Global Shift:
Total Data Quality For e-Retailing

Graham Rhind

Sponsored by Capscan

A Trading Division of GB Group

TABLE OF CONTENTS   Page

INTRODUCTION

When it comes to the reasons for poor data quality, the buck is often passed to the customer. Capscan's own research in 2008, 2010 and 2012[1] shows a substantial and stable feeling that a major cause of data quality is data entry errors by customers.

Customers are accused of laziness, deliberate attempts to mislead, skipping information, providing the wrong information in the wrong fields in the wrong format, and of being in too much of a hurry, often in an attempt to protect their personal data.

Customers are the owners of their own data and know how it should be written and formatted. Yet many systems require formats and use processes that the customer is ignorant of, and usually uninterested in. Customers will take some care to provide information that they consider relevant to the transaction, such as e-mail address or telephone number, or postal address if an item is to be sent, but see no reason to be the guardians of an organisation's data quality when entering these or other, less relevant, pieces of data.

THE NEED FOR BETTER SYSTEMS

Yet rarely are these errors down to the customers. They are often down to an organisation's poor design, poor processes, poor training and poor planning. If an organisation wants good data (which is the customers to give and not the organisations by right), systems that work need to be designed.

Trends in commerce and in our behaviour are shifting fast. We are finding new ways to shop, using different channels and from different devices. Retailers need to be agile to keep ahead of this movement, all the while embracing the demand for their products from a wider range of people in more countries than ever before. The need for a holistic and enterprise-wide Global Data Quality Management (GDQM) philosophy, with systems in place to match, has never been greater.

Yet most retailers are failing their customers and are suffering the consequences. The lack of an international data quality approach is apparent in many of the online shops that customers need to wrestle with.

> " The need for a holistic and enterprise-wide Global Data Quality Management (GDQM) philosophy, with systems in place to match, has never been greater. "

[1] http://www.capscan.com/white_papers.aspx

Postcode

1018 VV

Please enter your street number, building, mailpoint/floor etc. (if required) and begin entering your street. Pick your address from the list of any matches found that appears (your address will be automatically completed if there is an exact match):
Building/Floor etc.

Building Number

Street

Locali

Town

Coun

> Message from webpage
>
> ⚠ Please supply a complete postcode
>
> OK

Please amend your international dialling code if your phone number is located in a different country.

Telephone

+31 (0)

Next

Some companies which are attempting to retail internationally still use data collection systems which block foreign information.

> " Access to retail sites by customers from mobile devices is also booming – research from Reevoo[2] suggests an increase of 137% between 2011 and 2012. "

[2] Reevoo, "Mobilising Social Commerce", 2012.

E-COMMERCE IS BOOMING

E-commerce is one of the fastest growing markets in the world. In the UK online sales are already estimated to be over 13% of all retail sales, and this figure is growing fast in every country. Access to retail sites by customers from mobile devices is also booming – research from Reevoo[2] suggests an increase of 137% between 2011 and 2012. As e-commerce increases in importance, the number of shoppers from other countries, and their potential spending power, is increasing in gross value and are becoming significant sources of income for the retailers concerned. These retailers need to provide a flawless and frictionless shopping experience for every web shop visitor, regardless of their geographic location.

The boom in e-commerce removes the opportunity of personal interaction with the customer, especially in global automated systems online 24 hours per day, 7 days per week. The opportunities to guide

the customer through a process or have a dialogue with them are reduced, yet the need for high quality data collection remains the same. With processes moving from the personal to the digital, more thought and attention to detail is required to achieve sales and prevent the high level of shopping cart abandonment currently experienced by the industry.

MORE CHANNELS

The data collected is also being used in new and innovative ways. These might include location-based systems, fraud detection and prevention, and security, improved channel support, the move to new markets and channels, and identity verification.

Just as channels bringing data to an organisation are increasing in diversity and scope, so the uses to which data can be put is expanding. Rapid address input and validation is an essential part of any serious e-commerce operation, but it is no longer sufficient. New and improved channels, legal and financial requirements and the need to combat fraud all dictate that other data, such as name, telephone numbers (fixed-line and mobile) and e-mail addresses, also need validation. Validation using extensive reference tables not only ensures that data enters any system accurately and correctly, so that it can be trusted and used for customer contact, sales, marketing and business intelligence further downstream in the business process; it ensures that you are doing business with real people whose purpose is to purchase from you and not with bots or criminals with more nefarious motives.

