

MARCH 2006

Taking Data Quality to the Enterprise through Data Governance

By Philip Russom

TDWI REPORT SERIES

tdwi
THE DATA WAREHOUSING INSTITUTE

A 101communications Publication

Research Sponsors

Business Objects

Collaborative Consulting

DataFlux

DataLever

Firstlogic

IBM

Informatica

Trillium Software

Taking Data Quality to the Enterprise through Data Governance

By Philip Russom

Table of Contents

About the Author	2
About TDWI	2
About the TDWI Report Series	2
Acknowledgments	2
Executive Summary	3
Research Methodology	4
Demographics	4
The State of Data Quality	5
Defining Data Quality	5
Best Practices for Improving the Quality of Data	6
Evolving Perceptions, Realities, and Initiatives	9
Liability and Leverage—A Case for Data Quality	10
The Problems of Poor-Quality Data	10
The Benefits of High-Quality Data	13
Data-Quality ROI and Budget	14
Taking Data Quality to the Enterprise	15
The Scope of Data-Quality Initiatives	15
Most Data-Quality Trends Lead to Enterprise Use	17
Data Governance and Enterprise Data Quality	19
The State of Data-Governance Initiatives	19
Anatomy of Data-Governance and Stewardship Programs	21
Recommendations	23



About the Author

PHILIP RUSSOM is the senior manager of research and services at The Data Warehousing Institute (TDWI), where he oversees many of TDWI's research-oriented publications, services, and events. Prior to joining TDWI in 2005, Russom was an industry analyst covering BI at Forrester Research, Giga Information Group, and Hurwitz Group. He also ran his own business as an independent industry analyst and BI consultant, and was contributing editor with *Intelligent Enterprise* and *DM Review* magazines. Before that, Russom worked in technical and marketing positions for various database vendors. You can reach him at prussom@tdwi.org.

About TDWI

The Data Warehousing Institute (TDWI), a division of 101communications LLC, is the premier provider of in-depth, high-quality education and training in the business intelligence and data warehousing industry. TDWI is dedicated to educating business and information technology professionals about the strategies, techniques, and tools required to successfully design, build, and maintain data warehouses. It also fosters the advancement of data warehousing research and contributes to knowledge transfer and the professional development of its Members. TDWI sponsors and promotes a worldwide Membership program, quarterly educational conferences, regional educational seminars, onsite courses, solution provider partnerships, an awards program for best practices, resourceful publications, an in-depth research program, and a comprehensive Web site.

About the TDWI Report Series

This series is designed to educate technical and business professionals about new business intelligence technologies, concepts, or approaches that address a significant problem or issue. Research for the reports is conducted via interviews with industry experts and leading-edge user companies, and is supplemented by surveys of business intelligence professionals.

To support the program, TDWI seeks vendors that collectively wish to evangelize a new approach to solving business intelligence problems or an emerging technology discipline. By banding together, sponsors can validate a new market niche and educate organizations about alternative solutions to critical business intelligence issues. Please contact Wayne Eckerson (weckerson@tdwi.org) to suggest a topic that meets these requirements.

Acknowledgments

TDWI would like to thank many people who contributed to this report. First, we appreciate the many users who responded to our survey, especially those who responded to our requests for phone interviews. Second, our report sponsors who diligently reviewed outlines, survey questions, and report drafts. Finally, we would like to recognize TDWI's production team: Jennifer Agee, Bill Grimmer, Denelle Hanlon, Deirdre Hoffman, and Marie McFarland.

This special report is the property of The Data Warehousing Institute (TDWI) and is made available to a restricted number of clients only upon these terms and conditions. TDWI reserves all rights herein. Reproduction or disclosure in whole or in part to parties other than the TDWI client, who is the original subscriber to this report, is permitted only with the written permission and express consent of TDWI. This report shall be treated at all times as a confidential and proprietary document for internal use only. The information contained in the report is believed to be reliable but cannot be guaranteed to be correct or complete.

For more information about this report or its sponsors, and to view the archived report Webinar, please visit: www.tdwi.org/edqreport.

©2006 by 101communications LLC. All rights reserved. Printed in the United States. The Data Warehousing Institute is a trademark of 101communications LLC. Other product and company names mentioned herein may be trademarks and/or registered trademarks of their respective companies. The Data Warehousing Institute is a division of 101communications LLC based in Chatsworth, CA.

Executive Summary

Data quality is difficult to comprehend in its entirety, because of the diverse aspirations and actions collected under its broad umbrella. This includes standard technology and business practices that improve data, like name-and-address cleansing, record matching and merging, house-holding, de-duplication, standardization, and appending third-party data. Some of these tasks can be automated with software, while others—like entering data properly—are purely matters of business process.

Given this complexity, it's no wonder misconceptions abound, like thinking data quality is a one-time action that results in perfection. To the contrary, data quality is a complex concept that encompasses many data-management techniques and business-quality practices, applied repeatedly over time as the state of quality evolves, to achieve levels of quality that vary per data type and seldom aspire to perfection.

Of the organizations TDWI surveyed, 82.5% continue to perceive their data as good or okay. However, half of the practitioners surveyed warn that data quality is worse than their organization realizes, which explains why the number of organizations with a data-quality plan doubled between 2001 and 2005. Many companies took action on data quality because compliance provided a swift kick in the pants. Other kicks came from initiatives for business intelligence, customer service, global supply chain, and IT system consolidations and migrations.

Two-thirds of respondents have studied the problems of data quality, while less than half have studied its benefits. This indicates clearly that data quality initiatives are driven more by liability than leverage. In other words, organizations improve their data to avoid problems like direct-mail costs, misguided decisions, poor customer service, or faulty information in financial and regulatory reports. Of course, when these problems are fixed, data has greater leveragability. The benefits aren't completely overlooked, since most organizations surveyed claim a return on investments in data quality. Either way you look at it, the liabilities of poor-quality data and the leveragability of high-quality data should compel anyone to action.

Data-quality products and practices are evolving quickly as they move from technical to business users, from point products to suites, from batch to real-time operation, from data profiling to quality monitoring, from US-centric to global, and so on. All these trends boil down to the fact that data quality is broadening beyond its departmental roots into enterprise-scope usage. While this broadening is good for the data, it's challenging for the organization, which must adjust its business processes and IT org chart to adapt to enterprise usage.

