Capscan

The Hidden Returns On Your Data Quality Investment



Graham Rhind

Sponsored by Capscan Ltd



TABLE OF CONTENTS

	Page
Introduction	2
Data quality versus data quality projects	3
The outfield effects of data quality improvement	3
Minor improvements, major gains	4
The downstream effects of improved data quality	4
How do data quality improvements affect ROI?	6
Collecting data correctly	7
Case studies	7
XYZ Limited and their product release mailing	8-9
CDE PLC	10-11
Conclusion	12

© Capscan Ltd

INTRODUCTION

Improvements in data quality will always benefit an organisation in one way or another, in financial and other respects. Unfortunately, a culture where data quality is treated as the norm and as an essential part of all business processes has pervaded very few organisations. Organisations often treat data quality in a different way to other parts of their business. An airline, for example, would rarely request or expect a return on the investment it puts into the maintenance of its craft, as it is patently obvious how their business would suffer if poorly maintained aircraft started falling from the sky. Poor data quality can be equally pernicious, but it doesn't have the same immediate and obvious impact, and it tends to be given too little attention in most organisations. Gaining the support of senior management for spending on data quality is an essential but sometimes difficult task. Almost all organisations use some form of capital budgeting for their investment decisions, but there is a disconnect between those people who see and work with data, and have to grapple with the problems poor data quality brings with it on a daily basis; and those who make the business decisions and need to understand the impact bad data quality has on their financial bottom line.

Return on investment (ROI) is dependent upon how data is and will be used. ROI on improving the quality of data which is not used and, (importantly) will never be used, is zero. Traditionally, ROI measurements look at the immediate effects of introducing quality improvements, at the point of introduction and within a department's or organisation's boundaries.



© Capscan Ltd



DATA QUALITY VERSUS DATA QUALITY PROJECTS

There is a need to distinguish, though, between data quality improvements and data quality improvement projects. Some projects are ill-conceived, overly complex and fail to tackle the causes of poor data quality. In these cases, ROI from their implementation may be negligible or negative. It is the key to understanding the relationship between your data and your business, and also the relationship between the data and factors external to your business, as data quality can, for example, affect your ability to attract and hold customers.

Traditionally, data quality improvements can provide ROI and impact an organisation by increasing revenue (through business growth), increasing revenue by decreasing costs, and reducing risk by allowing compliance to laws and regulations.

THE OUTFIELD EFFECTS OF DATA QUALITY IMPROVEMENT

Measures of ROI must look much further than the point at which the service is being implemented. A measure, for example, that provides only an immediate financial benefit from reducing the number of keystrokes made per operator per customer in a call centre and presents that as ROI, is missing the huge scope and effect that data quality has, both within and outside an organisation.

As an example, an internal study of Dutch data files containing over 1 million records, found that the place name 's Gravenhage (The Hague) had been written in over 50 different ways, and that this inconsistency affected over 50% of all records. Introducing address validation software at the data entry point (for example, in the call centre) shows ROI at that immediate point – faster location of customer records with the commensurate reduction in staff costs, reduction in the number of duplicates created, and so on. Few ROI studies, though, look further. Any business intelligence using the inconsistent data is likely to underestimate the number and value of customers they have in that city by over 50%, and decisions made on that basis will be flawed and very costly for any organisation.

© Capscan Ltd

> "Using Matchcode has significantly improved the value and usability of the Butlins, Haven and Warner Breaks databases and we know that every new address that is entered into the Holiday Sales Systems is of the same quality and common format. We expect to roll out names data across all our brands, as it will further support accurate addressing and help identify real duplicates in each of the databases."

IT Manager at Bourne Leisure,
Jeremy Stancombe



MINOR IMPROVEMENTS, MAJOR GAINS

Enormous improvements in data quality can often be achieved with the smallest financial outlay. The secret is to identify the main causes of the problem and to tackle them directly and in the most efficient (and usually most cost-effective way). Many companies, for example, install and run expensive data validation and cleansing systems to improve the quality and value of data already held, but it is cheaper to prevent data quality issues at source, though showing the ROI of this is more complex and often results in those measures not being taken. Altering a web data-entry form so that your customers are shown only those fields which are relevant and actual for their country, in the correct order, without obligatory fields which they cannot complete, and in the language that they speak, is a very cheap and highly effective way of improving data quality, and it can be implemented quickly and with the minimum of fuss. Yet it is rarely attempted – organisations are structured in ways that militate against such common sense approaches.

