

The Cost of Poor Data Quality

***How to make better business
decisions and positively affect
your bottom line***

IVEY

Richard Ivey School of Business
The University of Western Ontario



Decide with Confidence

This is a joint research paper by D&B Canada and the
Richard Ivey School of Business, University of Western Ontario

© The D&B Companies of Canada Ltd.

Data quality: Why should I care?

Data quality affects your business' profitability at both the top and bottom line. Poor data quality costs the U.S. economy six hundred billion dollars annually — 5% of the American GDP. This is in real tangible costs such as unnecessary postage, printing and staff overhead. This is a significant number and should have executives' attention. After all, its pure cost and reducing it can directly affect your bottom line.

But it's not just the bottom line we need to consider. Good data quality also improves revenue through effective customer segmentation. This established practice only works well with top-quality data supporting the decision making process. Surprisingly though, through 2007, at least 25% of critical data within Fortune 1000 companies will continue to be inaccurate².

Recently, McKinsey & Company, a leading strategy consulting firm, identified the following as some of the critical concerns for the executive community in 2006:

1. *Emergence of Asia*
2. *Offshoring*
3. *Growth in the world's biggest companies*

The quality of information that a company has access to is the difference between being able to meet today's business challenges or being steamrolled by them. Executives may protest against data overload, but in reality the real issue is that executives don't just need data to succeed, they need quality data that leads to actionable information. A recent *Harvard Business Review* article laid out the case why managers need to make decisions based on evidence and not simply gut feel³. Good data is the foundation of good evidence, and ultimately good decisions.

The information that executives need in order to make good strategic decisions is information about existing and potential customers and suppliers — and this need is expanding rapidly. Given current practice in many companies this will result in one of three outcomes:

1. *Costs to fix problems will increase exponentially*
2. *Data quality will get worse*
3. *A new perspective about data quality will emerge*

This document is about that third perspective; it is an executive call to action about data quality within the organizations they lead. Data quality is not an IT issue alone. The solution lies in mobilizing the organization. C-level attention is required to face challenges head on and resolve the problems.

C-level executives need to be concerned with facts like these...

- Through 2007, at least 25% of critical data within Fortune 1000 companies will continue to be inaccurate.
- Data quality is not an IT issue alone. The solution lies in mobilizing the organization.

The good news is that this is not only about saving a little money and making minor process improvements. What's truly exciting is what good data quality can bring to your organization. Consider the following two enterprises:

- A major credit card issuing bank that, instead of sending direct mail or pre-approved applications, can send approved credit cards that just require the customer to call and activate the card.
- A vehicle leasing company that reduces its target audience drastically, focusing its resources much more intently on more probable sales.

Both examples illustrate the power of good quality data. The credit card company assumes what to others might be unacceptable risk because it has confidence in its selection of each customer and the company's ability to deliver that card to the right customer. The leasing company believes that it can concentrate its resources on prime prospects thereby de-emphasizing other potential prospects because they are not ideal. Good quality data in both cases allows executives to make the right decision. And that's why data quality is important to executives. It facilitates better — more accurate, more consistent, more timely — decision-making.

Why does data quality matter?

In short, data quality matters because it affects the bottom line and an executive's ability to make strategic decisions. In a recent Bain & Company study, customer relationship management (CRM) and customer segmentation were rated as two of the top five executive tools to optimize business success⁴. However, customer data issues and consolidation processes often derail these initiatives⁵.

Marketing has long seen these effects. CSO Insights surveyed senior marketing executives from 281 companies across multiple industries, and they found that **47% of respondents had seen either a noticeable or severe impact on their marketing campaigns from missing or incomplete data**⁶. Martin, a senior manager at a global software firm, recalled his CEO's embarrassment and concern when the CEO was unable to accurately answer a customer's question about how much business the two companies did together. The customer might have left the meeting with the impression that their business was not that important to the software firm after all.