Accomplishing anything with this kind of *enterprise data quality* (EDQ) requires close collaboration among IT and business professionals, who understand the data and its business purpose—collaboration made manifest in a data-governance committee or program. Data governance is rare today, but will proliferate as companies take data quality into broader enterprise use and move beyond mere stewardship. TDWI recommends data governance strongly, because it gives all data-management practices consistency, efficiency, and mandate as they reach for enterprise scale.

Note that the most critical success factor for EDQ via data governance is *mandate*. Data stewards and governors must induce technical and business managers beyond their purview to change their processes and data when opportunities for data improvement arise. Without a strong mandate (supported by an attentive executive sponsor) to drive pragmatic changes, EDQ, data governance, and data stewardship deteriorate into an academic study of data.

Data quality is a complex concept that encompasses many technologies and business practices.

Most business people think their data quality is good, while half of technical people say it's worse than perceived.

Data quality is driven more by liability than leverage.

Data quality is broadening beyond its departmental roots into enterprise-scope usage.

Data governance gives enterprise data quality the organizational structure it needs for collaboration between IT and the business.

Research Methodology

Report Scope. This report is designed for technical executives who wish to understand the state of data-quality initiatives today, as well as why they are evolving from departmental solutions to enterprise programs. The report describes a range of best practices for improving the quality of data, the drivers behind trends in data-quality practices, and how data governance is required for successful data-quality initiatives on an enterprise scale.

Terminology. TDWI defines *data quality* as the quality of data's content and structure (according to varying criteria), plus the standard technology and business practices that improve data, like name-and-address cleansing, matching, house-holding, de-duplication, standardization, and appending third-party data.

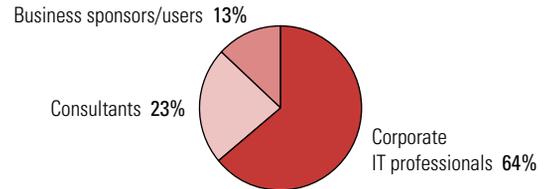
Survey Methodology. This report's findings are based on a survey run in late 2005 (with occasional references to a survey run in late 2001), as well as interviews with data-management practitioners, consultants, and software vendors. In 2005, TDWI sent an invitation via e-mail to the data-management professionals in its database, asking them to complete an Internet-based survey. The invitation also appeared on several Web sites and newsletters; 803 people completed all of the survey's questions. From these, we excluded the 53 respondents who identified themselves as academics or vendor employees, leaving the completed surveys of 750 respondents as the data sample for this report.

Survey Demographics. The majority of survey respondents (64%) are corporate IT professionals, whereas the remainder consists of consultants (23%) or business sponsors/users (13%). Hence, this market sample ably represents data-management practitioners and their business counterparts. Due to branching in the survey, some questions allow responses only from individuals who've had direct experience with data-quality or data-governance initiatives.

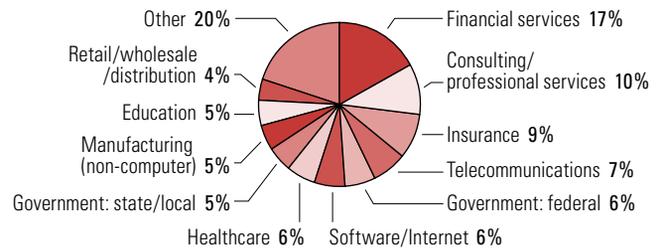
The financial services and insurance industries (26% combined) dominate the respondent population, followed by IT consultants (10%) and other industries (single-digit percentages). We asked consultants to fill out the survey with a recent client in mind. By far, most respondents reside in the U.S. (62%), trailed by Europe and Canada, respectively. Respondents are evenly distributed across all sizes of companies.

Demographics

Position

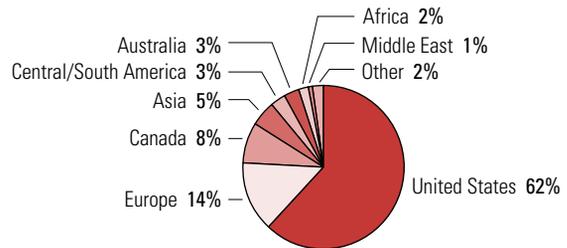


Industry

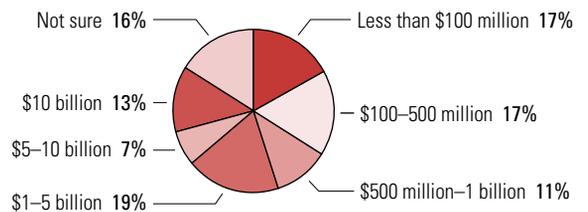


(The "other" category includes industries with less than 4% of respondents.)

Geography



Company Size by Revenue



Based on 750 qualified respondents.

The State of Data Quality

Defining Data Quality

Business and technical people apply the term *data quality* in many different ways. Yet, we all seem to grasp the diversity of aspirations and actions collected under its broad umbrella. Although no single definition is possible, we can identify three common focal points:

- **The quality of data.** The criteria for measuring quality vary according to the type of data, its use, business requirements, what's possible technologically, and how tolerant the technology and the business are to defective, incomplete, or nonstandard data. Quality assessments and adjustments are applied to data collections that range from individual fields and records to multiple databases spanning multiple enterprises.
- **Business-quality processes.** These are a means of assuring the quality of data as people create and handle it via applications. Processes range from localized standards for data entry to mid-level processes where people review and correct data by hand to global, enterprise-scope programs for data governance and stewardship.
- **Data-management techniques.** These are usually implemented through homegrown or vendor-built software solutions, the point being that software provides automation for measuring quality and adjusting data. These techniques automate many best practices, including name-and-address cleansing, record matching and merging, house-holding, de-duplication, standardization, and data enhancement (e.g., where geospatial data or demographic and firmagraphic data from a third party are appended to enterprise records).