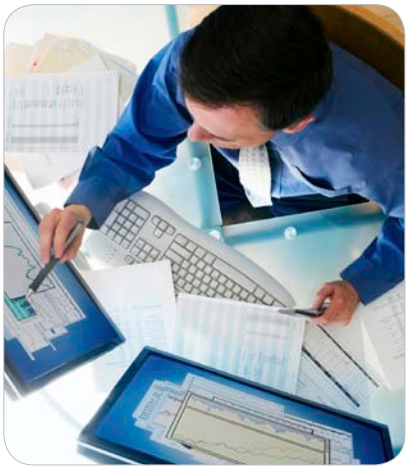
It is generally easier to show ROI and to get executive buy-in for upstream data quality improvements when a down-stream initiative is already in place. The downstream system will already be showing ROI figures, and upstream improvements to data quality will often eliminate many of the downstream costs.

THE DOWNSTREAM EFFECTS OF IMPROVED DATA QUALITY

Any data quality improvement will have effects on every user and the use of that data downstream from the improvement. Address validation will reduce costs for the call centre. It will also reduce duplication rates, improve marketing effectiveness, improve sales drives, improve business intelligence initiatives, allow business decisions to be made on better data, and so on; all of which will have immediate or future financial returns. Address validation increases revenue assurance and shortens the invoice to cash cycle. It reduces waste stemming from returned mail, failed deliveries and the extra staff required to deal with returns and the customer complaints stemming from them. It also improves the accuracy of de-duplication and house-holding procedures, with the resultant reduction in marketing and delivery costs.

Improved de-duplication reduces invoicing errors and the costs involved in communicating multiple times with the same customer. In this way, churn from dissatisfied and irritated customers is reduced.

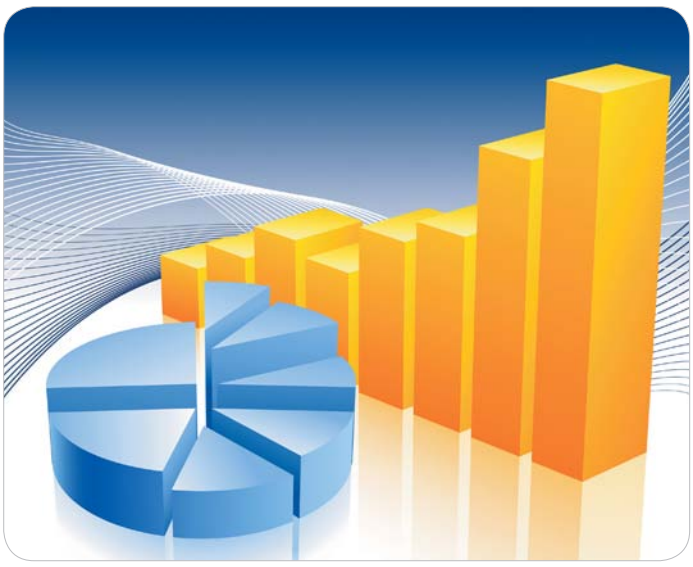
© Capscan Ltd



Equally, ROI should be assessed outside the organisation, and in terms of effect on customer acquisition and churn. How many fewer prospects are now replacing the receiver before speaking to a call centre operative because waiting times are reduced because of faster address entry and validation? How many more people achieve contact with your organisation through the web form which is formatted and validated for the country that that person is in? To what extent have you reduced loss of custom to competitors by not presenting prospects with a form demanding information that they cannot give, such as a state or ZIP code outside the United States? And how many people will be turned towards, or away from, your organisation on the basis of the stories your prospects and customers will be telling about their experiences with you?

All of these have an impact on the ROI of any data quality improvements. Cast the net far and wide and you will get a true impression of how improving data quality will positively affect your organisation.

