

Overall Approach to Data Quality ROI

By: William McKnight, founder and president,
McKnight Associates, Inc.



Sponsored by: Firstlogic, Inc.

THE IMPORTANCE OF DATA QUALITY

Data quality is an elusive subject that can defy measurement and yet be critical enough to derail any single IT project, strategic initiative or even a company as a whole. The data layer of an organization is a critical component because it is so easy to ignore the quality of that data or to make overly optimistic assumptions about its efficacy. Having data quality as a focus is a business philosophy that aligns strategy, business culture, company information, and technology in order to manage data to the benefit of the enterprise. Put simply, it is a competitive strategy. One day, like operational excellence, rich product features, everyday low prices, high product quality and short time-to-market, data quality will be expected by our markets. In the meantime, each company has the opportunity to differentiate itself through the quality of its data. Leading companies are now defining what the marketplace data quality expectation will be.

A parallel trend to data quality improvement in the marketplace is the come back of return-on-investment (ROI) measurement systems for technology-based initiatives. No longer is it acceptable to “throw money at problems,” target soft measures or lack accountability for results with technology spending. The approach of targeting ROI is even viable for efforts to improve the quality of data in a company and many executives are demanding payback for quality initiatives.

It is important to note that there are many benefits that accrue from improving the data quality of an organization. Many of these benefits are “intangible” or unreasonable to measure. Benefits such as improved speed to solutions, a “single version of the truth,” improved customer satisfaction, improved morale, an enhanced corporate image and consistency between systems accumulate but an organization must selectively choose those benefits to perform further analysis on and convert to hard dollars. ROI must be measured on hard dollars.

A program approach to data quality is required to measure data quality ROI. Data quality improvement is not just another technology to implement. We must change our way of doing business to fully exploit data. Investments in the technologies as well as in organizational changes are necessary to reap the full rewards. Data quality is right in the “sweet spot” of modern business objectives that recognize that whatever business a company is in, it is also in the business of data. Those companies with more data, cleaner data, accessible data and the means to use that data will come out ahead.

But how do organizations begin and justify data quality improvements? The cleansing process and maintenance will cost money and may require dedicated staff. Management may not release funds dedicated to data quality improvement until such time as the cleansing will improve the reliability and accuracy of key business processes like trending, analysis, or billing for product sales. Data quality can — and must — measure its success based on its contribution to the improvement of such objectives.

However, there has not been a methodology to articulate and improve data quality ROI, until now. You can't improve what you can't measure. So, we need a means for measuring the quality of our data warehouse. Abstracting quality into a set of agreed data rules and measuring the occurrences of quality violations provides the measurement in the methodology, which was developed in conjunction with IT best practice successful efforts with data quality.

OVERALL APPROACH TO DATA QUALITY ROI

Step 1: System Inventory

In an environment that has not focused on its data quality, there is likely to exist many systems that require clean data to operate well and make or influence correct decisions.

In the System Inventory step, company systems are surveyed and prioritized according to their use of and need for quality data. These high priority systems rely on its data being accurate in order to perform their intended function and can be found across the artificial boundary between operational systems and decision support or data warehouse systems. Systems needing clean data come from all categories including Enterprise Resource Management (ERM), Customer Relationship Management (CRM), Data Warehouse, E-Commerce and Call Center or customer service.

For the purposes of this paper, we will use a targeted marketing example to better explain the data quality ROI process. The concepts applied to this target market example can be applied to other IT systems whether it be your business intelligence, CRM applications or other critical applications.

In the next step, we'll drill down on the detail in those systems.

Step 2: Data Quality Rule Determination

Data quality can be defined as a lack of intolerable defects. There is a finite set of possibilities, which can constitute data quality defects and which categorize all data quality rules, as found in the list below. The rules generated in this step are the rules that you wish the data to conform to. These rules can apply wherever important data resides. This step will define the data quality rules that we would like our data to live up to.