With such a complex definition, it's no wonder that data quality suffers myths and misconceptions that merit correction here:

- **More than defect correction.** The uninitiated often jump to the conclusion that data quality is about finding and correcting defects. But this is just one of data quality's tasks, whereas others include standardizing (so multiple systems speak the same language) or repurposing data (to create data structures conducive to data analysis or reporting). Tasks like matching, house-holding, and de-duplication find records that are not wrong per se; they're just redundant, and so must be merged or deleted, usually to reduce the costs of direct mail or telemarketing.
- **Not a one-time action.** Being a data steward is like being a maid: as soon as you clean up a room, someone messes it up. For example, customer data degrades as your customers move, marry, and matriculate. To stay ahead of the degradation, periodic tasks monitor the quality of customer data, then manipulate it to maintain an appropriate state. Even so, data quality might be a one-time action in rare cases like database migrations or consolidations.
- **Seldom about perfection.** In some cases, it makes sense to reach for perfection in data quality, due to the dire ramifications of poor quality (as with financial or regulatory reporting). However, maintaining "just enough quality" is a more efficient use of resources and a more realistic goal, even though it demands case-by-case definitions. For example, data about supplies that's only seen and used internally has a high tolerance for low quality, whereas data seen by the customer needs a state of quality closer to perfection.

Data quality is a complex concept that encompasses many technologies and business practices.

Due to its complexity, data quality suffers many myths and misconceptions.

Summary. Data quality is a complex concept that encompasses many data-management techniques and business-quality practices, applied repeatedly over time as the state of quality evolves, to achieve levels of quality that vary per data type and seldom aspire to perfection.

Best Practices for Improving the Quality of Data

Many of the practices applied to the early phases of a data-quality project involve the discovery, inventory, and profiling of data assets. Among these, data profiling is crucial to success.

Data profiling—a practice on its own—is integral to successful data-quality efforts.

- **Every data-quality project must start with data profiling.** If you're in IT, your ultimate deliverable is improved data, not a profile of current data. So, it's hard to rationalize committing time and resources to data profiling prior to designing a data-quality solution. Avoid the urge to scrimp on data profiling, because it leads to more accurate scope, better defined business benefits, and a reduction of “gotchas” after deployment. Luckily, most users understand the importance of data profiling, as seen in their use of manual analysis and automated analysis (86% and 24%, respectively, in Figure 1).
- **Tool-based data profiling is more productive than manual methods.** There are two common routes to data profiling: mostly manual methods and vendor tools. Manual methods involve using a query tool or hand-coded queries; the user typically writes a document summarizing the data's structure and the state of its quality. The other route involves vendor tools dedicated to data profiling or functions within a larger tool that do “just enough profiling.” These features are now common in most data-integration and data-quality tools.

Today, manual methods are more common than tool-based ones. But the former are less productive and accurate than the latter, because of the time-consuming and error-prone process of moving profile information from a query to the document to the data-quality solution. Users must repeat this process many times, so they should rely on a dedicated profiling tool or the profiling functions of their data-quality or integration tools.

- **Data profiling, like data correction, is not a one-time task.** Over a project lifecycle, data profiling evolves from an ad hoc study of data in design and development phases to monitoring and measuring the state of quality (38% in Figure 2), thereby supporting data audits in deployment and maintenance phases (25%). Software automation enables companies to repeat these activities monthly, weekly, or daily, whereas manual methods aren't this repeatable.
- **Data profiling, integration, and quality are closely related data-management practices.** Data integration—regardless of its form as ETL, EII, or replication—involves accessing data in a source system, merging and transforming it, and moving it into a target system (which may be the same as the source). Data-quality operations that can be automated by software demand a fair amount of data integration. Like data quality, data integration demands a lot of data profiling. And data-integration operations inexorably uncover data-quality problems. Hence, data profiling, integration, and quality are so intrinsically linked that many companies practice them in tandem.

How does your organization determine the quality of data prior to a new IT project?

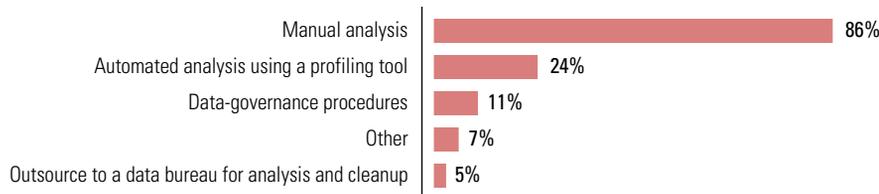


Figure 1. Based on 1,004 responses from 750 respondents in 2005.

How does your organization determine the quality of data after the IT project is deployed?

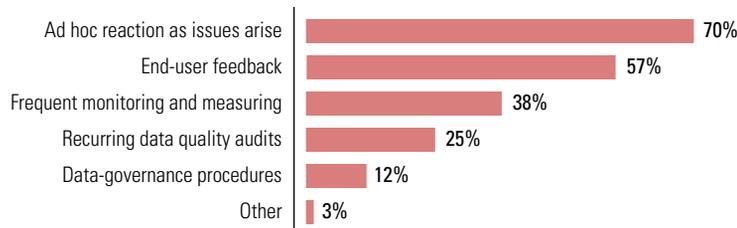


Figure 2. Based on 1,543 responses from 750 respondents in 2005.

How does your company typically fix data-quality problems when they arise?

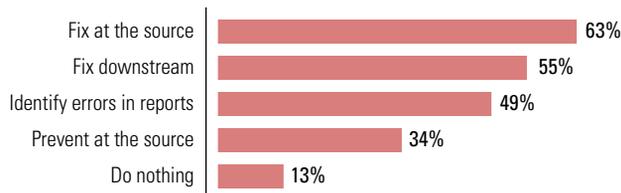


Figure 3. Based on 1,611 responses from 750 respondents in 2005.

Data profiling aside, data quality in most companies today is seldom outsourced, seldom governed on an enterprise scale (but will be eventually), and seldom proactive instead of reactive.

Data quality is seldom outsourced, governed, or proactive.

- **Data quality is largely an internal affair, seldom outsourced.** Numerous service bureaus specialize in various types of data processing, but most focus on name-and-address cleansing for consumers at residential addresses (or e-mail addresses), or sometimes businesses at commercial addresses. This is valuable, but it's a fraction of the corporate data that needs improvement, which explains why a mere 5% of survey respondents reported using a bureau.
- **Data governance is new but upcoming.** At 11% and 12% in Figures 1 and 2, data-governance procedures guide very few data-quality initiatives today. However, all the users TDWI interviewed for this report have recently deployed a data-governance program or are planning one, suggesting that this is a growth area among data-management practices.