The degree of accuracy improvement is an important aspect of ROI, especially when international data is involved. Often systems are set up for national data, and operators have the cultural and linguistic background to understand this data, so data errors tend to have relatively less impact than when international data is being gathered. Systems which cannot handle foreign languages or name and address formats, and operators without the background or linguistic abilities to recognise potential errors, will create major data quality issues which have major impacts on downstream processes. Thus, the ROI of improving data from one place may be greater than an improvement to data from another place.



© Capscan Ltd

Some less obvious effects of data quality improvement

- A complete, accurate and relevant view of each customer is available to all departments within an organisation. With this improved business intelligence, your organisation makes decisions based on accurate and reliable data. The business is healthier and more robust.
- Marketing and sales waste less time looking for data, trust their data sources more and can be more targeted and more effective
- Information Technology departments spend less time fighting data quality fires.
- Improved revenue assurance and a shortened invoice to cash cycle.
- Reduced costs from wastage – mail returns, customer complaints etc.
- Increased customer satisfaction and reduced customer churn from dissatisfaction and a negative brand image – gain and maintain customers more effectively.
- Improved fraud control.
- Better and easier compliance with the regulatory environment.

HOW DO DATA QUALITY IMPROVEMENTS AFFECT ROI?

Because ROI is dependent upon the data and how it is being used, it is impossible to provide an example on measuring ROI which is relevant for all organisations. However, we can look at examples of how introducing data quality improvement in some sample parts of the business can affect different parts of that business, from which ROI information can be gathered.

Introducing address validation at the point of collection has an immediate effect upon the quality of the data gathered. Data is gathered faster - operators need fewer keystrokes to find the information that they are looking for. This increases the efficiency of the operators – more customers can be helped and fewer call centre staff may be required. Prospects and customers spend less time waiting for calls to be answered, leading to increased customer satisfaction. Fewer people replace the receiver through frustration before their call is answered. Increased satisfaction increases an organisation's worth in the eyes of their customers and, through them, all of those friends, colleagues and members of their various networks that they choose to tell about their experiences with that particular organisation.

Address validation increases greatly the quality of the data being gathered. The data quality is affected less by the restricted linguistic and cultural background of the operators. They make fewer errors in trying to enter data with which they are unfamiliar. They create fewer duplicate records and can find customer data more quickly and easily.

The increase in data quality affects every worker, department and process using that data downstream inside and outside the organisation. Knowledge workers have to spend less time trying to locate records, and because there are fewer duplicates, they find the information that they need in the records that they find, rather than having to track down possible duplicates containing different information. Business analysts are able to produce better and more accurate summaries of the data that the organisation has. Through more trustworthy business intelligence, executive decisions can be made on a more solid basis, leading to a healthier and more robust business. The Information Technology department has more time and resources to spend on their core work instead of fighting the fires that rage when data quality is poor.

© Capscan Ltd

> "Capscan not only enables access to our master data in a way we've never been able to before but also benefits the organisation as a whole by reducing service delivery errors and helping to ensure those services are of a high quality and speedily executed."

CRM Development & Support Officer
at Trafford Council,
David Aspinall

COLLECTING DATA CORRECTLY

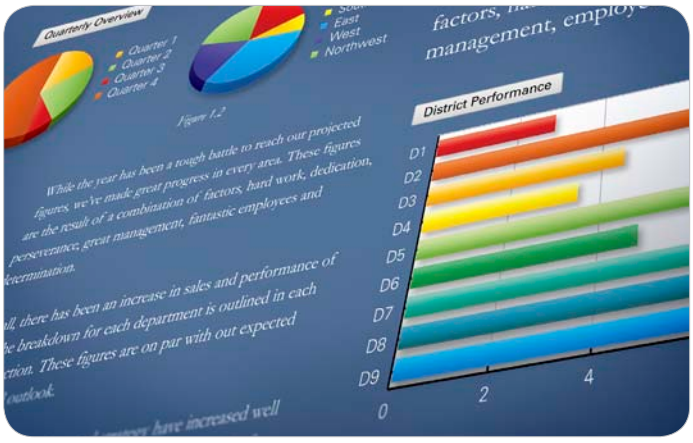
When data is already clean the cost of cleansing data downstream is significantly reduced (though not eliminated – some data, such as address data, decays over time, so the data needs to be maintained). Cleansing downstream is always less effective than cleansing at source – you cannot always identify and remedy quality issues, and you do not have the opportunity to hold a dialogue with your customer that you have at the data collection point. Sirius Decisions coined the 1-10-100 rule for this: "It takes $1 to verify a record as it is entered, $10 to cleanse and de-dupe it, and $100 if nothing is done, as the ramifications of the mistakes are felt over and over again"[1].

