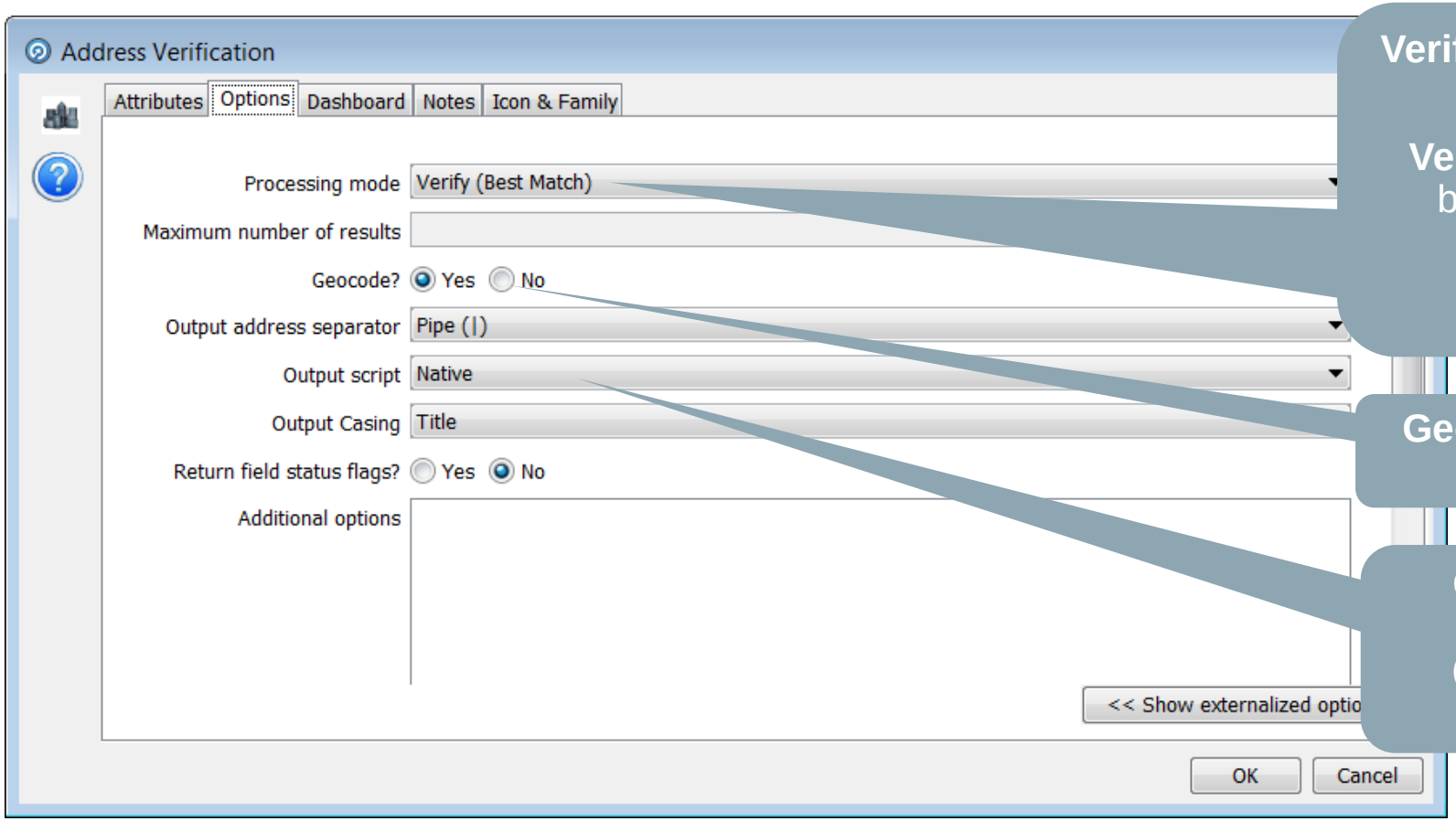
Supplying the Addresses to be Verified

- You can input:
  - Whole address in a single attribute
    - Or
  - Address elements in different attributes (Premise, Building, Thoroughfare, Locality, Administrative Area, Country etc.).
    - This can improve parsing and therefore verification results.
- Country must be input.
- Reasonably tolerant of omissions / non-standard representations.