- **Data assessments and improvements should be both reactive and proactive.** In too many cases, finding and fixing a data-quality problem is an ad hoc reaction to an issue or end-user feedback (70% and 57%, respectively, in Figure 2), instead of proactive measures like frequent monitoring and recurring data-quality audits (38% and 25%). In other words, unanticipated problems flare up and repeatedly push IT and the business into disruptive fire drills. To a degree, this is inevitable, given the sad state of data and the lack of governance in most organizations. And each organization may have unique definitions of what's reactive versus proactive. Even so, proactive order can reduce reactive chaos.

USER STORY

Data quality in practice needs both reactive and proactive procedures.

Improving data when it's created or changed is easier than finding it later, then assessing and altering it on the fly.

“Many organizations struggle to transition from a reactive process—which addresses data crises as they occur—to a mature, proactive one that measures conformance with data quality expectations early in the information flow, catching information flaws before they can negatively impact the business,” said David Loshin, the president of consulting firm Knowledge Integrity, Inc. “When developing a short-term reactive fix, users should take the time to understand the root cause of the problem and develop business rules that can be folded into long-term proactive procedures. Over time the proactive procedures get richer, while the short-term reactive fixes—which are inevitable—get less hectic. This transition is key to project goals like monitoring quality, assuring conformance with established business rules, and providing auditability.”

Upstream versus downstream. With a data lifecycle, assuring quality sooner is better than later.

- **Data quality, historically applied downstream, is moving upstream.** When asked where they first applied data-quality practices, most users interviewed by TDWI identified one of two areas: a data warehouse or a marketing database used for direct mail. Companies starting at one of these points have struggled to move their data-quality best practices upstream to improve application data. The fact that 63% of respondents claim they “fix at the source” means that data-quality practices—like salmon returning to where they were spawned—have succeeded in swimming upstream (see Figure 3). This is a worthy goal, since improving data early in its flow improves the many business and technical processes that touch it later. Even so, there's still a need to “fix downstream” (55%) and “identify errors in reports” and other products of downstream databases (49%).
- **An ounce of prevention is better than a pound of cure.** Many actions can degrade the quality of data, but most occur at or near the source. Hence, some best practices operate on data before it's allowed into a system, under the assumption that it's easier to correct a defect or enhance an incomplete record before it's mixed in with thousands or millions of other records. Figure 3 indicates that 34% of organizations “prevent at the source” by validating, cleansing, standardizing, and enhancing records that originate from human data entry or flat files before applying each record to an application database.
- **Garbage in, garbage out.** A common problem mentioned by interviewees is that business users make arbitrary decisions when using applications, which makes some database fields useless. The solution can be as simple as giving users training or more time for data entry, solutions that don't require a data-quality tool, a line of code, or any action from IT.

USER STORY

High-quality data entry leads to high-quality application data.

“When a call center rep closes a customer interaction, the rep classifies it by selecting a wrap-up code,” said Mannie Goldberg, director of data resource management at NSTAR. “Our reports showed that most interactions were tagged as general inquires or unknown, which didn't tell us anything useful. We showed this to the call center director and that led to data-entry changes and

additional rep training. The changes have given our call center management new insights into why customers make repeated calls. It’s amazing how a simple fix like this can improve both external customer interactions and internal business intelligence.”

Evolving Perceptions, Realities, and Initiatives

Perceptions of the quality of data haven’t changed much since TDWI’s 2001 survey. Figure 4 reveals that all ratings changed four percentage points or less between the two surveys. In both years, survey respondents painted a rosy picture, with a combined average of 82.5% perceiving their data as good or okay and only 17.5% as excellent or poor.

82.5% continue to perceive their data as good or okay.

As with many things in life, a grim reality balances the rosy perception (see Figure 5). Nearly half of survey respondents claim that the quality of their organization’s data is “worse than everyone thinks.” The wide majority of respondents are data-management practitioners who know data well and understand the ramifications of quality. Their ironic comparison of perception and reality warns us that most organizations need to take action to improve the quality of data.

Half of practitioners warn that data quality is worse than their organization realizes.

Our organization thinks the quality of its data is:

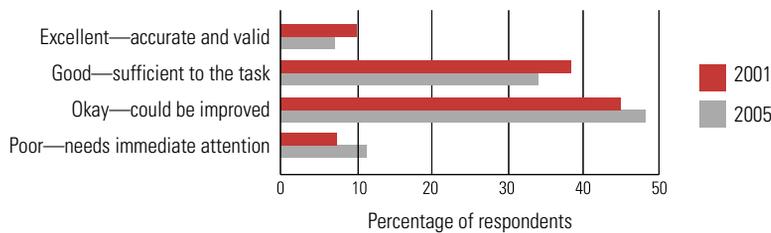


Figure 4

In reality, the quality of our data is:

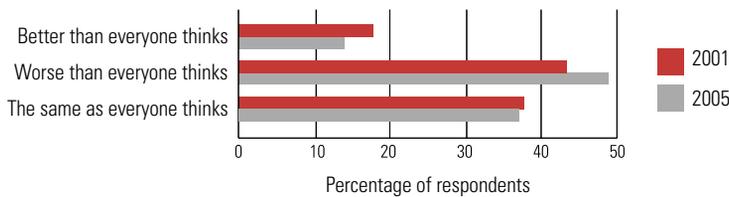


Figure 5

What is the status of your data-quality initiative?

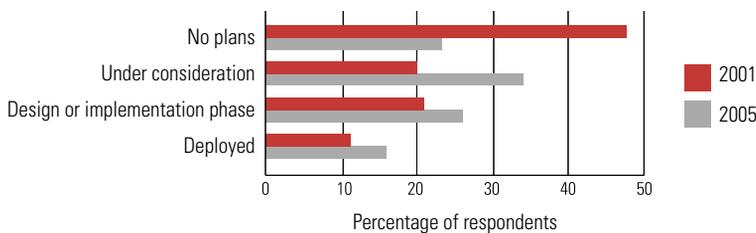


Figure 6. All three figures are based on 647 respondents in 2001; 750 in 2005.

