The case for address management

An international banking chain has recently been advertising its expertise in the field of local knowledge and customs. Never one to miss a challenge, I checked the company's website, particularly its feedback form. Inevitably, this form contained many of the errors that companies make in their dealings with international customers, especially with names and addresses - a single input form in a single language, inappropriate required fields, confusing field labelling and so on.

We all carry with us various personal and cultural imprints which colour the way we view the world, and this always effects localisation efforts. Even those of us who like to view ourselves as experts sometimes have problems trying to look at localisation issues with a neutral eye. Something that catches my eye, for example, is how many articles and web pages, intended for an international audience, are peppered with unexplained abbreviations and acronyms which simply do not mean the same thing to most of the readers as they do to the author. Authors know this if they stop to think about it, but that happens too rarely.

A further characteristic of globalising systems is that too often people involved in the process lose track of the required end result. I see many definitions of localisation, globalisation, internationalisation and so on as referring to the creation of software or another product which satisfies certain criteria. In my own eyes this is a fallacy - the purpose of the exercise is to aid your customers in some way, through enabling them to better understand your web site, by allowing them to enter data according to their cultural norms and so on. The software or product which is created to support this process is a means to the end, not the end in itself.

Personal names and addresses appear to be the poor relation in any internationalisation project, undeservedly so. Companies will spend fortunes creating web pages in a local language, perhaps satisfying local cultural norms in terms of layout, colour use and so on, and yet their forms for collecting names and addresses remain static and totally inappropriate for the job in hand. An appalling number of companies still, for example, insist on "state" as a required field, though a large majority of the world's countries do not have states. I have never yet found a company on the Internet which tackles the variety of name and address formats in any satisfactory way. A good deal of this problem is ignorance of the enormous variety of name and address formats which are used around the world, and the importance that these have to the local users themselves.

A failure to embrace global differences in name and address data collection forms leads to polluted and inaccurate data. As this data may be the core of both business intelligence analyses and sales and marketing efforts, any flaws in it can cause massive problems for its owners. Poor data quality leads to duplication, a waste of money and a cause of irritation amongst your customers, leading in turn to loss of custom. On a more basic level, if you are unable to correctly collect the data required for a postal service to

deliver your goods to the customer, or you collect the data in such a way that your goods are delays in the system, then clearly your business will suffer.

The world has 241 countries and territories, some 6000 different languages, around 120 different address formats and around 35 different personal name formats. It is somehow bizarre that this variety, which needs to be understood and embraced, has not been noted in the internationalisation radars of most companies.

The richness of the cultural differences between us as far as personal name and addresses are concerned can be illustrated with a few examples. On many web sites a customer is asked to provide *First name* and *Last name* data, meaning given name and family name. Often these are required fields. However, the majority of the world's population writes their name in a different order, for example with the family name first, so without proper field labelling data gathered in this way will mix up given and family names in the database. Furthermore, the inhabitants of many countries do not have family names, so providing them in a required field can be difficult!

In the same way, asking for a *prefix* to gather a form of address (Mr, Mrs etc.) and a *suffix* to gather a seniority indicator (Sr, Jr, III etc.) or an academic title (Ph.D. etc.) will create the same confusion. Whilst Americans write their form of address as a prefix and their seniority or academic title as a suffix, Germans, for example, write both their form of address and their academic titles as a prefix and the Japanese write their form of address as a suffix.

Whereas in English-speaking countries personal names are normally always written in the same way, for names in other languages this is not the case. Polish, for example, is a language that has declensions, and names may be feminised. This, together with specific freedoms allowed to people in choosing which version of their names they give, means that a married Polish woman can write her name correctly in 72 different ways! I wonder how many companies have taken account of this in their data processing…

Address structures display similar variety, not only in the order in which the data is written, but also the data contained. Addresses range from a single line for a Pacific island territory to eight or nine lines for United Kingdom addresses. Web forms with *state* and *postal code* as required fields will collect greatly polluted data from those numerous countries that have neither of these things in their addresses. The variety of numbers used within addresses is also instructive. Whereas in some countries each building has its own number, in others each apartment within a building is numbered separately. A single number is sufficient to deliver a mail piece in one country, but a whole set of numbers and symbols indicating building number, door number, staircase number and floor number are required for other countries. Whereas in the United Kingdom many streets have buildings number 1-2-3 down one side of the street, in The Netherlands it is more usual that the numbering is even (2-4-6) on one side of the street and odd (1-3-5) on the other. In many Alpine villages, without true streets, buildings are numbered as they are built, and the village name replaces the street name in the address. In Japan also buildings are numbered in the order that they are built, so the importance of

the local knowledge of the postal delivery person is paramount to the deliverability of any mail piece. Even country name collection on most web sites is not always accurate, as drop downs, based often on the flawed ISO 3166 list, do not contain all postally or culturally correct country and territory names, missing, for example, the Isle of Man and the Channel Islands, not part of the United Kingdom but too often treated as such.

A full understanding of international personal name and address issues is not only difficult to master, it also requires constant work to keep the information relevant. Countries come and go with surprising frequency, postal code systems are introduced, altered significantly or dropped in an average of four countries every year, and so on. The dynamism requires appreciation when working with international data. Address management does not begin and end with data collection, but a constant monitoring of changing postal systems, on a national and local basis, is required to ensure that your data remains accurate.

When companies have understood the richness of cultures within the world, as they pertain to data gathering, they then need to apply the knowledge to their electronic media to improve both their customer relations and their data quality. When your customers are unable to enter their details correctly, in full and/or in the right place, they become frustrated and irritated and your database becomes polluted or you lose their custom. Companies seem to perceive it as acceptable to spend large amounts of money cleaning and validating data collected via an Internet form, but don't appreciate that a smaller amount of money spent on optimising input forms for their international visitors will save them far larger sums of money further down the line.

Electronic media such as the Internet are able to provide companies with unparalleled abilities to adjust their pages and input forms to the visitor concerned, but companies are not taking advantage of this for handling personal name and address gathering. A recent survey names its top language-compliant international web sites, but a quick check of their personal name and address data collection pages shows no attempt at internationalising this part of the process. If they are the top of the pile, there's a long way to go.

Information resources and software components do exist which can resolve this problem. Requesting initially, for example, the country of residence and the language in which they would like to see the form allows a web page input form to be built which perfectly mirrors the requirements of the customer, in terms of fields presented, field order and field labels. Other components allow real-time validation of an entered address against postal files so that all data entered can be ensured accurate, consistent and valid. Using these resources can produce better data for the website owner and a happier customer, which in turn can translate to commercial success.

Alas, sales of these components are known to be minimal, and this is clearly to be seen by looking at almost any web data collection form. Clearly, we still have a long way to go before we are able to pull companies out of this particular rut.

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