Copyright © 2014 Oracle and/or its affiliates. All rights reserved. |

Setting the Address Verification Processor's Options



Verify (Best Match) – returns the best single match against AV data files.

Verify (Allow Multiple Results) – returns best single match if possible, or multiple matches if ambiguous.

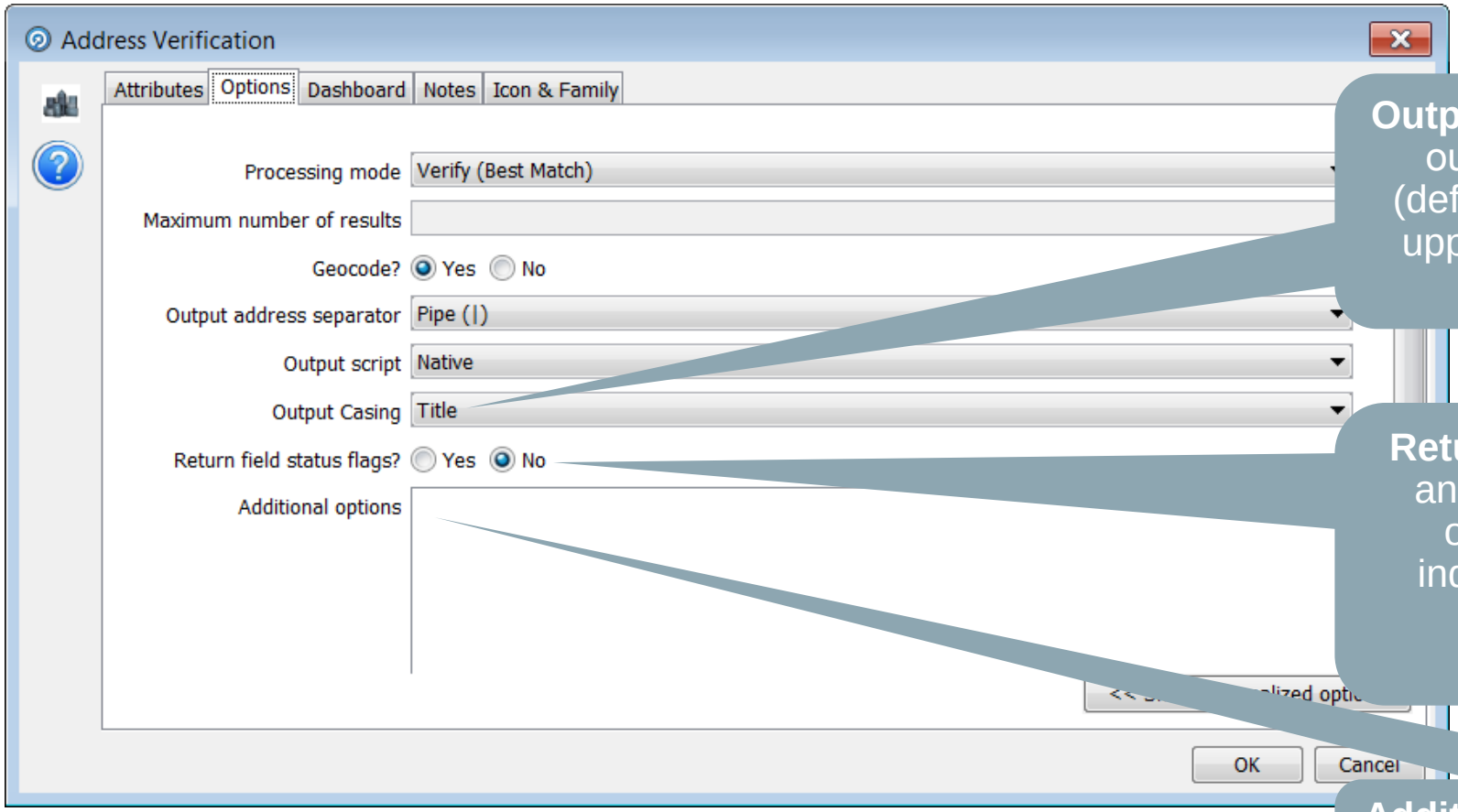
Search – returns multiple matches.