Organizations with a data-quality plan doubled between 2001 and 2005.

Despite its problems, data quality has witnessed substantial action in recent years. A frightening 48% of respondents reported having “no plans” for a data-quality initiative in 2001 (see Figure 6). By 2005, the percentage of respondents without a plan had dropped dramatically to 24%, while responses increased for data-quality initiatives considered, designed, or deployed. This dramatic change reveals that a significant number of corporations took action on data quality between late 2001 and late 2005. When asked in a TDWI Technology Survey in May 2004 (97 respondents), this question yielded percentages similar to those of 2005, suggesting that the increase in data-quality usage was well under way by early 2004.

Compliance is a swift kick in the pants—but not the only kick.

The most obvious change in the business climate during this period is, of course, regulatory reporting and compliance. This affects both corporate oversight and homeland security, as mandated by US federal legislation like the Sarbanes-Oxley Act, the TREAD Act, and the USA PATRIOT Act. Most users interviewed in the course of this research pointed to regulatory reporting and other compliance tasks as events that changed the way they and their employers handle data. No doubt, compliance has driven some companies to take action recently, but there are many other good reasons for doing something about the quality of your data.

USER STORY

Compliance can be a Trojan horse that gives enterprise data quality its mandate for change.

Getting buy-in to the need to implement information quality management (IQM) disciplines with a clear mandate means you need to find an important goal at the executive level that is impacted by non-quality, non-managed information. “Compliance is a foot in the door for us, in terms of getting a mandate to change upstream application data,” says an IQM thought leader in a leading European telecommunications company. “Once tackling IQ for the sake of compliance is proven and our methodology has delivered successfully in support of the important goal of compliance, we anticipate broadening our mandate to other systems where the pain of non-quality information is just as important but perhaps not as urgent.” Hence, compliance isn’t necessarily a cost center when you leverage your accomplishments in compliance by applying them to other initiatives to address other business goals.

Liability and Leverage—A Case for Data Quality

When making a case for a data-quality initiative or project, organizations cite both liability and leverage. They need to reduce costs by alleviating the liabilities of poor-quality data or they want to increase revenue by leveraging the benefits of high-quality data. Either way, the case can be compelling, such that most organizations claim a return on investments in data quality.

The Problems of Poor-Quality Data

Roughly half of organizations suffer tangible problems due to poor-quality data.

In the surveys of 2001 and 2005, TDWI asked: “Has your company suffered losses, problems, or costs due to poor quality data?” (See Figure 7.) Numbers of respondents answering “yes” grew from 44% in 2001 to 53% in 2005, which suggests that data-quality problems are getting worse.

In the same period, however, respondents admitting that they “haven’t studied the issue” dropped from 43% to 36%. It’s possible that the two trends cancel each other out, such that problems haven’t necessarily increased. Rather, more organizations now know from their own study that data-quality problems are real and quantifiable. Averaging the two years together, 48.5% (or roughly half) of organizations now recognize the problem. Since this is far higher than the 12% denying any

problem, we conclude that problems due to poor-quality data are tangible across all industries and exist in quantity and severity sufficient to merit corrective attention.

Poor-quality data creates problems on both sides of the fence between IT and business (see the 2005 data in Figure 8). Some problems are mostly technical in nature, like extra time required for reconciling data (85%) or delays in deploying new systems (52%). Other problems are closer to business issues, like customer dissatisfaction (69%), compliance problems (39%), and revenue loss (35%). Poor-quality data can even cause problems with costs (67%) and credibility (77%) on both sides of the fence.

Problems due to poor-quality data impact both IT and the business.

The Origins of Poor-Quality Data. Survey responses show that problems unquestionably exist. But where exactly do they come from?

- Problems originate in both IT and the business.** (See Figure 9.) Problems arise from technical issues (conversion projects, 46%; system errors, 25%), business processes (employee data entry, 75%; user expectations, 40%), and a mix of both (inconsistent terms, 75%). Problems even come from outside (customer data entry, 26%; external data, 38%). Hence, data quality is assaulted from all quarters, requiring great diligence from both IT and the business to keep its problems at bay, with both internal processes and external interactions.

Data quality is assaulted from all quarters, requiring great diligence to keep problems at bay.

Has your company suffered losses, problems, or costs due to poor-quality data?

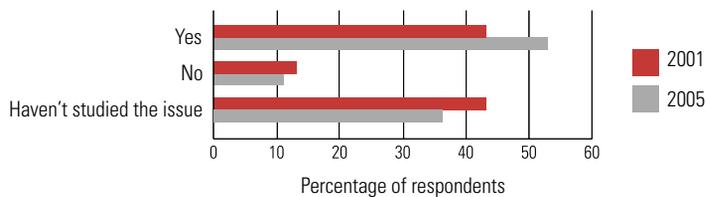


Figure 7. Based on 647 respondents in 2001; 750 in 2005.

Which problems has your company suffered from due to poor-quality data?

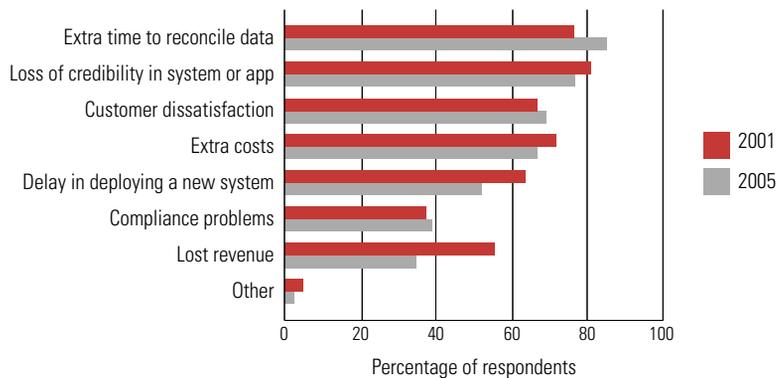


Figure 8. Based on 1,340 responses from 286 respondents in 2001; 1,703 responses from 399 respondents in 2005.