While these effects are pronounced, the impact of poor data quality is not limited to sales and marketing but extends throughout the enterprise. Gartner Vice President, and Data Management and Integration analyst, Ted Friedman, in a podcast⁷ earlier this year said that nearly "every business function will have direct costs associated with poor data quality". This includes sales, marketing,

"Every business function will have direct costs associated with poor data quality." Investments in enterprise systems "are at risk if the quality of data is poor. The result is a lack of acceptance and trust by the user community and pretty soon the investment is wasted."

Ted Friedman
Gartner, Vice President
and Data Management &
Integration Analyst

financial reporting, supply chain management and internal operations. Companies have recognized the need for enterprise-wide data integration by adopting tools such as Enterprise Resource Planning (ERP), Business Intelligence (BI) and CRM systems. However, Friedman remarked that [investments in enterprise systems “are at risk if the quality of data is poor.](#) The result is a lack of acceptance and trust by the user community and pretty soon the investment is wasted.” Without good quality data, enterprise system implementation is doomed to disappointing results or even outright failure.

What is good data quality?

So now we know why we want good quality data, but what is that exactly? There was a time when data quality simply meant that the data was free of explicit errors. However, emerging definitions of data quality focus not just on the data's characteristics but rather on its fitness for use.

For example, CFOs are being encouraged by their fellow executives to have their internal customers rate the quality of the data being provided by the finance department. In fact, in some companies this is becoming an important measure of the CFO's performance⁸. The data is evaluated on how easily and effectively it helps other departments in the company understand the business and make decisions.

In a study that investigated users' *perception* of data quality (as opposed to a technical evaluation of it) the following four dimensions of data quality were defined⁹:

Dimension	Characteristics
Intrinsic	Believable Accurate Objective Reputable
Contextual	Value-added Relevant Timely Complete Appropriate amount
Representational	Interpretable Easy to understand Consistent Concise
Accessibility	Available Secure

From this list, it is obvious that the faith in the process used to collect, analyze and distribute the data is an important factor in a user's trust in the quality of data. For example, whether a piece of data is believable or reputable will be significantly influenced by the process used to gather and manage it. Similarly,

ensuring that data mean the same thing in different contexts and are represented in a similar manner relies more on the process by which it is managed than simply by the end result.

Process is the key

Checking for quality issues deals only with the symptoms and does not get to the root of the problem. [As manufacturing companies learned long ago, the only way to ensure good data quality is to build quality improvement into the process.](#) Thinking about data quality as a business process rather than a technical object reveals it as an executive issue. The goal is to create an environment where all employees are empowered with the right data to make better strategic and operational decisions. Broadly speaking, there are three main phases to this process: gather, integrate and leverage.

Gather

There are hundreds of potential data sources for most companies. These include, but are not limited to, news and media, payment information, legal and corporate filings, and industry directories. Data collection is made more difficult when you are seeking information about a company that is recent (rather than several months old), private (rather than public), international (rather than local), and detailed (sales to a specific standard industrial classification code [SIC] as opposed to summary information such as total sales).

As information pours into an organization, it potentially overwhelms systems, processes and people.

In addition to company information, it is often important to gather industry or even firmographic data. A change in the sales of automotive vehicles is an important piece of data for making decisions about any firm in the automotive sector, not just original equipment manufacturers (OEMs). Further, a change in the value of certain foreign currencies can be important information about particular customers or suppliers.

Integrate

As information pours into an organization, it potentially overwhelms systems, processes and people. It can be difficult to figure out where to put each piece of incoming data, likely creating a large pile of miscellaneous "stuff", stored on sticky notes, random files or just "remembered".

It's not simply a matter of developing a good filing system. For each new piece of data, one of three implications could be drawn.

1. *What data needs to be updated?*

If a company has collected some credit information, you should be sure to assign it to the right customer or supplier in your database.

2. *Does this data mean that other data needs to be either deleted or archived?*

If you see that a customer has filed for bankruptcy, at a minimum their credit terms need to be reassessed.