VALIDATION AND VERIFICATION

As part of a Total Data Quality Management (TDQM) regime, regular batch cleansing of back-end data is advisable because of our fast changing world and your mobile and dynamic customers, who move house, change or add telephone numbers and e-mail addresses, get married or divorced and so on. Batch-cleansing is not sufficient, though. Data should enter any system clean – it should be validated and verified at source. It is the optimal time to ensure the best data quality – it is never possible to achieve the same level of cleansing during back-end batch routines.

> **Poor data quality affects every part of a business, and organisations suffer when mission critical decisions are based on low quality data.**

Furthermore, in larger systems data is injected and exchanged within and between data files, so that back-end cleansing becomes more complex. Often, when data is cleansed in one file it is overwritten by incorrect data from other systems, making the issue far more challenging. Data of every type, from addresses though telephone numbers to e-mail addresses and beyond, can be validated at source far more cheaply, effectively and in a more controlled way than validation at any later point. Poor data quality affects every part of a business, and organisations suffer when mission critical decisions are based on low quality data.

> Canadian Subscribers: When entering your postal code, please omit spaces.
> International Subscribers: Please put your postal code at the end of your second address line.

Placing complex and unnecessary burdens on the customer to enter data is particular ways will have them leaving your site in droves.

THE SHIFTING DATA PARADIGM

The amount and range of data being collected, how it is being collected, and where from, is increasing all the time. Data is now gathered through internet interfaces on personal computers but also via mobile systems (m-commerce) – the way we communicate with each other and with business is changing fast. These mobile systems also differ widely, from tablets to smart phones. Users of mobile devices, especially those with small screens and without keyboards, benefit particularly from well-designed rapid entry data collection systems. New data being gathered includes various electronic addresses, such as e-mail and social media URLs, and much more often now that data can originate from any point on the planet. Organisations have been slow to adjust their systems to keep up with these developments. Whereas national verification systems of some basic data such as postal addresses are not uncommon, the verification of other data such as telephone numbers and e-mail addresses and of international data is far less common; and where such verification exists, it often exists in only a small number of the data gateways an organisation may use. A company might, for example, use a rapid address collection and verification system in a call centre, but use badly designed and unsuitable web forms in shopping cart software in their online e-commerce sites, ensuring not only that data collected is of a poor quality, but that large numbers of customers abandon their carts before completion of a purchase. All sources suggest very high abandonment rates. Customer

> **What works here won't work there. Companies like to think globally but they need to act locally.**