If data is to be used to communicate with the customer, ROI is clearly evident in these processes too. Fewer customers are called or mailed multiple times, leading to an improved customer perception of your organisation. Less needs to be spent on (tele)communications, handling and production costs. Returns from mailings or product deliveries are reduced, so fewer resources are required to manage them; and more customers get their goods and get them on time. Invoices can be sent earlier and fewer refunds need to be made, so the financial cycle time is reduced.

CASE STUDIES

As mentioned previously, ROI will always depend on how the business and data interact – ROI will be different for each organisation. However, some examples can demonstrate how immediate ROI can be calculated for some typical scenarios.

[1]"The Impact of Bad Data on Demand Creation", Sirius Decisions, January 2009



XYZ Limited and their product release mailing

XYZ Limited has a data file with 100,000 customers and prospects, which they'd like to mail with news of an updated product release. The file is uncleansed and was collected unvalidated, and it has a 30% duplication rate (i.e. 40000 customers appear only once in the file, 30000 appear twice). The cost for each mail piece (design, printing, handling, postage) is GBP 1. The cost of sending 100,000 mail pieces (to 70,000 individuals) is GBP 100,000. Response rate amongst the non-duplicates is 2%; amongst the duplicates only 1%, as they are irritated at getting multiple duplicated mailings. The value of each response is GBP 100, giving a return of GBP 110,000, a very small profit on the exercise, but resulting in a loss of brand value for the company due to its sloppy use of their customers' (duplicated) data.

XYZ Limited invest in de-duplication software, which locates 15,000 duplicates (50% of all duplicates), leaving a file with 85,000 records so that the total cost of sending the mailshot is reduced to GBP 85,000. Response rates remain the same, but as the number of duplicates is lower, the number of responses rises, giving a response value of GBP 125,000 (for a domestic mailshot).

If that data is international, XYZ Limited would have to contend with higher duplication and non-deliverable rates, and higher costs per communication made, as data from unfamiliar language and cultural areas is entered with a greater range of number of errors without validation. Data quality improvement in this case would bring greater rewards.

Investing also in international data validation and cleansing software, so that data is validated at source and standardised for other processes increases the effectiveness of de-duplication software, reducing the duplication rate of the file further thus reducing the size of the mailshot to 75,000, costing GBP 150,000. Response increases again to provide a return value of GBP 135,000 (for an international mailshot, with a higher cost for each mailpiece sent).



© Capscan Ltd

XYZ Limited, domestic mailshot

WITHOUT DATA QUALITY INITIATIVES	WITH DE-DUPLICATION	
100,000	85,000	Number of mail pieces sent
GBP 1	GBP 1	Cost per mail piece
GBP 100,000	GBP 85,000	Total cost (cost per mail piece x number of mail pieces)
	GBP 15,000	Saving
1,100*	1,250**	Positive response (actual)
GBP 110,000	GBP 125,000	Income from response (number of responses x [GBP 100] average value of response)
	GBP 15,000	Improved income
	GBP 30,000	Total amount saved through de-duplication initiatives for this project alone

* 2% response rate from the 40,000 customers who receive the mailing once (800 responses) and 1% response rate from the 30,000 customers who receive the mailing twice (300 responses)
** 2% response rate from the 55,000 customers who receive the mailing once (1,100 responses) and 1% response rate from the 15000 customers who receive the mailing twice (150 responses)

XYZ Limited, international mailshot

WITHOUT DATA QUALITY INITIATIVES	WITH DE-DUPLICATION, VALIDATION AND LEANSING	
100,000	75,000	Number of mail pieces sent
GBP 1.5	GBP 1.5	Cost per mail piece
GBP 150,000	GBP 112,500	Total cost (cost per mail piece x number of mail pieces)
	GBP 37,500	Saving
1,100*	1,350**	Positive response (actual)
GBP 110,000	GBP 135,000	Income from response (number of responses x [GBP 100] average value of response)
	GBP 25,000	Improved income
	GBP 62,500	Amount saved through data quality initiatives for this project alone