3. *The data might enrich the existing dataset.*

You might collect information about two or three new potential customers located in new markets.

Therefore, through integration processes, new data can update, retire or enrich a company's existing database.

Leverage

Having a large store of data is one thing but it's not going to create any value on its own. The first step in value creation is to leverage the relationships within the data to create better information. For example, an executive's perspective is likely to be different if he realizes that three seemingly distinct customers are actually divisions of the same company, and there are four more divisions that currently aren't customers! Or, the CFO might be less inclined to extend additional credit if she realizes that that customer's total accounts payable have been increasing each month for the past year.

Executives who enable the process of gathering, integrating and leveraging data create a valuable resource that empowers managerial and executive decision-making.

Executives who enable the process of gathering, integrating and leveraging data create a valuable resource that empowers managerial and executive decision-making. However, a critical question is whether to develop this capability in-house or to draw on the expertise of a third party.

Don't go it alone

The good reasons for using an external partner — specialized expertise, broader experience, focus on quality, overall lower costs — are highly applicable to the data management process. **A third party data partner brings focus.** While for most organizations "good" data is enough, for a company focused on data provision "as good as possible" is the benchmark. Through relationships and specialized investments, such firms have access to a much larger range of data than most other organizations are able to secure and manage.

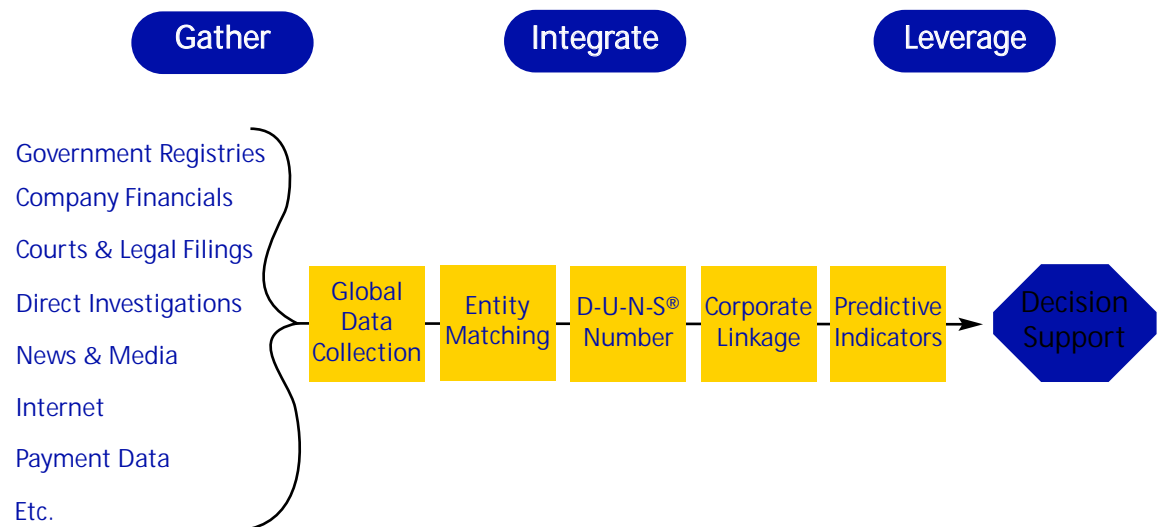
For example, many organizations are expanding or initiating business activities in China. Gathering data on Chinese firms is difficult. Even recording the names of Chinese companies is tricky for many systems. As new information on these firms is added to databases, the probability that inaccurate and duplicate entries are created is significant. This will make it difficult for business leaders to have reliable information upon which they can base confident decisions.

Most companies don't need, want, or can afford, to make the investments necessary to ensure data quality. Organizations that specialize in gathering and cleansing information are already investing in and developing methods to manage data quality — especially in new, emerging markets like China.

