



Written in association with Allies Computing

Address validation and United Kingdom Addresses

An increasing globalization of business, new data collection channels such as the Internet and call centers and advances in data management software and computer hardware have allowed information to be stored, transferred and utilized for various purposes in greater quantities than ever before. With an increase in data storage and use, the challenges involved in data quality management increase and quality requirements tend to lag behind in the allocation of resources.

Postal address management is one of those data quality issues where attention is required and where, fortunately, solutions exist to improve address data quality. This white paper looks at the benefits of employing address validation, with particular attention to United Kingdom addresses,

How is address management beneficial?

Address management should be an integral part of the data quality program of any company holding a database, whether it is for internal or external use.

A major reason to put resources into address management is the need to prevent irritation amongst your customers. Little irritates people more when receiving a postal communication from a company than seeing their name and address details written incorrectly, in the wrong format, in the wrong language, in the wrong order and so on. It indicates to them that you do not respect the information about them that you hold and that you take a cavalier attitude to it – and, by extrapolation – have the same attitude towards your customer. A recent survey in Britain indicated that 23% of business respondents wanted direct mailers to work at improving address accuracy. When a customer receives a communication with properly formatted, validated and correct information you prevent these negative feelings.

In itself, this can increase your return on investment – a dissatisfied customer will not only be less likely to buy (or buy again) from your company, but they will often be sure to pass their dissatisfaction on to the people around them, whose purchasing decisions will similarly be affected. Data validation can save money in a number of other ways. A properly formatted address which contains all the required components is deliverable. Undelivered or delayed mail costs money. Despite the numerous resources available to companies to validate address data, Royal Mail reported in 2002 an 18% increase in the number of undeliverable postal items.

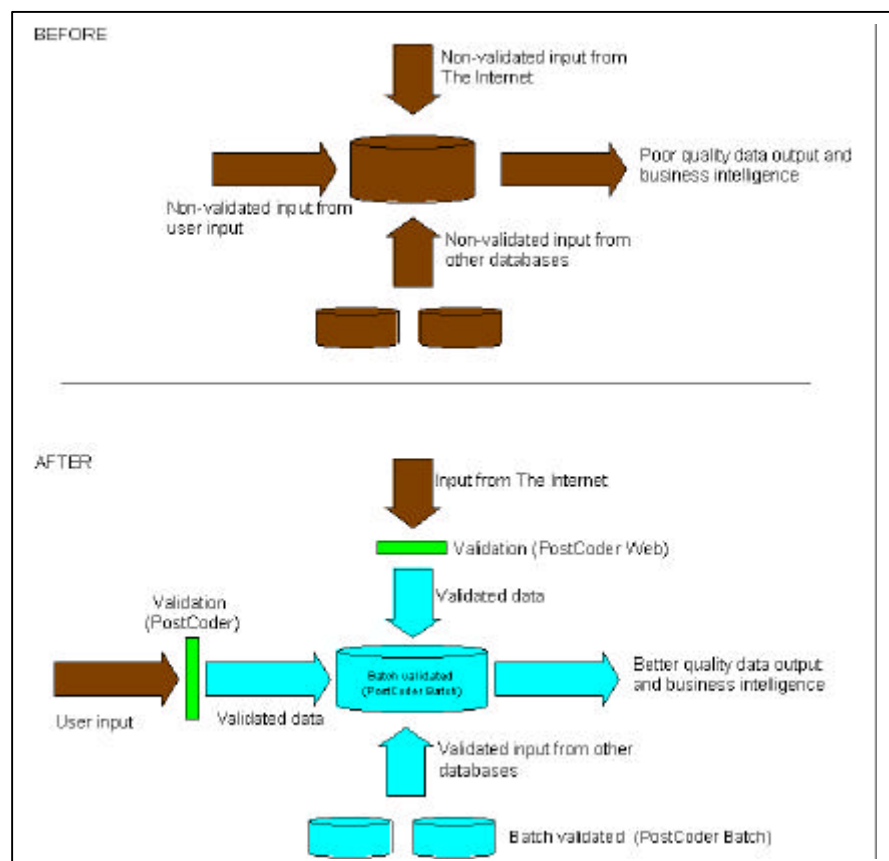
Furthermore, an address holds an important place in data processing, often being used for searching and ordering, match-key creation, de-duplication, merge-purge and so on. Inconsistency in address data and its storage causes a great deterioration in the

effectiveness of these processes, so that data quality suffers. Apart from the negative effects that receiving duplicate mailings will have on your customers, it costs you money and reduces return on investment. Equally, if you are not using your data for mailings or order fulfillment, but only as a resource for internal management information, then decisions based on data with many of duplicates or incorrect data which has been incorrectly merged or matched can have serious consequences for your company.

Good data quality, including good address data quality, improves database flexibility and consistency, accuracy and accessibility.