* 2% response rate from the 40,000 customers who receive the mailing once (800 responses) and 1% response rate from the 30,000 customers who receive the mailing twice (300 responses)
** 2% response rate from the 65000 customers who receive the mailing once (1300 responses) and 1% response rate from the 5000 customers who receive the mailing twice (50 responses)

© Capscan Ltd



> **"Matchcode International has minimised the number of keystrokes the data entry team needs to make to identify a full, valid and accurate address from fifty to eight to ten keystrokes. We can also be confident that in such a data-sensitive market, the addresses at the point of capture are to the required standards of the country postal services and crucially, our client."**

Group Project Manager, Andy Boniface,
Graphic Data UK Ltd

This is a classic ROI calculation, but it fails to take into account the many subsidiary positive aspects of improving data quality within an organisation – customers have increased trust in you and will maintain or increase their custom with you; staff in downstream processes spend less time searching and sifting through data to find information they need; business decisions are made on data whose quality provides a firmer foundation upon which to build an organisation's future.

CDE PLC

CDE PLC provides support and takes orders via a call centre (employing 100 operators). Operators enter customer data in its raw form, as they understand it from the customer on the line, and, if the system does not throw up a duplicate, will create a new record for that person. This results in major duplication within the system. There are multiple views of the customer, products sent out are often returned as undeliverable and require handling, invoices get lost, delayed or are sent out multiple times. CDE are spending an average of GBP 4.5 per customer just to collect and correct the data as telephone calls need to be longer than necessary, complaint call numbers from customers caused by duplicates in the system are high and intensive interaction with the customer is required.

CDE introduce a data entry system into its call centre using rapid addressing technology – operators can quickly find a customer record using only partial address information, reducing keystrokes and time that the customer needs to be on the line, and increasing productivity within the call centre. As data is validated and enters the system in a standardised form, duplication goes down dramatically. Fewer packages and invoices are returned, so less handling is required, and invoices arrive on time and are paid quicker, reducing the need to employ staff to chase up unpaid invoices. The amount of money spend per customer spent in the call centre is reduced from GBP 4.5 to GBP 0.9.