Geocode - controls whether or not to return Latitude and Longitude coordinates .

Output script – controls which writing system should be used for the output. ('Native' returns the output in its original script).

Copyright © 2014 Oracle and/or its affiliates. All rights reserved. |

Setting the Address Verification Processor's Options (Continued)



Output Casing – controls the letter case for output fields. Select from 'Upper', 'Title' (default - the first character of each word is upper case, the others are lower case), or 'Lower'.

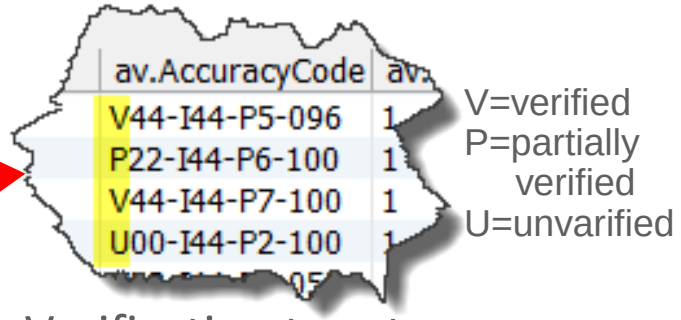
Return field status flags? – If set to Yes, an additional flag will generated for each output attribute, with a numeric value indicating the verification status of each. See EDQ's online help for more information.

Additional options – enables you to set any other AV option (see loqate support website for full details).

Copyright © 2014 Oracle and/or its affiliates. All rights reserved. |

Setting the Processing Mode

- There are three options:
  - Verify (Best Match)
    - One-to-one: Checks the reference data and returns the best match.
    - Answers the question: which elements of this address exist?



  - Verify (Allow Multiple Results)
    - Attempts to verify the input address one-to-one, but allows Address Verification to return multiple possible results if the input address has an ambiguous verification result. Results are returned as arrays.
    - You can configure the maximum number of results to be returned.
  - Search
    - One-to-many: checks the reference data and returns multiple matches in arrays.
    - You can configure the maximum number of results to be returned.
    - Answers the question: what addresses similar to this one exist?

Copyright © 2014 Oracle and/or its affiliates. All rights reserved. |

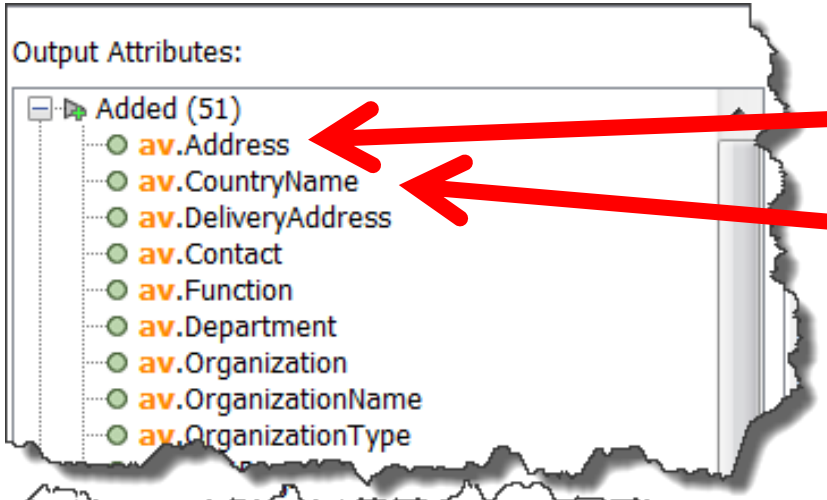
Powersearch

- Powersearch:
  - Is designed for ‘auto-complete’ usage.
    - For example:
      - A user enters the first part of an address such as “21 Holdsworth Road, South...”
      - Every few key strokes, AV is called in real-time and returns a list of matching addresses.
      - The user can select one of the matching addresses.
  - Is invoked when:
    - **Processing mode** option is set to **Search**
      and
    - Powersearch data is installed (must be purchased from loqate).
      and
    - Only **Address lines** and **Country** input fields are populated.