- Referential Integrity (RI) refers to the integrity of reference between data in related tables. For example, product identifiers in sales records need to also be found in the product tables. RI is essential to prevent the users from navigating to a “dead end.” Many transactional systems utilize database-enforced RI, but many data warehouse systems do not due to the performance and order restrictions it places on data loading. But RI is still important and must be enforced somehow. In the targeted marketing system, for example, bad RI may mean sales with an undefined product and/or to an undefined customer, which effectively takes that sale out of the analysis.
- Field(s) assigned to uniquely identify a business entity, like a product identifier, should also be unique. It would constitute a violation when the same identifier is used to identify multiple products. This also has implications for RI. Not only do we not want “dead end” navigations in the data, but we also want a singular reference. In the targeted marketing system, for example, a single product for product identifier number 123 would be expected.
- Cardinality restricts the volume of referencability for any given value. For example, for the targeted marketing system, we may expect to find between one and three addresses for a customer – no more and no less. Other relationships may be based on 1-to-1, 1-to-many (infinite) or 0-to-many. The 0-to-many is interesting because it allows for the entry of a product without a sale. There are reasonable conditions on the cardinality in this case whereby these products would be “new” products – perhaps products with an entry date in the last 14 days only.

These first three components of cleanliness can be derived, for databases, from review of the data model.

- Subtype/Supertype constructs may also be found in the model. For example, the targeted marketing system may represent a customer's marital status as married, divorced, widowed or single. Different constructs may apply to each state, like former-spouse data for divorced persons. Subtype/supertype bounds the states of a given business entity. The logical progression through the various states for a business entity (i.e. prospect, customer, good customer, former customer) is another quality check.

Having a robust logical data model from which one can glean rules is an asset in determining rules for your data quality process. If the data model is missing or inadequate, this may be required before completing this step.

There are also measures associated with data value appropriateness such as the next two rules:

- Are there reasonable domains for values? This is especially useful for analyzing numeric data, whose data types can allow an enormous range of values, only a subset of which are reasonable for the field. These include items like incomes $> \$2,000,000$ or < 0 ; last-year purchases $> \$500,000$ (for example a department store) or < 0 ; and offers mailed $> 1,000,000$ or < 1 . All of these values would be considered unusual at the least and quite possibly incorrect. That's why we want to flag them for investigation.
- Are data columns being used for multiple meanings? Data types constrain the values in a column ... to a degree. Any mix of 0-20 characters can go into a character (20) data type column. However, if this is a Name field, there are some characters that you would not expect to find in certain columns such as % and \$. These would be flagged that the field contained inappropriate data. There are also numerous misspellings and incorrect alternative spellings of last names.

There are various formatting errors that can be found in the field. Typical of these formatting errors found in name columns include, for example:

- A space in front of a name
- Two spaces between first and last name and/or middle initial
- No period after middle initial
- Inconsistent use of middle initial (sometimes used, sometimes not)
- Use of ALL CAPS
- Use of "&" instead of "and" when indicating plurality
- Use of a slash (/) instead of hyphen (-)

In environments where original data entry is “free form,” unconstrained and without the use of master data as a reference, many of the violations on the previous page will be found.

Another type of typical data quality error involves columns that should only contain data if certain data values (or null) are present elsewhere. For example, in the customer table where some customers are organizations and others are individuals, only organizations will have SIC code, organization size, etc. (unless the presence of that data actually represents data for the individual’s organization).

Some systems derive data from data in other systems. For example, while one system may contain discount amount and unit price, the discount percentage — a simple division calculation — could be a new calculation. Checking the veracity of calculations like customer lifetime value for our targeted marketing system is a common quality check.

The existence of data consistently through all of its history, without gaps, indicates users will not be left to wonder what happened during missing periods, is another common quality check.

There is also incorrect data. For example, the inconsistent use of initials and misspellings or inappropriate spellings (i.e., William McNight instead of William McKnight).

Finally, does the data conform to the expected set of “clean” values for the column? For example, in the gender column we would expect to find M, F, U (unknown) or the like.

Data quality rule determination is determining the rules that will govern our data. Data characterization, the next step, determines where we are with the quality of the data.

Step 3: Data Profiling

It's one thing to sit back and pontificate about to what rules the data ought to conform. The next step is to determine the data quality with a data profiling and prioritization exercise. Usually no one can articulate how clean or dirty corporate data is. Without this measurement of cleanliness, the effectiveness of activities that are aimed at improving data quality cannot be measured.

Measuring data quality begins with an inventory. By taking account of important data across several tangible factors that can be used to measure data quality, you can begin to translate the vague feelings of dirtiness into something tangible. In so doing, focus can be brought to those actions that can improve important quality elements. Ultimately data quality improvement will be performed against a small subset of data elements, as you will find most elements already conform to standard. The subset must be selected carefully, however. Another way to put it is that data quality initiatives will not be comprehensive across all corporate data elements and all possibilities.

Data profiling can be performed with software programs like IQ Insight™ from Firstlogic, SQL or similar queries against the data showing the spread of data in the systems and checking for rule adherence. For our targeted marketing example, we may find that the gender column is distributed as shown in Figure 1 below.

