

MatchSoft

The power behind information

Matchsoft is an advanced Name and Address matching service that helps organisations achieve an accurate and consistent register of their customers in the UK & Ireland.

Outlining the Issues

Organisations typically have large and sometimes disparate homegrown databases that include duplicate and therefore inaccurate or outdated references to the same customer. This could be the result of having a variety of data input points with different validation rules or from the integration of external datasets. It could also be caused by separate departments in one organisation using different databases containing references to the same entity but in different formats; the components of an address field, for example, might have a different structure in separate datasets.

Sample Scenarios

Scenario 1:

A government agency or large corporation that has responsibility for payments based on entities that exist within a database. For example, an agency that has responsibility for making payments to medical professionals on behalf of the health service. Should the patient records become duplicated through demographic change there is the possibility that duplicate payments will follow. MatchSoft can easily identify these duplicates and flag them for removal based upon de-duplication rules that are specified by the agency. An example of a simple de-duplication rule would be to keep the most recent record and disregard the rest.

Scenario 2:

A pharmaceutical chain has 64 retail outlets throughout major cities of the UK and Ireland. Each retail outlet does its own purchasing and maintains its own list of suppliers in a format of their choice. To maximise purchasing power and keep more accurate records the company wishes

to centralise purchasing and warehousing to a central site. They wish to amalgamate their 64 supplier lists into a single database to form the basis of a new purchasing system. Each shop has in excess of 500 supplier records. MatchSoft will group the duplicate suppliers, even if the original address formats are widely different, and supply a report in CSV, HTML and/or PDF format. This can be used to automatically load the preferred supplier record of each supplier group into the new database.

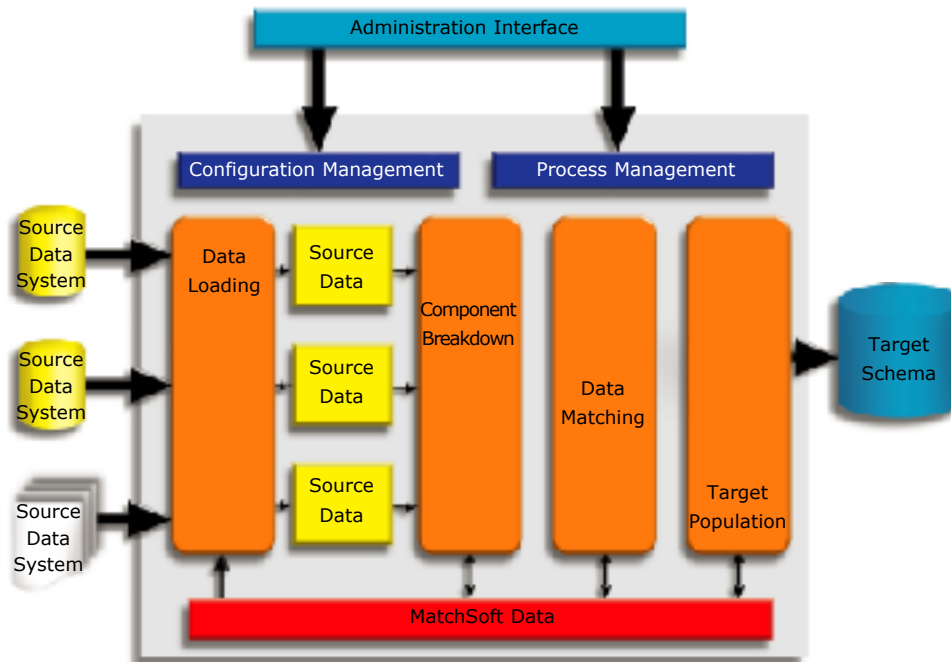
Scenario 3:

A Local Authority is concerned that it is holding data for several different services in separate MS Access databases, the address formats are typically of the address 1, 2, 3, 4 standard. The Authority would like to merge this data in a single database with a new tighter address standard imposed. The MS Access databases cumulatively hold 250,000 address records, meaning that manually changing the format of each address is time consuming and labour intensive. MatchSoft has multiple component types such as house name, house number, street, district, county, postcode etc. It can be used to break these addresses down into discrete chunks using a combination of positional intelligence and reference data. This newly formatted data may then be directly loaded to the new database thus solving the problem.

Foyle Technologies Ltd.
Level 1 Rosemount Complex
65 Rosemount Avenue
Derry City
County Londonderry BT48 0HL

Phone: +44 (0)28 71282620
Fax: +44 (0)28 71371264
Email: info@foyletech.com
Web: www.foyletech.com

How MatchSoft Works



Outlining The Solution

Equivalents / Reference Tables:

MatchSoft has a comprehensive list of reference data for UK and Irish datasets including forenames, districts, towns etc. The rationale behind this is that equivalence data may be generated such that...

- 'ST' becomes 'STREET'
- 'JIMMY' becomes 'JAMES'
- 'LTD' becomes 'LIMITED'
- 'SCH' becomes 'SCHOOL'
- 'CTGE' becomes 'COTTAGE'
- 'MUINEBEAG' becomes 'BAGENALSTOWN'
- 'MUINE BHEAG' becomes 'BAGENALSTOWN'

Common spelling mistakes and abbreviations in towns and counties may also be recognised through equivalencing. This also gives MatchSoft the advantage of recognising where it is within an address structure allowing it to make intelligent decisions on the nature of other address components, we call this 'positional intelligence' which overcomes the problem of poorly structured address data.

Spelling Inconsistencies:

A system of phonetic and soundex algorithms are

used to determine matching strings even if one or more of them has been spelt incorrectly:

String Phonetic Soundex

- Jason SN 250
- lasun SN 250
- lasen SN 250
- laysun SN 250
- Limewood MWT 530
- Lionwood NWT 530
- Linewoud NWT 530
- Crickle RKL 624
- Crickler RKL 6246

The reason we use both is to catch the 'Limewood' comparison above where it would fail a phonetic test but pass a soundex check. The reverse is true in other circumstances such as the 'Crickle' check.

Foyle Technologies Ltd.
 Level 1 Rosemount Complex
 65 Rosemount Avenue
 Derry City
 County Londonderry BT48 0HL

Phone: +44 (0)28 71282620
 Fax: +44 (0)28 71371264
 Email: info@foyletech.com
 Web: www.foyletech.com