Copyright © 2014 Oracle and/or its affiliates. All rights reserved. |

What information does the Processor Return?

- Verified addresses and address elements.
- Flags indicating verification status and accuracy.



Standardized whole address.

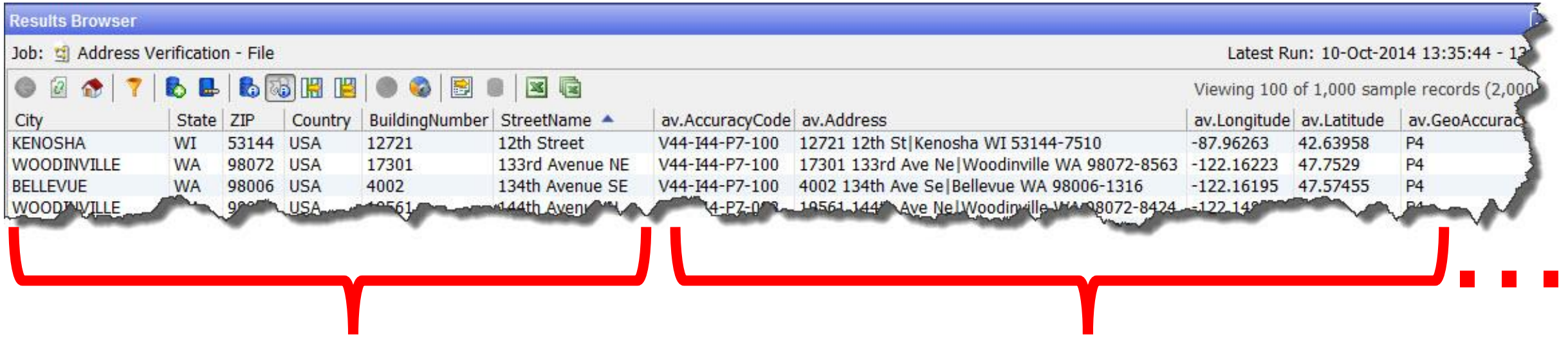
Elements of address parsed into own attributes.

Latitude and longitude returned if Geocode option is selected.

Accuracy codes output as flags.

Copyright © 2014 Oracle and/or its affiliates. All rights reserved. |

Address Verification Processor Results



Inputs

Outputs

Copyright © 2014 Oracle and/or its affiliates. All rights reserved. |

Part 3 – Understanding the Flags

Copyright © 2014 Oracle and/or its affiliates. All rights reserved. |

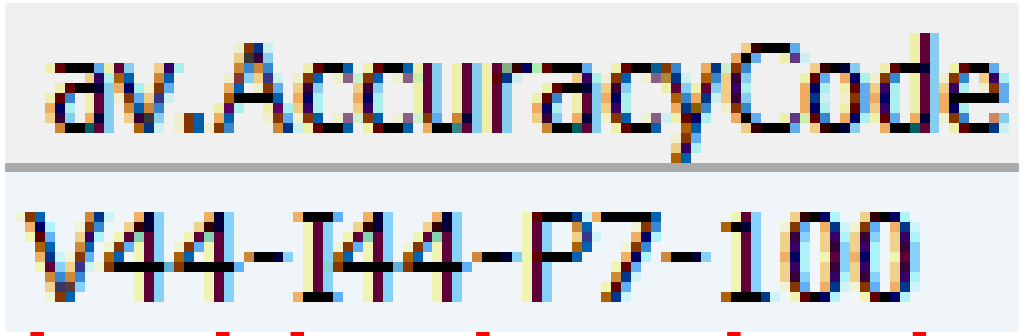
How Much Should you Trust your Address Data?