This distribution shows our valid values of 'm' (male), 'f' (female) and 'u' (unknown) occurring in 42,855, 44,583 and 5,986 records respectively. Some of the other values (1, 2, y, x) could be caused by certain data entry people following a different legend (i.e., 1=male, 2=female). The '!' could be a flat out error. These questionable values need to be analyzed.

Note that, depending on context, 'unknown' may actually be considered a bad value as well. In our case, we are simply recording the fact that we don't know.

Figure 1: Data Distribution

Value	Num. Occur.	Pct. of Overall
m	42855	0.43
f	44583	0.45
u	5986	0.06
1	4070	0.04
2	1265	0.01
y	986	0.01
x	254	0.00
!	1	0.00

Step 4: Data Quality Scoring

Once the systems and the data quality rules are identified and the data is characterized, scoring the data quality needs to be performed. Scoring will represent the state of the data quality for that rule. System scores will be an aggregate of the rule scores for that system and the overall score will be a prorated aggregation of the system scores.

Scoring is a relative measure of conformance to rules. For a given rule, it could be as simple as the percentage of opportunities for rule enforcement that are positively met:

$$\frac{\text{Adherence}}{\text{Possibilities}}$$

For example, if 94 percent of genders conform to the desired values, as in the Step 3 example, that score is 94 percent.

Since many of the results will (hopefully) be above 99 percent, to provide a higher level of granularity to the scoring, you may set a floor at 50 (or any number) and measure the adherence above that number. For example, using the gender example, the score would be (94-50)/50 or 88 percent instead of 94 percent.

Averaging all the data quality scores from the system results in an overall system score.

Figure 2: Rule Reference Example

Rule 1: All product identifiers in sales found in sales table	98%
Rule 2: All product identifiers unique	95%
Rule 3: Only products added in last 14 days don't have an associated sale	92%
Rule 4: Former spouse information kept for divorced persons only	95%
Rule 5: All customers moved through states of prospect, customer, good customer, former customer in correct progression (no prospect to good customer, no time gaps in assigned states)	93%
Rule 6: All incomes between 0 and \$2,000,000	89%

Rule 7: All last-year purchases between 0 and \$500,000	93%
Rule 8: Number of mailings for offers between 1 and 1,000,000	99%
Rule 9: Names are without % and \$	84%
Rule 10: No space in front of name	99%
Rule 11: Two spaces between first and last name and/or middle initial	96%
Rule 12: No period after middle initial	100%
Rule 13: Consistent use of middle initial	97%
Rule 14: Only organizations have SIC codes	97%
Rule 15: All names verified correct	89%
Rule 16: Gender has only M, F or U	94%

Rule score average = Data Quality Score for system = 94.375 percent.

Simple averaging may not be as effective as a weighting system that weighs the more important scores higher. Regardless, consider the scoring of a system be set such that it approximates the following:

Figure 3: Data Quality Scoring

Score -----	Meaning -----
99-100	Excellent data quality, well fit-for-purpose
96-98	Good data quality, but data quality still needs proactive work
90-95	Average data quality, data quality needs proactive work
81-89	Below average data quality, data quality is likely having a negative impact on system's operation
65-80	Poor data quality; system is likely producing poor results
0-64	Very poor data quality; do not run system; mobilize effort to improve data quality immediately

Step 5: Measure Impact of Various Levels of Data Quality

ROI is about accumulating all returns and investments from a project's build, maintenance, and associated business and IT activities through to the ultimate desired results all while considering the possible outcomes and their likelihood. Using ROI for justification is reducing the proposed net change in activities to their associated anticipated cash flow. Often, a cost of money is used to reflect in today's numbers the present value of expected cash flows in the future.

Various ROIs can be computed in anticipation of a project for its justification showing the various potential outcomes distributed across their probability of occurrence – a probability distribution for the project. The variables include numerous critical components of which data quality is one. The appendix shares some industry-specific examples of projects and how data quality might affect outcomes.

All things being equal, the data quality scores of a system will lead to different system results, and hence different ROIs.

In our targeted marketing example, it will be impossible to mail to bad addresses and we will get returned mail at incorrect addresses. More importantly, we will segment customers inappropriately and therefore market to them inappropriately using poor quality data. A bad product reference in sales, for instance, leads to incorrect profiling of sales habits. Incorrect customer state assignment or income level assignment also leads to incorrect profiling and lower returns from marketing efforts.

