

Address Cleansing

Modern business activities increasingly depend on reliable and accurate address data. This is particularly true for CRM and data warehousing applications where unclear structures and different formats reduce the value of data.

Organisations that want to communicate effectively with their commercial partners must take steps to ensure that their data is structured correctly as everybody identifies closely with their own address.

Addresses consist of multiple elements which all-too-often result in a variety of problems. Sometimes the capitalisation is wrong, sometimes addresses are spelt incorrectly or have the wrong abbreviation. Sometimes they are based on names which have changed or names which are in the wrong word order. More mundanely, there are often spelling &/or typing errors.

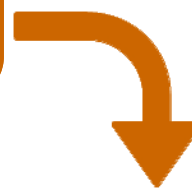
A particular problem which affects Romanian address data is missing or incorrect diacritics. Also databases which haven't kept abreast of the changes from î to â.

The following genuine examples from a single database illustrate what addresses sometimes look like and how they can be improved:



A.PRINCIPATELE UNITE 74 S.4, B
 ACADEMIEI 21, S.1, B
 AL. DECEBAL,27,IASI
 ATELIERULUI 22 A S.1
 AL.AMURGULUI,NR.8,DOROHOI,BT.
 B-DUL UNIRII NR.39 BC

Does your database look like this?



The result is easier to manage and use for the future.

Str. Principatele Unite Nr.74, Sc.4, București
 Str. Academiei Nr.21, Sc.1, București
 Al. Decebal Nr.27, Iasi
 Str. Atelierului Nr.22, Sc.A, Sector 1, București
 Al. Amurgului Nr.8, Dorohoi, Botoșani
 Bvd. Unirii Nr.39, Bacau

Reference Data

All addresses are standardised according to locality names in SIRUTA (issued and maintained by Institutul National de Statistica) and street names from Poșta Română; post-codes are also according to the Poșta Română INFOCOD standard.

However, experience shows that official sources do not contain details of all streets - especially in smaller localities. For this reason, every effort is made to match street names to other public sources where available.

Data Structure and Format

Cleansed addresses can be delivered as Excel files or Access files according to customer preference. In addition they can be provided in a number of standard formats:

- With capitals -or- proper case
- With abbreviations -or- without abbreviations
- With diacritics -or- without diacritics

A further variant relates to the street prefix: in general, the designation e.g. Aleea will be used i.e. the form of prefix that would be used to address an envelope. However, the street type (e.g. Alee) can be provided on request.

Address Enhancement

Post-codes

Post-codes can be added to all addresses which exist in the Poșta Română database. However, for streets which are not listed in the INFOCOD database, mapping sources are used to allocate streets to the correct postcodes wherever possible.

Coordinates

If required, coordinates can be added to all cleansed addresses at various levels of accuracy according to their location in Romania. The accuracy codes are:

- ADD - Address (coordinates for individual buildings)
- SSC - Street Centre Centroid
- SC - Street Centroid
- DC - District Centroid
- LC - Locality Centroid

The most accurate coordinates are applied wherever possible e.g. ADD in the larger towns, ranging down to LC in villages (satul) and hamlets (cătunul).

Geo Strategies is a licensed partner of Poșta Română.

Contact us

For more information, please contact us on:

T: +44 (0)1223 205080
F: +44 (0)1223 205081
M: +40 (0)722 244940
E: mosaic@geo-strategies.com
W: www.mosaic.geo-strategies.com