- The address verification processor generates flags for each address. There are two particularly important flags:
  - AV.AccuracyCode:
    - Tells you:
      - Whether and to what extent the address was verified.
      - How much it was changed during standardization.
  - AV.GeoAccuracy:
    - Tells you:
      - How the longitude and latitude were determined.
      - To what level they are likely to be accurate.

ORACLE

Copyright © 2014 Oracle and/or its affiliates. All rights reserved. |

The AV Accuracy Code: Four Segments



Verification status and match levels

Could the address be found in the Global Knowledge Repository?

Parsing Status and match levels

Was it possible to parse (extract) the elements of the address?

Postcode status

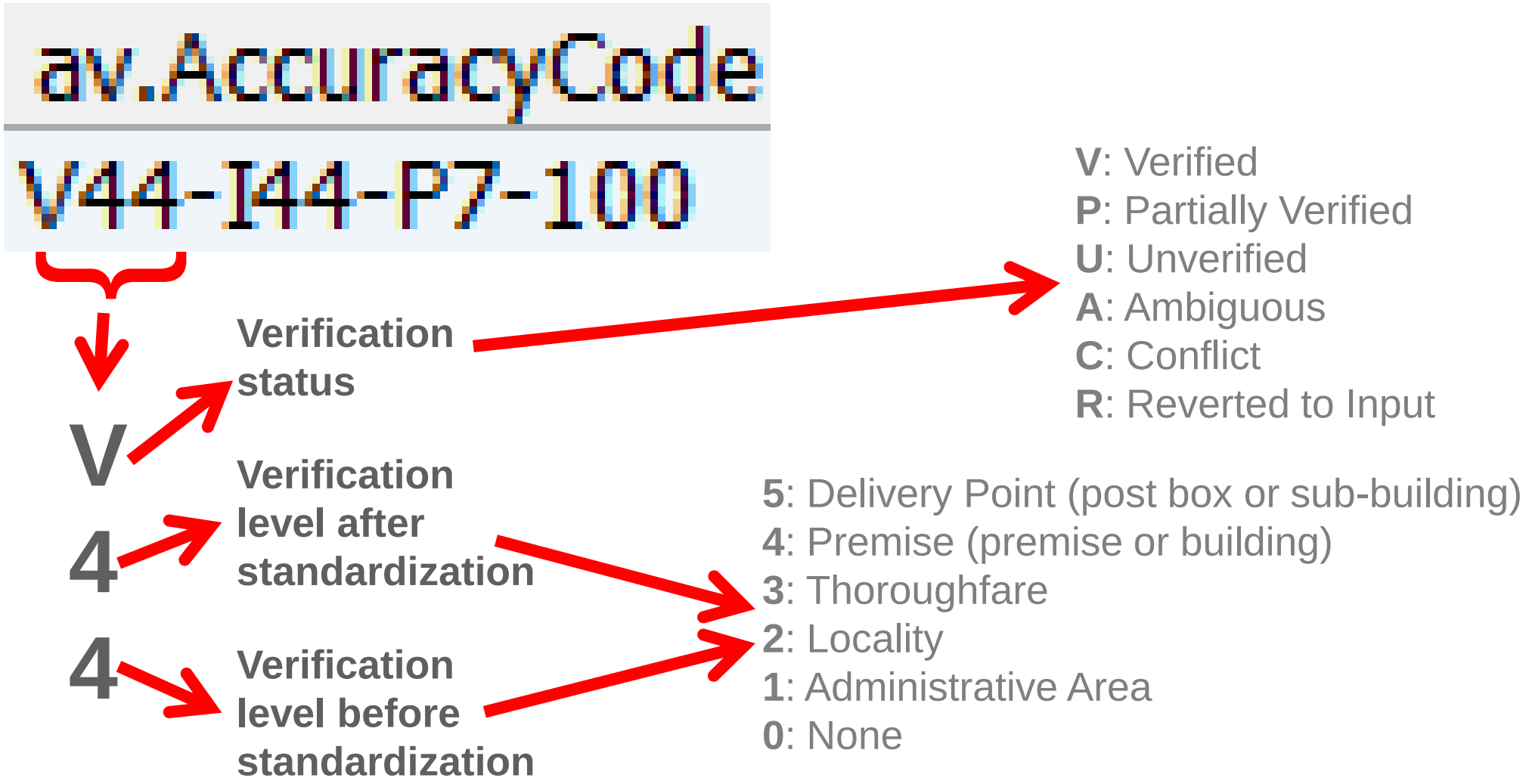
To what extent was the postal code verified?