The measurement of all of these quality rules is the data quality score for the system. Another way to look at it is the data quality rules were arrived at not by intellectually determining how the data should look like but at the cost to the function of the system if the data lacked quality.

For example, if the data quality score for the Targeted Marketing database is 85, we can expect to achieve a 3 percent return on our marketing programs which leads to a 101.92% percent ROI. A score of 80 underachieves the potential return and yields 2.5 percent return on marketing dollar which leads to a 64.47% percent ROI. A score of 90 produces a higher achieving Targeted Marketing Program, perhaps a 4 percent marketing return and 175% percent ROI.

The ROI of this project would be calculated by the profit of the products sold tied to this promotion, divided by the investment in creating and distributing the promotions. Potentially some form of customer lifetime value could also be added to the returns if the company deems new customer relationships to be long-term. With better quality data in the system, allowing for more opportunity for the supposedly good promotions to reach their proper audiences, the return on marketing efforts is higher and the overall ROI higher.

At a DQ score of 90, potentially there are 110,000 prospects with quality data instead of 100,000 at the level of 80. Also, the quality of the data would yield higher accuracy in the targeting algorithms meaning 4% are likely to respond instead of 2.5%. All this at a nominal increase in investment for the additional mailings and more targeting activity as shown in the following table.

Figure 4: Cost/Benefit Analysis

DQ Score	Prospects Reached	Return on Marketing	Avg. profit	Return (prospects X ROM X avg. profit)	Investment*	ROI (return-investment/investment)*
90	110,000	0.04	\$250	\$1,100,000	\$400,000	175.00%
85	105,000	0.03	\$250	\$787,500	\$390,000	101.92%
80	100,000	0.025	\$250	\$625,000	\$380,000	64.47%

*Without considering costing for DQ improvements

So, it behooves us to improve the quality of the data to improve the anticipated return on the project. But at what cost? In the Data Quality Improvement step, we will detail the data quality actions and apply cost to arrive at the ideal data quality level.

Step 6: Data Quality Improvement

It is much more costly to fix data quality errors in downstream systems than it is at the point of origin. Primarily, if data entry in operational systems can be constrained to enter valid sets of values, this will be a great help.

For example, if a name is being entered, automatic fill-in is a handy data quality assist. If a product is being entered, it should be from an up-to-date drop-down list. If appropriate, data being entered that is “new” can be compared against previously entered similar (including phonetically similar) data and the entry function can be verified with “are you sure” type prompts.

Consider a field that serves no apparent operational purpose, but has a strong analytical purpose. If 50 percent of the data being entered into this field is unusable, our transformation strategy has no hope of salvaging much use out of this field. This is often the case with call center data entry where agents are rewarded based on the volume of calls taken. It’s easy to ignore certain fields on the screen and accept the default (or worse – enter what’s convenient but incorrect) when this is the case.

Fixing early-entry systems is not always possible since many deal with complex, older technology operational systems that are not easily changed. Many of these systems were written without an understanding of data quality. Changing them will be many months and involve a major cross-functional effort.

Downstream systems, such as data warehouses, benefit from quality improvement in systems that improve the quality at the point of origin.

Wherever the changes need to take place, there is a reasonable set of actions that can be performed on data to increase its quality:

- a. **HOLD OUT:** Record(s) that are being brought into the environment from other environments are held out of the main data areas in cases of gross rule violation. They may be placed into ‘holding’ tables for manual inspection and action. If adapting this approach, be sure procedural reviews are held quickly because data will be held out of the system until it is accounted for.
- b. **REPORT ON IT:** Data quality violation is reported on but data is loaded and will remain in the data store for the system. Typically the handling of these violations involves nothing more than a post-load report on just-loaded data that creates visibility into the situation. These reports are not useful unless they are reviewed from the appropriate systems and business personnel.

It is also quite possible that referential sets of data, not singular rows, are affected when other forms of data quality defects are detected and the appropriate action should be taken on that data as well as the data directly affected by the quality defect. As with all data quality checks, the data just loaded is the only data that needs to be checked.

- c. "CHANGE DATA": Transform data to a value in a master set of "good" values (i.e., Texas is changed to Texas) or otherwise apply rules to improve the data.

The scoring in step 4 should occur AFTER any data quality improvements are done to the data during its normal cycles. For example, apply the scoring to a data warehouse system that changes operational data after it makes the changes during its extraction-transformation-loading cycle.

Scoring should only be affected by how the data looks after its data quality actions complete (or periodically if the system receives no batch-style loading).