Experience Management (2010) suggested a rate of 63%[3] and Forrester a staggering 71% in 2011[4]. Whilst some of this is due to poor website design, the requirement to register or hidden charges by the retailer, much is also due to difficulties in navigation through the checkout process, particularly those parts relating to data collection. These tests have been carried out almost exclusively on national shoppers. For international retail, information capture becomes a much greater part of the problem.
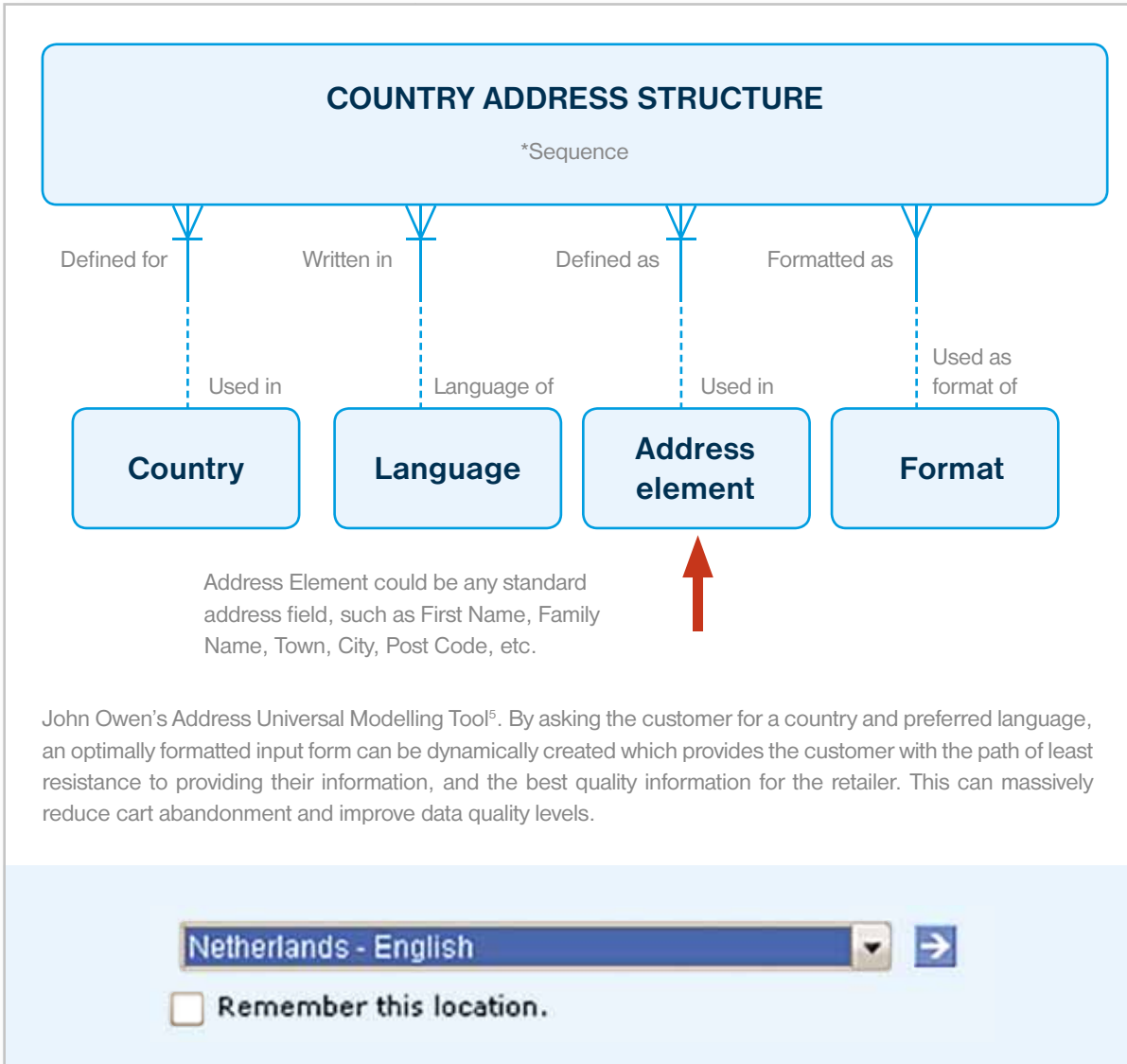
THE GLOBAL MARKET

What works here won't work there. Companies like to think globally but they need to act locally. Whilst some companies spend time and money to localise their websites to local cultures and markets, very few extend this line of reasoning to data collection or checkout systems. We live in a world of over 240 countries and territories, containing inhabitants speaking one or more of over 6000 languages, written in one or more of over 60 scripts, each written in different directions across the page. There are over 130 address formats, over 40 personal name formats, numerous date, time and number formats and a mobile global population. Very few organisations are aware of this diversity, and fewer still take it into account in their systems. Whereas asking the customer for their country and language of choice can allow input forms to be dynamically created which mirror the format of that person's expected input, most forms and shopping basket systems are fixed. They request input using inflexible forms designed for other markets, using terms that may not be universally understood, often in a language foreign to the customer, requesting information that the customer may be unable to provide, and/or leave no place for essential information to be added.

Country	United Kingdom ▼
House No	17
Postcode *	ox10 6ry

Rapid addressing systems, designed for each country's unique addressing system and requesting the least amount of information from the customer to provide a correct and verified address, greatly improve the usability of check out systems, increase conversion rates, reduce cart abandonment, provide the least friction for a customer to complete the form, and ensure best quality data enters an organisation's data stream.