Third-party data providers however, cannot simply take over your entire data management process. They are partners, there to improve the quality and intelligence of your information. The goal of the partnership is to complement your internal data collection and storing processes and to work with you to develop an optimal information environment.

The DUNSRight™ process

D&B is a leading global provider of business information. Founded in 1841, D&B's business database has information on more than 106 million businesses around the world. In order to assure their customers the delivery of high quality data, D&B developed and refined the DUNSRight Data Quality Process. Data quality is defined by how accurate, complete, timely and consistent the information is. D&B's DUNSRight process is depicted below as it relates to the "gather, integrate and leverage" approach discussed earlier.



Global Data Collection

No other company can match D&B's worldwide data collection. D&B's global database holds information on more than 106 million businesses. Globally, D&B updates information 1.5 million times each day and information sources include news and media, court and legal filings, payment data, financial records, directories and most importantly, direct investigations.

While this may sound impressive, what is its value? Ultimately, it's in the decisions that become possible as a result of having depth and breadth of quality information. You can confidently:

- Assess risk of companies where it's otherwise difficult to gather data, such as international or private firms

- Identify prospects from data drawn from multiple sources so you get a complete picture of a business
- Target markets with value-added information such as Standard Industrial Classification (SIC) and payment history so you reach your most profitable prospects
- Monitor customers and suppliers for financial and operational changes

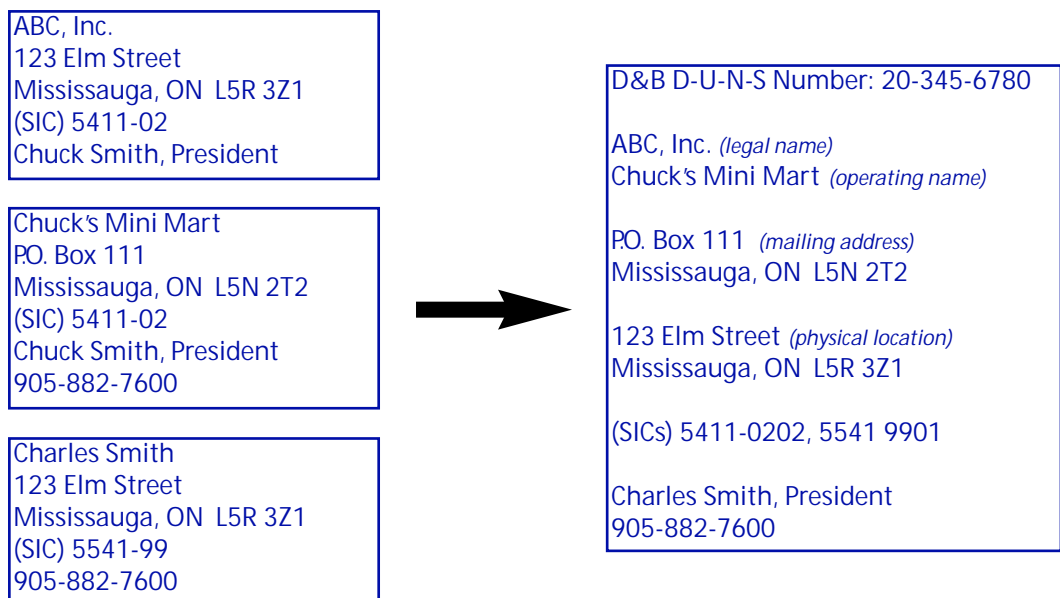
Many of these decisions are empowered with financial trade information that includes accounts receivable data on how quickly a company pays its vendors. This kind of information is unavailable to most companies and is but one example of the strength of D&B's data collection capabilities.

To be clear, **D&B information is not simply captured and made available**. Rather, it is subject to a rigorous data quality assurance process that includes over 2,000 checks and validations at the time of data entry as well as manual "reasonability reviews". In the past several years, D&B has increased its global data collection in emerging markets such as China, India, Brazil and Russia.