The willingness to spend on data quality improvements should be entirely determined based on the ability to advance the data quality scores, which have been correlated to project return.

Figure 5: Cost/Benefit of Adding Data Quality

DQ Score	Prospects Reached	Return on Marketing	Avg. profit	Return (prospects X ROM X avg. profit)	ROI (return-investment/ investment)*	ROI (return-investment/ investment)*	Add'l cost to improve DQ to this level	Revised Project ROI (return-investment + DQ cost/ investment+ DQ cost)
99	115,000	0.0425	\$250	\$1,221,875	\$410,000	198.02%	\$300,000	72.10%
90	110,000	0.04	\$250	\$1,100,000	\$400,000	175.00%	\$200,000	83.33%
85	105,000	0.03	\$250	\$787,500	\$390,000	101.92%	\$75,000	69.35%
80*	100,000	0.025	\$250	\$625,000	\$380,000	64.47%	0	64.47%

*Without considering costing for DQ improvements

In this case, the spending necessary to get the data quality score of 90 would be well worth it, given the improved ROI it will generate. Note a DQ score of 99 would actually REDUCE the project ROI due to the high cost of achieving that level of DQ. Data quality is directly related to the success of most projects. High quality scores in the data used generally mean higher project success. But neither the quality level nor its influence on project ROI needs to be vague and unstructured. By following the steps of the methodology, you can understand the current level of quality, continually improve it and generate higher returns for the bottom line through a data quality program.

SUMMARY

There's little doubt that in the past, placing a value on the data being collected in organizations' systems was a difficult proposition. Today, we're in a market economy in which capital spending is more closely scrutinized than ever before. Those at the executive level of organizations demand tangible returns on the investments they make in technology. This paper has outlined six key steps that you can take and put into process to help you realize that tangible ROI on your data quality initiative. Those six steps include:

1. System Profiling
2. Data Quality Rule Determination
3. Data Profiling
4. Data Quality Scoring
5. Measure Impact of Various Levels of Data Quality
6. Data Quality Improvement

We're in an exciting time in the history of corporate computing. Organizations are beginning to wake up to the fact that the data they collect and manage should be viewed as a corporate asset. It is the one thing that separates competitors, since they don't know how each other are leveraging their critical information. Ultimately, the quality of your data can be your competitive advantage or disadvantage — the choice is yours.

APPENDIX: INDUSTRY FOCUS

Data quality is applicable to many cross-industry initiatives, such as the targeted marketing example used in this paper. However, industries tend to cluster around a set of specific needs that lend themselves to data quality initiatives. This section will highlight those common initiatives in selected industries that can be most served by data quality initiatives to be used in Step 2 of the methodology — Data Quality Rule Determination.

In these applications, data quality scores that are properly set up are highly correlated to project return. Scores for multi-purpose systems can be used for all the purposes of that system. For example, the data warehouse data quality score relates to all of its applications.

RETAIL

Targeted Marketing –

This ROI comes from the net sales of the products sold on marketing efforts geared to specific customers and prospects due to their characteristics and aptitude to buy. Targeted marketing uses customer data to validate macro customer characteristics and develop customer-oriented marketing programs.

A high data quality score in the targeted marketing database means customers will have unique reference, consistent spelling, and names and addresses that are valid. This increases the percentage of customers that will be “contact-able” as well as those who will accurately participate in demographic analysis and clustering for promotions. A customer who receives a promotion in the mail (versus one who does not) is much more likely to buy.

Likewise, demographic and other profiling is done because of the additional knowledge it imparts into the marketing process and hence the higher ROI generated. A low quality score indicates a lack of sound, empirical information being brought into that process commensurately leaving ROI on the table.

Churn Management –

The ROI for churn management comes from the lifetime values of the customers retained due to special efforts extended on the company's part to retain those customers. A high data quality score in the churn management system implies that there is accurate tracking data for those items important to a customer's propensity to churn. Accurate tracking data could include contacts and survey results appropriately attributed to customers and appropriately attributed to its positive/negative connotation and sales with accurate and complete product and customer information.

A high data quality score here also means the measures associated with customer lifetime value determination, namely sales and demographics, are accurate. When these are accurate, we correctly identify the correct customers to create churn prevention programs for. We apply dollars wisely towards the retention of the best customers and we will generate higher ROI from our churn management program.

Market Basket Analysis –

The ROI for market basket analysis comes from special efforts extended on the company's part to place product according to intra-transaction buying patterns. Once the market basket is known, logical placement and marketing decisions use the information to establish products to place together or separate depending on the goals.