Match score

How similar were the input and output addresses?

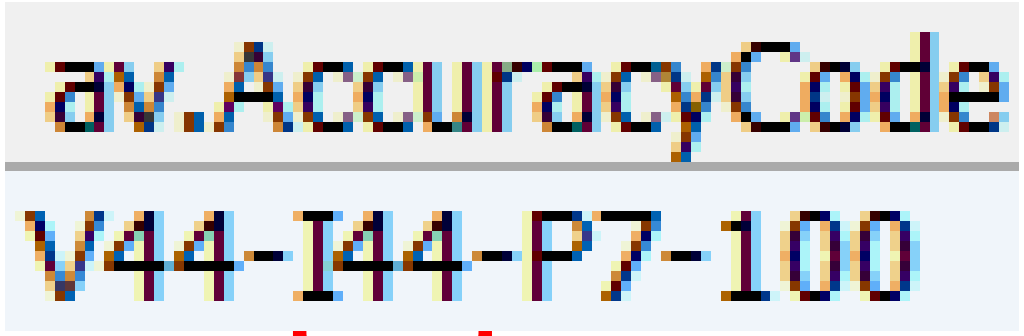
Copyright © 2014 Oracle and/or its affiliates. All rights reserved. |

Verification Status and Match Levels



Copyright © 2014 Oracle and/or its affiliates. All rights reserved. |

Parsing Status and Lexicon and Context Match Levels



Parsing status

I

Identification match – lexicon

4

Identification match – context

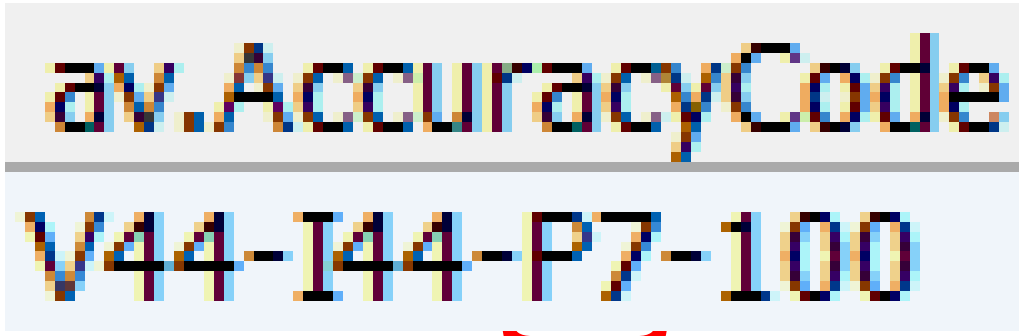
4

I: Identified and Parsed
U: Unable to Parse

5: Delivery Point (post box or sub-building)
4: Premise (premise or building)
3: Thoroughfare
2: Locality
1: Administrative Area
0: None

Copyright © 2014 Oracle and/or its affiliates. All rights reserved. |

Postal Code Status



P7 → **Postcode status** →

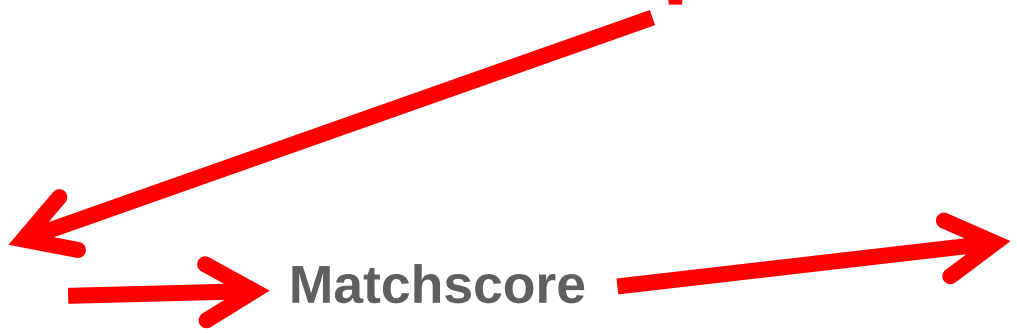
P8: PostalCodePrimary and PostalCodeSecondary verified
P7: PostalCodePrimary verified, PostalCodeSecondary added or changed
P6: PostalCodePrimary verified
P5: PostalCodePrimary verified with small change
P4: PostalCodePrimary verified with large change
P3: PostalCodePrimary added
P2: PostalCodePrimary identified by lexicon
P1: PostalCodePrimary identified by context
P0: PostalCodePrimary empty.