Entity Matching

For most companies, receiving this amount of information would leave thousands, if not millions, of data elements as single, separate pieces of information. Using complex and patented analytics, D&B matches data elements resulting in a single, accurate picture of each business entity. When D&B receives data, it is usually matched to an existing company record. If there is no matching company, it is held in an "unmatched data repository" until enough data exists to either match it to a known company (perhaps an existing company has a new operating name) or a new company is created.

In the figure below, the three companies on the left appear unrelated. When accurately matched, you realize this is one company. A clean and accurate database will help you more effectively understand and manage your customer & supplier databases.



For a client, this helps ensure that there are fewer duplicate or orphaned records in a company's customer and/or supplier database. In turn, this directly enables faster reporting and consolidation and less rework. Typos or missing data that naturally occur, when front line employees are serving customers, are cleaned up. Without this stage, it would be difficult to even know how many suppliers and customers you actually have.

D-U-N-S® Number

Once a business is confirmed and activity is evident, it is then assigned a D-U-N-S Number. D&B's most famous asset, [the D-U-N-S Number was defined in 1963 to exploit the rapid development of computing communications technology and enabled D&B to link and analyze information in entirely new ways](#). As a unique global identifier attached to operating entities, the D-U-N-S Number is never reassigned to another company, in any place, at any time. The D-U-N-S number has gained wide acceptance globally and is a common standard in the European Commission, United Nations and the United States government. It is recognized, recommended and/or required by more than 50 global, industry and trade associations.

At the heart of any CRM or ERP system is a relational database. This database relies on a unique identifier or primary key to establish relations between transactions and report consolidated results. In many companies, divisions create unique supplier and customer numbers.

However, these are rarely unique between divisions so that when a company tries to merge business information between divisions the numbers become irrelevant. Using names and addresses are often even more difficult. Using a primary key like the D-U-N-S number works across the company and provides tremendous value when trying to understand your customers and suppliers and your untapped market opportunities.

The D-U-N-S number has gained wide acceptance globally and is a common standard in the European Commission, United Nations and the United States government. It is recognized, recommended and/or required by more than 50 global, industry and trade associations.

Most executives are not going to get excited about a D-U-N-S Number. It's a made-up, arbitrary number applied to a business. It doesn't mean anything in and of itself. But it does open up tremendous opportunity to connect data across your organization so when two marketing divisions are talking about the same customer, they *really* are talking about the same customer. If two companies merge, systems based on such a number can be consolidated more easily, reducing expenses and increasing shareholder value. And finally, information can be acquired from external sources about potential new customers or suppliers. If this information is keyed to a unique identifier, it can be directly introduced into an organization's systems and brought into the strategic decision process. While the D-U-N-S Number may seem insignificant, it's the key to ensure that data can be represented consistently throughout the organization.