A high data quality score in the market basket system means there is minimal missing product data in the sales and products are logically and uniquely identified. Although always difficult to measure, this will lead to higher returns from the product placement activities undertaken as a result of market basket analysis.

FINANCIAL

Cross-Selling-

Cross-selling is very popular in financial intuitions these days as companies look to mine their customer econometric, lifestyle and life stage data to sell products in different product categories to customers. This is similar in nature to targeted marketing and market basket analysis under retail where a high quality score greatly improves the veracity of the effort. By having high sales and customer quality, cross-selling efforts can be trusted to generate higher returns as opposed to “winging it.”

Channel Parity-

Consistency of information across customer-facing channels is a means of generating high levels of customer satisfaction, leading to reduced churn and more reception to cross-selling efforts. ROI from these projects is difficult to measure as they get into the higher customer satisfaction/higher customer lifetime value chain. However, better parity means better customer satisfaction, which starts the chain towards ROI.

High data quality in the various channels needs to include the consistency of quality information between systems that interface with customers as ascertained through the cross-checking of data.

HEALTHCARE

Procedure Analysis –

Perhaps no application of data quality is more important as data quality for procedure analysis – the use of data to route procedures to best-of-breed providers likely to provide the best medical service and reduce the likelihood of follow on procedures and claims. High quality in this system means better-served and healthier customers as well as high ROI.

A high quality score means procedures are completely and accurately tracked and can sway the ROI of such projects by as much as 25 points. These systems rely highly on mining of the data to determine unhealthy and undesirable claim follow-up activity. Missing or incorrectly attributed data will always bring down the ROI potential of mining since mining needs as much data as it can in order to make accurate analysis.

MANUFACTURING

Inventory Tracking –

Inventory tracking can almost be considered part of the “lights on” operation of a manufacturing business. However, higher levels of inventory tracking can certainly lead to high ROI by the improvements they make over simple systems. A high quality data score in the inventory tracking system means the data accurately tracks inventory by accurately tracking products and locations with valid dating of product movement data.

A high data quality score is correlated to high ROI through the increased savings in inventory.

TELECOMMUNICATIONS

Contact Center Analysis –

Those chartered with the origination of data in the contact centers are stewards of its potential to the organization beyond the efficient running of the contact centers on an individual basis. Mined contact center data provides the foundation for customer relationship management strategies, agent activity and performance management, product management, promotion effectiveness management and infrastructure management.

High data quality means call data is appropriately logged, assigned to customers and typecast according to content. All of the above returns are highly correlated with high quality data. For example, agent activity and performance ROI manifests itself in more efficient agent utilization, which reduces agent costs and improves service levels which, in turn, results in more satisfied customers and enhanced ability to retain new and profitable customers. Service levels are improved through an understanding of agent capabilities gained from monitoring contact loads. Higher quality contact and agent data can create more efficient staffing situations with improved ROI from a lower TCO of the cost to serve via a contact center.

About The Author



William McKnight is founder and president of McKnight Associates, Inc. (www.mcknight-associates.com), a consulting company specializing in data warehousing and CRM solutions. Mr. McKnight is an internationally recognized expert in data warehousing and CRM with over 15 years of experience architecting and managing information and technology services for organizations.

Mr. McKnight is a frequent and highly rated speaker at major worldwide conferences and private events worldwide, providing instruction on customer intimacy, return-on-investment, architecture, business integration and other business intelligence strategic and architecture issues. Mr. McKnight is a published author and a columnist in Data Management Review for the column “The CRM-Ready Data Warehouse.”

Mr. McKnight is a regularly featured expert on data warehouse/business intelligence and CRM at major conferences, is widely quoted on data warehouse issues in publications such as Computerworld and Information Week and has been featured on several prominent expert panels. He is an expert witness and a judge for Best Practices competitions. Mr. McKnight is the former executive of a recognized Best Practices information management program.

About Firstlogic



Firstlogic specializes in enterprise-wide information quality solutions and services for e-business, data warehousing, and mail applications. Firstlogic provides information quality tools that enhance and match business-critical information anywhere it is touched, stored, or moved across the enterprise. In addition to developing commercial solutions, Firstlogic partners with many systems integrators, consultants, and original equipment manufacturers to provide its unique technology to their end-user customers. Founded in 1984, Firstlogic today serves thousands of customers worldwide, including Fortune 1000 companies in the e-business, financial, insurance, healthcare, direct marketing, higher education, and telecommunications markets. For more information, visit the company’s Web site at www.firstlogic.com