The Process

Address management is a continuous process. A single address validation run on a database is not sufficient for a number of reasons. Apart from a continuous input of potentially invalid addresses into the database from external sources such as the Internet, call centers or registration cards, the quality of a cleaned database without input



from other sources will deteriorate over time because in countries such as the United Kingdom, large numbers of addresses and postal codes change regularly, as new estates are built, streets are renamed, buildings are demolished and postal codes are altered to ensure approximately the same number of delivery points within each postal code area.

Address validation software, such as the PostCoder group of products from Allies Computing, ensures that address information can be data-entered and stored in an accurate and consistent manner, preventing many of the problems that arise through poor address data quality. Address validation is of particular benefit when used for United

Kingdom addresses because of their unusual complexity. This can be well demonstrated by comparing address formats used in The United Kingdom and The Netherlands

Dutch and British address structures compared

On the surface, the British and Dutch postal systems would appear to be similar. Both postal authorities use a postal code system which identifies a small area, ranging from a single user or building to a small group of habitations. In theory, in both cases, a postal code and a house number (or building or company name in the case of the United Kingdom) only are sufficient to make a postal piece deliverable. Here the similarity ends. Whilst Dutch addresses are highly standardized and formatted, British addresses are widely flexible and descriptive. Not for nothing does every person I meet working on international addresses tell me that they work first at getting their systems to work with UK data – “if I can get the UK right, I can get it to work for any country!”

In 99% of Dutch addresses the same five pieces of information are required, written always in the same order. These are the street name, the building number, a sub-building indicator (for example, a floor number, if present), a postal code and a postal town:



Though the way a user writes the data within these components can differ, for example in the casing of the components, spelling errors, non-standard abbreviations and so on, it is clear that such a simple

and structured system aids address data management to a great extent. This structure is well understood and widely used by the Dutch population, so it is always clear to both data users and data owners whether an address is complete and correctly formatted.

Addresses in the United Kingdom tell a different story. Through a mixture of historical influence, cultural inertia and a somewhat foggy policy from the postal authorities, users often write UK addresses in highly variable and descriptive ways. They may contain a mixture of components such as:

- One or more sub-building indicators
- One or more street names
- One or more building numbers
- Zonal indicators, such as industrial estate names
- Up to three settlement names
- A postal code
- An administrative area name (also for areas no longer in existence, such as Middlesex).

Thus, though strictly only the house number or name and the postal code are required for delivery, an address may look like this:

7th Floor
Unit 2
32 High Road
Seven Elms Industrial Estate
Off Ealing Broadway
Didsbury
Manchester
M23 1FK
Greater Manchester

This is a fictitious example, but it shows clearly how, with such a variable address format, there is scope for immense error in data entry, data storage and data use. This same address may be found as:

32 High Road
Manchester
M23 1FK

which is also deliverable. In fact, it may be written by different people in different sources in any number of ways with any number of possible combination of components between the 9-line version and the 3-line version above.

In data coming from differing sources, the amount of variance and therefore the possibility of error is immense. Unlike Dutch addresses, even an expert eye will have difficulty in deciding what part of an address as given is required and which is not, and, furthermore, British consumers are well known for often having individual ideas when it comes to the way that their address is written.

Clearly, then, to ensure standardized storage, optimal processing and correct output, address validation software is essential for British addresses. The first step is the batch processing, cleaning and de-duplication of existing databases, a process which needs to be repeated at regular intervals to prevent data deterioration as the postal system changes. In parallel with the batch processing, data being entered into the database from any source should be entered clean, so address validation at data entry points should be implemented. Recognizing the importance of address information in data quality processes and following these few simply steps can save immense amounts of money and improve business intelligence in a relatively short term.

About the author:

Graham Rhind is an acknowledged expert in the field of data management. He runs his own data consultancy company, GRC Database Information, in The Netherlands, (<http://www.grcdi.nl>) where he researches postal code and addressing systems, collates international data, runs a busy postal link website and writes data management software. Graham also regularly speaks on the subject and is the author of *Building and*

Maintaining a European Direct Marketing Database, The Global Sourcebook of Address Data Management and Practical International Data Management - a guide to working with global names and addresses.

Allies Computing PostCoder range includes:

[PostCoder Web S2S](#)

[PostCoder Web B2S](#)

[PostCoder API Professional](#)

[PostCoder API Compact](#)

[PostCoder Server](#)

[PostCoder Batch](#)

[PostCoder Pro](#)

[PostCoder Lite](#)

[PostCoder Local](#)

[PostCoder for SAGE Line 50](#)

[About Allies Computing](#)

End

Graham Rhind
GRC Database Information
Postbus 15213
1001 ME AMSTERDAM
020-422 93 10 (telephone)
020-422 93 11 ('fax)
graham@grcdi.nl
<http://www.grcdi.nl>