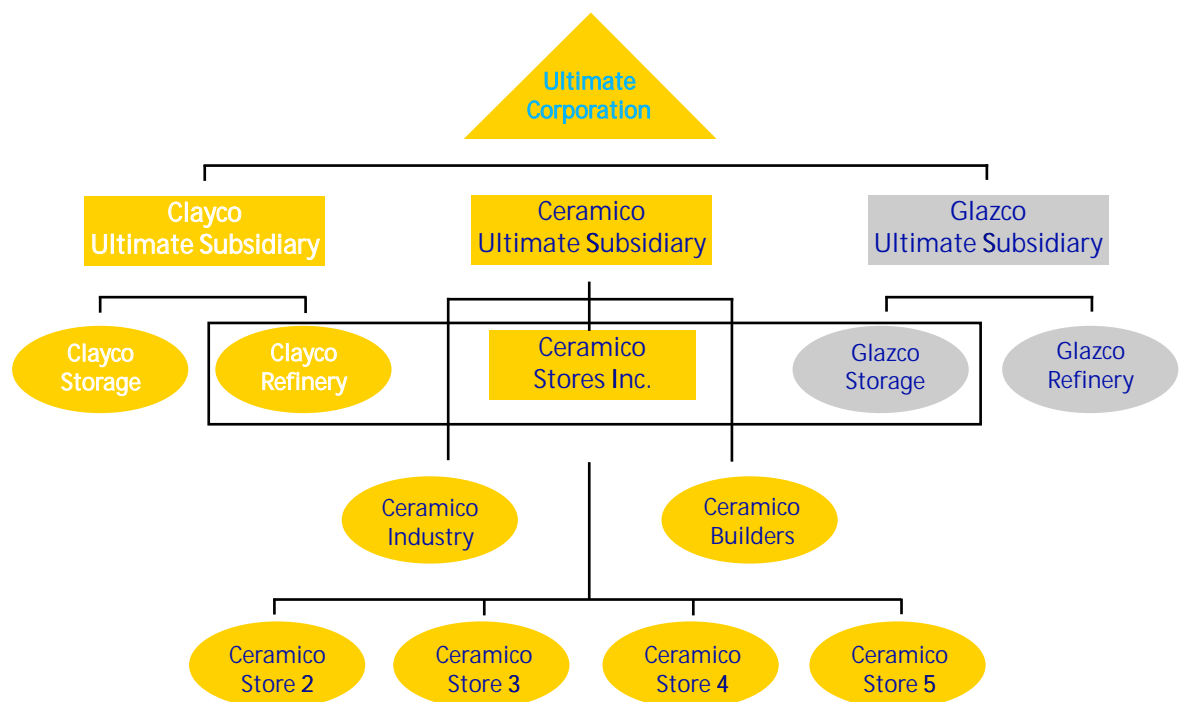
Corporate Linkage

A D-U-N-S number is tied to a specific operating entity. This means that a company with a headquarters, a distribution centre, a manufacturing plant and three retail outlets would have six different D-U-N-S Numbers assigned to it. **It would be impossible to get a single view of a company without understanding how companies are related.** Related entities, including subsidiaries and branches, are linked to one another via the D-U-N-S number.

Complex corporate structures make it difficult to see the relationships between different trading partners. For example, three seemingly distinct customers that look like this:



might actually be part of one larger organization that looks like this:



Lacking the detailed information about corporate linkages, it would be difficult for an executive to develop a strategic plan around supplier consolidation, risk exposure, and sales and marketing approaches.

Being able to link corporate entities provides very powerful capabilities for an executive.

- It becomes possible to understand your company's total exposure to a supplier or customer
- Warm calls for sales can be identified as sister divisions may already be customers

- You can begin to understand whether your suppliers are also your customers and make decisions about where to allocate your purchases
- You can also optimize your supplier spend directly impacting your bottom line
- You can investigate for money laundering and other financial fraud related activities

Without corporate linkages, it is simply not possible to ensure that your decisions are based on having an understanding of the complete picture. By having this insight, an executive can really begin to leverage the data that has been collected and integrated into their own corporate databases, which sets the environment for strategic planning.

Predictive Indicators

The final step in D&B's data quality process is the development of predictive indicators — a point of significant leverage. D&B has developed models within three categories:

- *Descriptive Ratings: an overall grade of past performance as a potential indicator of future performance*
- *Predictive Scores: the likelihood of a business remaining as an ongoing entity*
- *Demand Estimators: how likely a customer is to purchase a product*

Case Study for Predictive Scores

A bank developed a customer acquisition model that identified clients that were likely to have low acquisition cost and high spend potential and good credit risk and high retention value. This allowed them to focus their efforts on only these ultra-high potential clients.

Predictive indicators allow an executive to leverage information to gain the most rare executive resource: time. By identifying either more promising opportunities or more immediate risks with the use of these indicators, an executive can focus their time where it will make the most difference.

Predictive indicators truly leverage an organization's raw data into high quality information ready to support important strategic decisions.