- **Inconsistent data definition is a leading origin of data-quality problems.** Too often, the data itself isn't wrong; it's just used wrongly. For example, multiple systems may each have a unique way of representing a customer. Application developers, integration specialists, and knowledge workers regularly struggle to learn which representation is best for a given use. When good data is referenced wrongly, it can mislead business processes and corrupt databases downstream. With 75% of survey respondents pointing to this problem, it ties with data entry as the most common origin of data-quality problems.
- **Data entry ties for "worst place" as an origin of data-quality problems.** This problem has been with us since the dawn of computing and will probably be there at its sunset. At least the problem is lessened by user interfaces that require as little typing as possible, validation and cleansing prior to committing entered data, training for users, regular data audits, and incentives for users to get it right.

USER STORY

Data entry via mobile computing results in low-quality data.

As if the data-entry problem weren't severe enough, multiple IT directors interviewed for this research lamented that mobile computing lowers the quality of data entered. This is because a handheld device has low usability and a notebook PC is hard to use in the field. To compensate, these organizations cleanse mobile data carefully before synchronizing it with enterprise systems.

Which of the following most often contribute to data-quality problems in your organization?

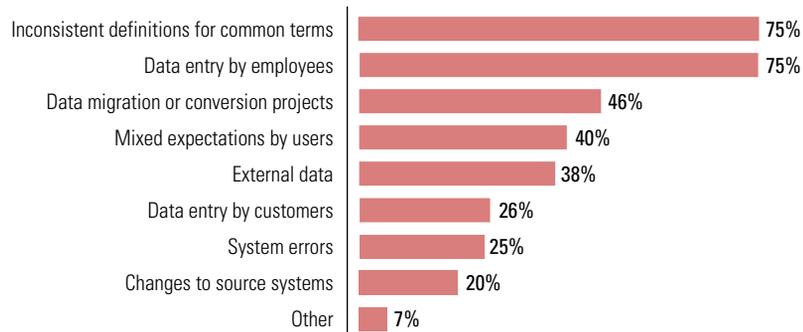


Figure 9. Based on 1,522 responses from 399 respondents in 2005.

Which types of data are especially susceptible to quality problems in your organization?

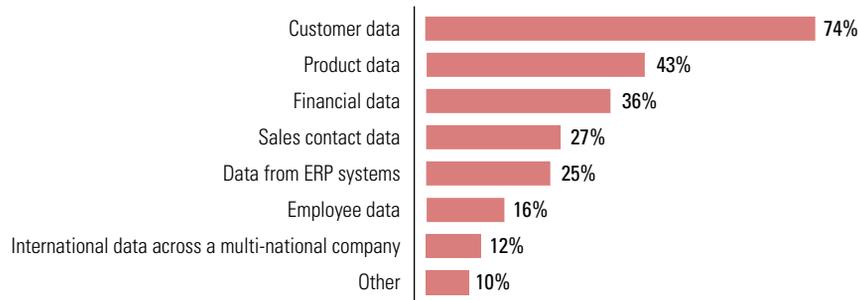


Figure 10. Based on 966 responses from 399 respondents in 2005.

Data representing certain business entities (like customer and product) are more prone to data-quality problems than data about other entities (like finances or employees). (See Figure 10.)

- **Data about customers is the leading offender (74%).** That's because customer data changes constantly as customers run up bills, pay bills, move to new addresses, change their names, get new phone numbers, change jobs, get raises, have children, and so on. The customer is the most highly changeable entity in most organizations, along with equivalents like the patient in healthcare, the citizen in government, and the prospect in sales force automation. Unfortunately, every change is an opportunity for data to be entered incorrectly or to become out of date. Since customer data is often strewn across multiple systems, synchronizing it and resolving conflicting values are common data-quality tasks, too.
- **Product data (43%) is in a distant second place after customer data.** Defining "product" is challenging, since it can take different forms, including supplies that a manufacturer procures to assemble a larger product, the larger product produced by the manufacturer, such products traveling through distribution channels, and such products available through a wholesaler or retailer. Note that this list constitutes a supply chain. In other organizations, the chain is not apparent; they simply acquire office supplies, medical supplies, military munitions, and so on, which are consumed in the production of a service. Hence, one of the greatest challenges to assuring the quality of product data is to first define what "product" means in an organization.

The procurement office of a Canadian provincial government follows one of the tried-and-true practices of procurement: diversify by acquiring critical supplies from multiple suppliers. This allows them to change suppliers when prices change or one supplier cannot deliver. The catch is that each supplier has a unique product identification number and description. This user solves the problem with a data-quality tool with probabilistic matching. They fuzzy-match supplies that are equivalent commodities, so they know their level of spending and inventory per commodity (not just per supplier and unique supply), which enables smarter budget and inventory management.

The Benefits of High-Quality Data

Roughly half of respondents reported they "haven't studied the issue" of data-quality benefits (49% in Figure 11), whereas only a third haven't studied its problems (36% in Figure 7). With more time spent studying problems instead of benefits, data quality is clearly driven more by liability than leverage. Even so, benefits exist, and 41% claim to have derived them, compared to a mere 10% denying any benefit (see Figure 11).

The top three benefits of high-quality data identified by respondents all relate directly to data warehousing (see Figure 12), namely greater confidence in analytic systems (76%), less time spent reconciling data (70%), and single version of the truth (69%). This is to be expected, since data-quality projects have a track record of success in data warehousing. Other benefits are more business-driven, like gains in customer satisfaction (57%), cost reduction (56%), and extra revenues (30%).

Data about customers and products creates the most problems, and so gets the most attention.

USER STORY

Matching equivalent commodities enables smarter budget and inventory management across suppliers.

Data quality is clearly driven more by liability than leverage.

Has your company derived benefits or gains from high-quality data?

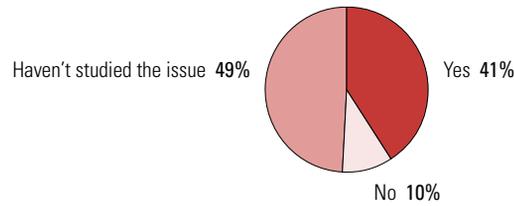


Figure 11. Based on 750 respondents in 2005.

Which benefits has your company derived from high-quality data?