[3] http://www.alienationdigital.co.uk/basket-abandonment-online.html

[4] http://www.retailtechnology.co.uk/market-analysis/reducing-basket-abandonment

COUNTRY ADDRESS STRUCTURE

*Sequence

Defined for	Written in	Defined as	Formatted as
Used in	Language of	Used in	Used as format of
Country	**Language**	**Address element**	**Format**

Address Element could be any standard address field, such as First Name, Family Name, Town, City, Post Code, etc.

John Owen's Address Universal Modelling Tool[5]. By asking the customer for a country and preferred language, an optimally formatted input form can be dynamically created which provides the customer with the path of least resistance to providing their information, and the best quality information for the retailer. This can massively reduce cart abandonment and improve data quality levels.

Netherlands - English

☐ Remember this location.

[5] http://integrated-modeling-method.com/address-universal-modelling-tool/

POOR FORM

Many forms, including those used for shopping carts in e-commerce, have been designed to collect data optimally from a single country, often the United States, and make few, if any, concessions for shoppers from other countries. When shoppers are faced with forms like this, many have difficulties in providing the required information, and many give up before the process finishes. For those using devices where input is less easy, such as mobile devices, abandonment is greater. Forms will, for example, expect a state to be added to an address, even in countries which do not have them, and any drop down for this information often contains only states, provinces and territories for the United States and Canada. Most systems require a postal code, though more than 50 countries and territories do not have them. This has become such an issue for some smaller territories that a single country-wide postal code has been introduced simply to allow citizens to take advantage of online shopping.

First Name:	Last Name:
Email Address:	Address:
City:	State/Province:
	Choose one...
Zip/Postal Code:	Country:
	Netherlands

Demanding data which a customer may not be able to provide, such as a state or a postal code in countries, which do not have them, either cause customers to abandon the process or to enter false information in order to make a purchase, creating numerous data quality issues for the organisation downstream.

Some forms expect a minimum text length in each field, even though there are people on the planet with single-letter names living in single-letter places. Telephone number fields often expect no more than 10 digits and the numbers to be formatted in the North American manner, even though some countries use longer numbers. Personal forms of address are often limited in scope to those showing gender,

A Trading Division of GB Group

and those only being relevant for single culture. The list of cultural pitfalls that an organisation without an effective global data quality management system in place trips over goes on. Forms which do not take global differences into account do nothing to aid a customer in purchasing a product. They leave customers struggling to complete them to the satisfaction of the underlying coding, and immensely increase cart abandonment or refused transactions. Those customers who do not abandon the process often have to add unverified data in such a way that downstream systems using that data are clogged and extra cleansing and processing becomes required.

CONCLUSION

The data world is developing rapidly. We are able to interact with more people, in more ways, through more devices, in more places, in greater quantity and through more channels than ever before. The world is within reach, but it remains diverse. Automated 24/7 data collection and e-commerce systems need to be agile and strong enough to take over from human interaction in the sales and check-out process. Data needs to be cleansed and verified, both in the back-end system and at the point of capture. This verification, however, cannot be a one-size fits all system designed for only one cultural group if retailers are reaching out to customers throughout the world. Dynamic data collection forms needs designing which take into account the world's diversity, supported by verification and cleansing software supporting data for all countries and territories, not just one. Without a holistic global data quality management system and philosophy in place, organisations will be unable to optimally take advantage of the worldwide market, which is there for the taking.

About the author

Graham Rhind is an acknowledged expert in the field of data quality. 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 four books on the topic of international data management.

Web: www.grcdi.nl
Email: graham@grcdi.nl
Twitter: www.twitter.com/grahamrhind

About Capscan

Capscan, a trading division of GB Group, was originally established in 1969 and is a leading supplier of customer registration and international data quality management (IDQM) solutions. The company is headquartered in London with more than 1800 customers worldwide across a wide range of commercial and public service sectors. In the UK alone, there are currently more than 140 different government departments, agencies and local authorities using Capscan products or services. In the private sector, Capscan's customers include a wide range of leading blue chip companies as well as small to medium sized businesses.

A Trading Division of GB Group

Capscan

Capscan
Moorgate House
Dysart Street
London
EC2A 2GU

Tel: +44 (0)20 7428 1255
Fax: +44 (0)20 7267 2745
Email: enquiries@capscan.com
Web: www.capscan.com