The DUNSRight Impact

Quality Dimension	Where DUNSRight Contributes Most Significantly				
	Global Data Collection	Entity Matching	D-U-N-S Number	Corporate Linkage	Predictive Indicators
Intrinsic	●			●	●
Contextual		●		●	●
Representational			●	●	●
Accessibility	D&B works with its clients to develop distribution and dissemination methods				

Summary

Data quality problems sneak up on an organization. And, there's not one big bang to fix them. With mergers and acquisitions, startups and bankruptcies, globalization and growth, the business environment changes too quickly for data to ever be "fixed" without implementing a management process. This process affects the entire organization and can only be led by a C-level effort. It's too important to delegate and too critical to delay.

Executing this process is demanding. It requires excellence in gathering, integrating and leveraging data. Most firms will do it better, faster and cheaper in partnership with a third party. In this paper, DUNSRight has been presented as a path to good data quality. If you take the time to uncover the symptoms of poor data quality in your organization, this process may be part of your path toward a stronger bottom line.

About D&B

Dun & Bradstreet (D&B), is the world's leading provider of Knowledge, insight and business information and has been enabling business-to-business commerce for 165 years. D&B owns detailed information on almost 106 million businesses worldwide and we've helped our clients make business decisions with confidence. D&B helps organizations minimize risk in their accounts receivable portfolio, optimize opportunity within their existing customer base, identify potential customers with a high propensity to purchase, identify companies with a high level of financial stability, leverage their supply chain to reduce expenses and meet legislative compliance expectations.

The D&B database is powered by a patented data quality process called DUNSRight™ which helps ensure accurate, timely and consistent information on businesses in almost 200 countries, in 95 languages or dialects, and representing more than 180 monetary currencies. At D&B we collect, aggregate, edit, and verify data from thousands of sources daily so that our customers can use the information to make profitable decisions. D&B's ability to turn an enormous stream of data into quality business information is what sets us apart from our competitors.

It's no wonder that 90 percent of the Business Week Global 1000 rely on D&B as a trusted business partner. For more information please visit www.dnb.com

About the Author

Darren Meister is an Assistant Professor of Information Systems and the Robert V. Brouillard Faculty Fellow at the Richard Ivey School of Business. His research has appeared in journals such as Management Science and MIS Quarterly. For ten years, he has taught in MBA and Executive Education programs in Canada, Hong Kong, Mexico and Europe. His PhD is from the University of Waterloo and he is also a graduate of the University of Cambridge. Darren can be reached at (519) 661-3859 or dmeister@iveyca.



Notes

- ¹The Data Warehousing Institute, Data Quality and the Bottom Line: Achieving Business Success Through a Commitment to High Quality Data, February 1, 2002.
- ²Gartner Report , Predicts 2006: Emerging Data Management Drivers and Strategic Imperatives, Ted Friedman, Bill Gassman and David Newman, November 2005.
- ³Jeffrey Pfeffer and Robert I. Sutton, Evidence-Based Management, Harvard Business Review, January 2006.
- ⁴Bain and Company, Management Tools and Trends, Darrell Rigby and Barbara Bilodeau, 2005.
- ⁵Forrester Research, The Forrester Wave™: Customer Data Integration, Q2 2005, June 13, 2005.
- ⁶CSO Insights, 2005 Executive Report: Target Marketing Priorities Analysis", 2005.
- ⁷Forrester Research, The Cost of Poor Data Quality, 1 March 2006, http://www.gartner.com/it/products/podcasting/attributes/attr_145611_301.mp3
- ⁸Thomas Wunder and Jeff Thomson. "Want to Be a More Effective CFO? Ask the Audience!" Strategic Finance, Jan 2006, vol. 87, no. 7, pg. 36-42.
- ⁹Richard Y. Wang and Diane M. Strong, "Beyond Accuracy: What Data Quality Means to Data Consumers", Journal of Management Information Systems, Spring 1996, vol. 12, no. 4, pg. 5-33.



Decide with Confidence

www.dnb.com