ORACLE

Copyright © 2014 Oracle and/or its affiliates. All rights reserved. |

Matchscore

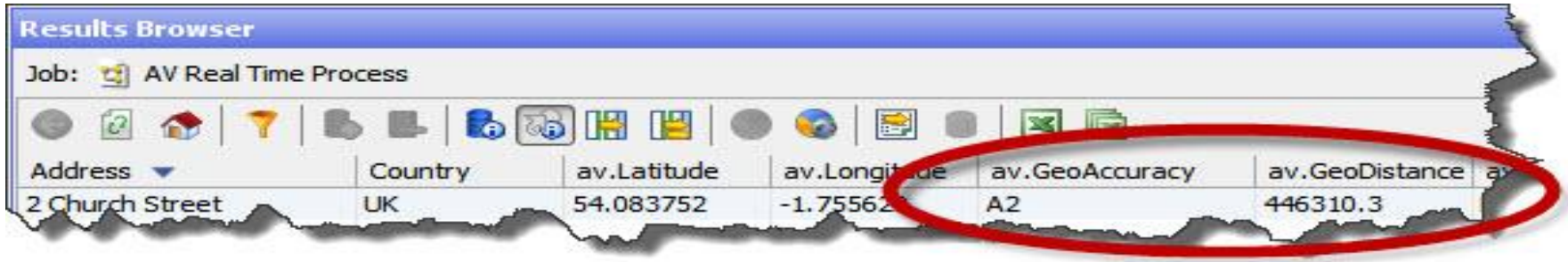
av.AccuracyCode
V44-I44-P7-100



- Similarity between input data and output data.
- Expressed as a percentage.
- 100% = complete similarity to the verification level (ignores added data).

Copyright © 2014 Oracle and/or its affiliates. All rights reserved. |

Geo Accuracy Code and Geo Distance



Geo Accuracy Code: Geocoding Status

- **P**: Point – a single geocode was found matching the input address.
- **I**: Interpolated – a geocode was interpolated from the input address's location in a range.
- **A**: Average - multiple candidate geocodes were found to match the input address, and an average of these was returned.
- **U**: Unable to geocode - not able to generate geocode for input address.

Geo Accuracy Code: Geocoding Level

- **5**: Delivery Point (PostBox or SubBuilding)
- **4**: Premise (Premise or Building)
- **3**: Thoroughfare
- **2**: Locality
- **1**: AdministrativeArea
- **0**: None

Geo Distance

Radius of accuracy in meters – indicates likely maximum distance between geocode and physical location.

Copyright © 2014 Oracle and/or its affiliates. All rights reserved. |

Using the Flags

- The AV flags provide important management information about:
  - The level to which you can trust your address data.
  - The level to which you can trust the latitude and longitude.

- The flags may influence the way in which you use your address data, including in downstream EDQ processing.

Copyright © 2014 Oracle and/or its affiliates. All rights reserved. |

Enterprise Data Quality Address Verification Self-Paced Training

- Address Verification Self-Paced Training is available from the Oracle Learning Library: https://apex.oracle.com/pls/apex/f?p=44785:24:17531673927993:PRODUCT:::P24_CONTENT_ID,P24_PREV_PAGE,P24_PROD_SECTION_GRP_ID:10244,141,3019

Copyright © 2014 Oracle and/or its affiliates. All rights reserved. |

ORACLE®