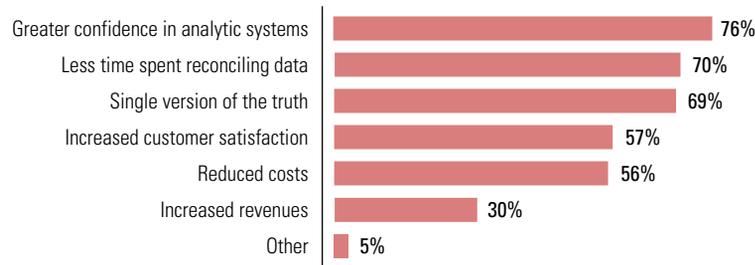


Figure 12. Based on 750 respondents in 2005.

USER STORY

Rich, standard product data greatly increases outbound supply chain efficiencies.

Large consumer packaged good (CPG) companies all participate in a global sync network, enabling data about their products to flow to their suppliers, distributors, and retailer partners. The network enables very efficient handling and delivery, such that a CPG company can get products onto retailers’ shelves much faster than was possible without the network. “But all this assumes very high-quality product data,” said a director in the supply chain organization of a CPG firm. “If data is inconsistent or incomplete—say, it doesn’t follow the standard or a value for product weight is missing—then the data won’t flow on the network, which means products won’t ship.” To leverage the benefits of this highly useful network, member companies have all invested considerably in standardizing and enhancing product data.

Data-Quality ROI and Budget

Investments in data quality are safe, growing, and likely to yield a return.

TDWI’s 2005 survey asked: “Does your company believe it can achieve a positive return on investment (ROI) by investing in a data-quality initiative?” In Figure 13, 43% of respondents reported that their organization believes ROI is possible, whereas 19% don’t; 38% are honest enough to admit they don’t know. This is similar to the response given when TDWI asked this question in 2001—40%, 19%, and 41%, respectively. Based on the respondents’ appraisal, ROI is a distinct possibility with data quality, though not an overwhelming probability.

Consistent with the recognized possibility of data-quality ROI, a combined 80% of respondents report that data-quality budgets will “stay the same” or increase, versus a miniscule 4% anticipating a budget cut (see Figure 14). Some interviewees described their data-quality initiative or team as a “cost center,” though it is in transition toward becoming a “revenue center.” Given users’ growing budgets and belief that ROI is possible, investments in data quality are safe, growing, and likely to yield a return in a reasonable amount of time.

Does your company believe it can achieve a positive return on investment (ROI) by investing in a data-quality initiative?

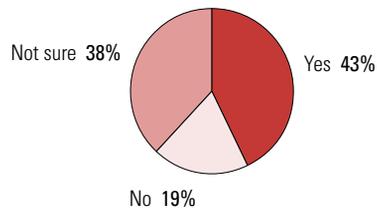


Figure 13. Based on 750 respondents in 2005.

How will the budget for your data-quality initiative change for the next budget period?

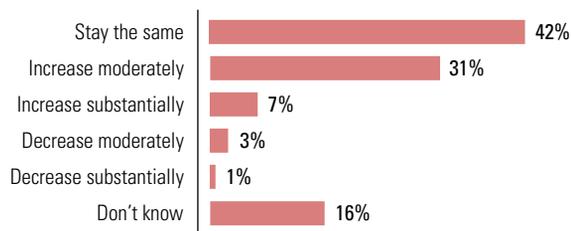


Figure 14. Based on 750 respondents in 2005.

Summary. The liabilities of poor-quality data and the leveragability of high-quality data should compel anyone to action. Organizations that depend on their data can't afford to ignore its quality. Furthermore, data-quality efforts are likely to yield a demonstrable return, and your peers in other organizations are increasing investments accordingly. Bottom line: you should, too.

Taking Data Quality to the Enterprise

The Scope of Data-Quality Initiatives

Historically applied to isolated silos in departments or single databases, data quality is progressively applied with more breadth across enterprises. TDWI research shows that a surprisingly large number of organizations (39%) are already applying data quality in some form to the “whole enterprise” (see Figure 15). An additional 26% apply it in a “single department that spans the enterprise,” such as the IT and marketing departments. Meanwhile, low percentages continue the older tradition of applying data quality mostly in “a single department” (14%) or “a single business unit” (13%).

The gist of the market data is that many organizations are well down the road to enterprise-scope use of data-quality techniques and practices—enterprise data quality (EDQ). Users interviewed for this research reported similar progress, corroborating the survey data. But most interviewees quickly added that they had only recently arrived at EDQ, typically in 2003 or 2004. Hence, the data-quality marketplace and user community has only recently crossed the line into EDQ. TDWI

Many organizations are well down the road to enterprise-scope use of data quality.

suspects that many more organizations are on the cusp, and will cross into EDQ in coming years. Of course, occasional departmental usage will continue alongside EDQ.

What is the scope of your data-quality initiative?



Figure 15. Based on 569 respondents in 2005 (only those who have a data-quality initiative).

Which business initiatives does your data-quality initiative support?



Figure 16. Based on 1,153 responses from 569 respondents in 2005.

Which software solutions does your data-quality initiative support?

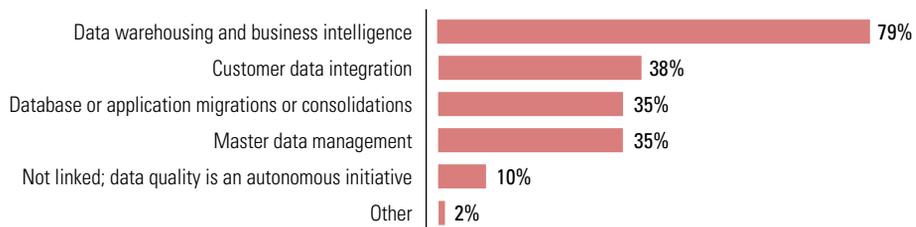


Figure 17. Based on 1,137 responses from 569 respondents in 2005.

Data quality is often piggy-backed atop related initiatives, which carry it across the enterprise.