© Capscan Ltd

CDE PLC, call centre costings

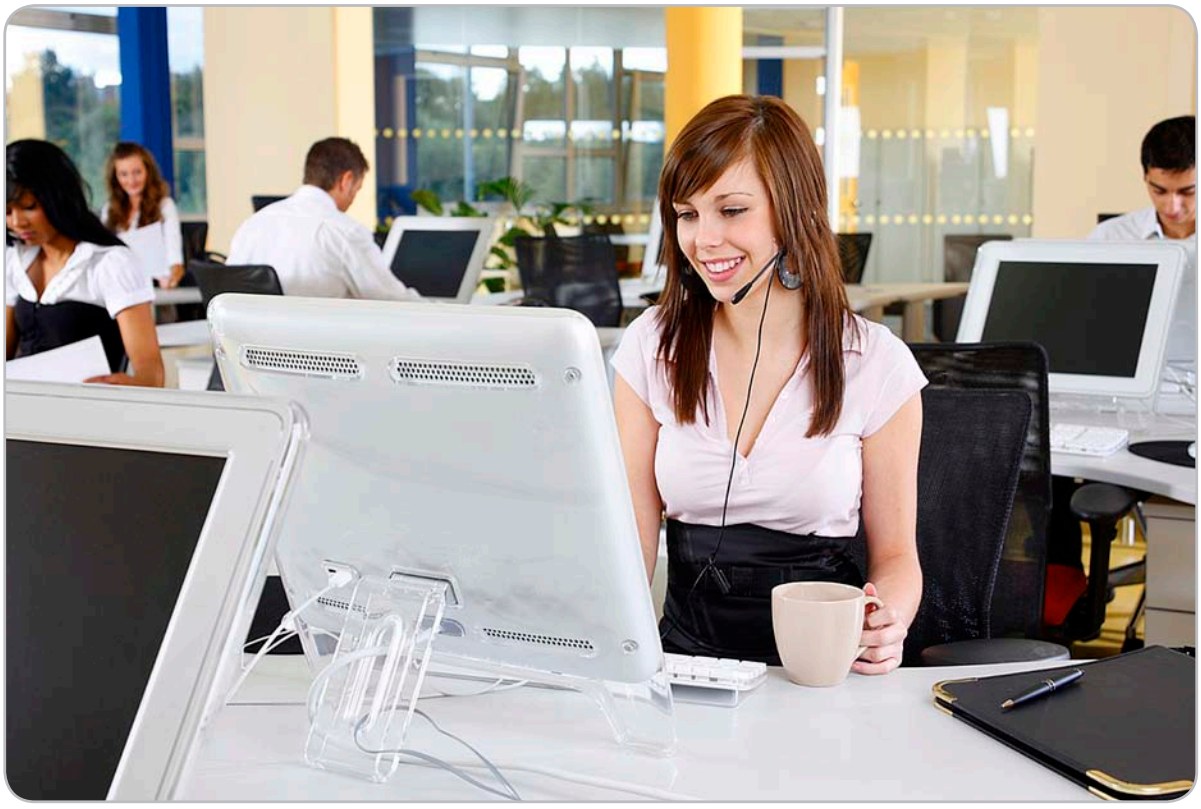
WITHOUT DATA QUALITY INITIATIVES	WITH DATA QUALITY INITIATIVES	
GBP 0.005	GBP 0.005	Cost per second of call centre operative *
900	180	Seconds required to handle call
GBP 4.5	GBP 0.9	Total cost to handle call (cost per second x seconds required)
3.33	16.66	Calls handled per operative per hour**
100	20	The number of the original operatives required to man the call centre after improvements***
0	GBP 2,764,800	Example savings per annum directly from call centre staffing reduction ****
0	Etc.	Additional infrastructure savings (electricity, telephone costs etc.)

* This figure is based on an average UK call centre operative cost of GBP 18 per hour, and includes wages, training, insurance, employee infrastructure etc.

** Assuming an average effective working time of 50 minutes (3,000 seconds) in the hour.

*** This is assuming a stable number of incoming calls. Improved data quality actually decreases the number of calls because businesses are not giving their customers cause to complain or with the need to run after the company to correct errors, so this number is likely to be even lower in reality than in this example.

**** Assuming an annual UK call centre operative cost (as opposed to wage) of GBP 34,560 (40 hours per week and 48 effective working weeks per annum). The staff number has decreased from 100 to 20, a reduction of 80. Cost savings = GBP 34,560 x 80.



© Capscan Ltd

CONCLUSION

Many of these returns on data quality improvement are hidden and are not always seen and taken account of in traditional data quality ROI reporting; yet even without these prizes, ROI can be seen to be high even within the immediate environment of the improvement taking place. It would be desirable for organisations to build data quality into the very fabric of their structure and processes, but whilst that is the case, it may still be necessary to demonstrate ROI to those decision makers who do not have immediate access to the raw data and are unable to see the advantages of improving it.

No improvement of quality of any data which is, or will be, used, is ever wasted in an organisation. Improvements reduce the load and improve the effectiveness of every person and department using the data downstream. They spend less time finding data, they are able to do their jobs more effectively and, where they have customer-facing jobs, they can serve the needs of the customer more quickly and more efficiently. Less time needs to be spent on fire fighting, and organisations can make operational and policy decisions on the basis of correct and trusted data. Organisations which work with incorrect and poor quality data will, ultimately, fail.



© Capscan Ltd



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 Ltd

Capscan is a leading supplier of international address management solutions and data quality services. The company's flagship product, Matchcode, is a fully functional address management system, available as a stand-alone programme for data capture, a web-based tool for online data capture and as a tool for batch cleansing of commercial databases. Capscan now supports addressing functionality for 240 countries or territories worldwide. In addition, Matchcode UK address data can be integrated with Ordnance Survey data sets to allow mapping and logistics rationalisation. Capscan is also expert in the integration of lifestyle and business data sets into address databases to support and rationalise canvassing or marketing activities throughout Europe.

Capscan Ltd
Grand Union House
20 Kentish Town Road
London NW1 9BB

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













