The Importance of Information Quality Management for your Business

Interview with Danette McGilvray for ZipCode, a Brazilian company

August 9, 2011

1. What is data quality and why is it important?

First, a quick definition: data or information quality is the degree to which information and data can be a trusted source for any and/or all required uses. It is having the right set of correct information, at the right time, in the right place, for the right people to use to make decisions, run the business, serve customers, and achieve company goals. (Is there a difference between data and information? I see data as facts of interest to your business. When those facts are put in context they become information. Other than that I often use the words interchangeably.)

What that means to you is whenever you pull up a customer name and address to take an order, process a shipment, send marketing materials, solve a customer complaint, or create an invoice you can depend on that information being correct. You can count on it to efficiently complete the business transaction.

It also means you can trust your business intelligence and analytics. Whenever you bring together information about your top customers across the enterprise, you can confidently adjust your strategy based on those reports.

Why do these things matter? The more *effectively* we take care of our customers the less it costs the company. This also increases the likelihood our customers will do business with us again, which then increases revenue.

2. Which are the main symptoms that lead the company to identify the need to invest in the quality of information?

If you are sending marketing materials via email or postal mail and you are getting many returns - you have a data quality problem.

If your staff is spending more time arguing about whose report is correct (when the numbers should agree, but do not) - you have an information quality problem.

If your people are spending a lot of time on correction activities, rework, reprocessing orders, handling returns or dealing with customer complaints – you may have a data quality problem. Many of these activities do not appear to be associated with information quality, when in fact they are. Since processes and functions are distributed across an organization and many people, the cost and scope of data quality problems are often not visible.

If you are implementing a new system, such as a Customer Relationship Management (CRM) application, and you are integrating and moving data from various source systems - you will have an information quality problem.

I can say this with confidence because every project I have personally been involved with or heard about that integrates and moves data has had data quality problems that had to be addressed for the application to work correctly when put into production. Data that may have fulfilled the needs of one particular functional area are now combined with data from other functional areas – often with very poor results. If this aspect of your project is not taken care of, the business *will* experience problems.

Why do we care about that? We are not just talking about a piece of technology working. We are talking about information being available that is essential to your people to do their jobs – allowing them to take and ship that order, send the marketing material, create the invoice, and process the payment.

3. How can we justify to the board/Directors the need of an investment in data quality?

Asking to invest in data quality in a *general* sense will not find you many supporters. Data and information quality must be tied *specifically* into whatever is important to the company. But surprisingly it is not always clear which data is associated with what the business really cares about. To get started, use my quick Connect-the-Dots technique to tie the data to your business needs:

1. What are the most important business needs in your company? These can be goals, strategies, issues, or opportunities.

2. What are the people, processes, and/or technology associated with those business needs?

3. What is the associated information that supports those processes, people and organizations or is used in the technology?

4. What is the data that makes up the information?

Once these questions are answered, it connects the data with the business needs. This technique is easy to do and is the starting point for showing there is value to having high-quality data and information.

Follow this by using other relevant business impact techniques, which are qualitative and quantitative measures for determing the effects of data quality on your organization. These techniques range from less complex (such as collecting examples or stories about the impact of poor data quality) to more time consuming techniques (such as quantifying the costs and revenue impact of poor quality data).   Having real results of business impact combined with good communication skills will go a long way towards getting support for your data quality efforts.

If you are interested in learning about these techniques for assessing business impact see *Executing Data Quality Projects: Ten Steps to Quality Data and Trusted Information™* by Danette McGilvray (Morgan Kaufmann, 2008). I mention the book because often people think there is no way to show the value of information quality. We actually can do that, but we don't have time to discuss all the techniques today.

4. Where should we start, taking into consideration the stages of a company (start, growth, maturity, and decline)?

No matter what stage your company is in, start with the most important business needs and the data and information associated with those needs.

5. **Which kind of tools and suppliers are required to develop a project of data quality?**

There are various tools on the market which can help with your data quality. *Discovery tools* look at multiple sets of data, identify data available and find relationships and tranformations based on the business rules hidden in the data. *Data profiling* tools use analytical techniques to understand the validity, structure, content, and other basic characteristics of data. Start with these tools to understand the magnitude and location of your data quality problems.

Data cleansing tools (also known as data quality tools) are used to identify duplicates, link related records, standardize, and update records. Other tools may not be labeled as data quality tools per se, but if used properly will enhance the quality of your data. These tools include name and address verification, data enriching or augmentation, data modeling, metadata management, and ETL (Extract-Transform-Load) tools. Application development tools often have modules that include some of the functionality mentioned.

As important as tools are, it is equally critical to have a good process for taking care of data quality (such as the Ten Steps™ approach) and people working together with knowledge of your business, data, and technology.

6. **Would you like to leave any comments that you consider important for companies?**

Though I have used customer information as examples, what I have shared here works with all types of data - financial, employee, product, manufacturing, medical, research, and so forth. Whether your organization is a for-profit business, a government agency, a charity, an educational institution, or related to healthcare, all of these ideas apply – because every organization depends on information to support its goals and to deliver on its commitments. If your company is concerned about data quality, now is the time to get started! Thanks for the opportunity to talk with you!

 Note*:* This article contains some material published in the book *Executing Data Quality Projects: Ten Steps to Quality Data and Trusted Information™.* For more information on this title and other similar books, please visit *elsevierdirect.com.*

ABOUT THE AUTHOR

Danette McGilvray is president and principal of Granite Falls Consulting, Inc., a firm that help organizations increase their success by addressing the information quality and data governance aspects of their business efforts.

She is the author of *Executing Data Quality Projects: Ten Steps to Quality Data and Trusted Information*™ (Morgan Kaufmann, 2008). An internationally respected expert, Danette's Ten Steps™ approach to

information quality has been embraced as a proven method for both understanding and creating information and data quality in the enterprise. The Chinese-language edition is now available and her book is used as a textbook in university graduate programs. She spoke at the QIBRAS conference May 2011 in Sao Paulo, Brasil, and subsequently taught a 1.5 day workshop on her Ten Steps™ approach to data quality which was hosted at the ZipCode facilities.

You can contact Danette at danette@gfalls.com or through her website at www.gfalls.com.