One of the reasons data-quality usage is spreading is that it is often piggy-backed atop related initiatives that carry it across the enterprise. Just about any data-intense initiative or software solution will ferret out data-quality problems and opportunities. IT and business sponsors have realized this over time, so it's become commonplace to include a data-quality component in initiatives for governance (49%), CRM (42%), marketing campaigns (34%), compliance exercises (35%), and supply chain management (16%). (See Figure 16.) Anecdotal evidence suggests that data quality gives these initiatives better planning, a more predictable schedule, and a higher-quality

deliverable. Likewise, data-quality software is often integrated with software for other solutions, like data warehousing and business intelligence (79%), customer data integration (38%), migrations and consolidations (35%), and master data management (35%)¹. (See Figure 17.)

“Our recent data warehouse initiative proves that clean, integrated data and good decisions go together,” says Darren Taylor, vice president of the Information Access Division at Blue Cross Blue Shield Kansas City (BCBSKC). “Our longer-term vision is to leverage data to help customers make sound healthcare decisions. This vision demands a high degree of data quality.” BCBSKC is applying its expertise with data warehousing; future plans include utilizing portal-based solutions to provide important healthcare quality and efficiency metrics in an effort to improve the health of the community. “When there’s an issue with a published metric, the first thing people do is challenge the data,” Taylor continued. “So, the data you deliver outside your organization has to be rock solid.”

USER STORY

Standards for external data may be higher than those for internal data.

Most Data-Quality Trends Lead to Enterprise Use

The practice of data quality is in a state of transition, as its every aspect is currently evolving due to a strong trend. Most of these trends have a general effect in common—they result from or cause a broadened use of data-quality tools and practices across an enterprise (see Figure 18).

Data-quality products and practices are evolving quickly...

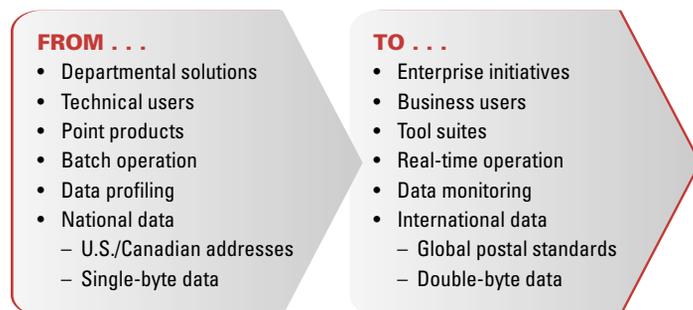


Figure 18. The evolution of data-quality products and practices.

From departmental solutions to enterprise initiatives. Most IT directors interviewed for this report spoke of their first data-quality projects as supporting data warehousing or marketing functions like direct mail. A few mentioned product data issues, like catalog matching and cleansing. Regardless of the isolated silos where they started, data-management professionals now pursue the quality of data broadly across their organizations. But many are at a fork in the road: either they keep deploying data-quality silos in more departments, or they fall back and regroup into a centralized team that gains efficiency and consistency across all efforts. TDWI recommends the second route, because it establishes a structure that leads to further progress in the long run.

From technical users to business users. Long story short, the data-quality user community gets more diverse all the time. At one extreme, technical users design and code large, scalable solutions for matching and consolidation rules. At the other extreme, semi-technical marketers handle name-and-address cleansing and other customer data issues. With the rise of stewardship years ago and

¹ Data-quality software supports these initiatives; yet poor-quality data can act as a barrier to them. As Colin White points out, “data quality issues are the leading inhibitor to successful data integration projects.” See the TDWI report *Data Integration: Using ETL, EAI, and EII Tools to Create an Integrated Enterprise*, November 2005, www.tdwi.org/research.

data governance recently, there's a need to support business users who are process and domain experts, with little or no technical background. This trend affects the tools that vendors provide; most were designed for one of these user constituencies, and now must support them all. Likewise, organizations must staff their data-quality initiatives carefully to address all these user types, their needs, and their unique contributions.

From point products to tool suites. TDWI defines data quality as a collection of many practices, which explains why few vendors offer a single “data-quality tool.” Instead, most offer multiple tools, each automating a specific data-quality task like list scrubbing, fuzzy matching, geo-coding, and so on. As organizations broaden data-quality usage, they use more of these point products, then suffer the lack of interoperability, collaboration, and reuse among them. To supply this demand, vendors have worked hard to integrate their point products into cohesive suites.

From batch to real-time operation. Some data-quality tasks are still best done in batch, like list scrubbing and matching records in large datasets. However, given that data entry is the leading source of garbage data, there's a real need for real-time verification, cleansing, and enhancement of data before it enters an application database. As with most data-management practices, data-quality software has evolved to support various speeds of “right-time” processing.

From data profiling to data monitoring. These are similar in that each results in an assessment of the quality of data. On the one hand, data profiling has the additional step of data discovery and is done deepest prior to designing a data-quality solution. Monitoring, on the other hand, is about measuring the quality of data frequently while a data-quality solution progresses, so stewards and others can make tactical adjustments to keep a quality initiative on plan. The trend is to do profiling more deeply, then embrace monitoring eventually. TDWI strongly recommends both.

From national to international data. Companies that are multi-national or have a multi-national customer base have special problems when expanding data-quality efforts across a global enterprise. Name-and-address cleansing is a straightforward task when done for U.S. and Canadian addresses and postal standards; yet it becomes quite complex as you add more languages, national postal standards, and information structures (like Unicode pages and double-byte data). Users must deal with these and other issues as they take data quality to an international enterprise.

USER STORY

Name-and-address cleansing (usually departmental) also works in enterprise-scope usage, even on a global scale.

TDWI recently interviewed Kelly Williams, the corporate data architect at BMC Software. One of his many mandates is to assure and improve the quality of data. Eventually he'll chair a data-governance committee that fulfills this role enterprisewide, and his proposal for data stewardship is currently under review with his management. In the meantime, he's part of a team starting work on some of BMC's name-and-address lists, by scrubbing them and applying postal standards. So far, the team has solutions in place for lists in the U.S. and Canada. Japan is next, due to pressing problems with the quality of its lists. Eventually, the team will deploy data-quality solutions for name-and-address lists—then other datasets—in 25 European and 14 Asian nations.

Summary. Data quality proved itself in its data warehouse and direct mail origins, and has now moved beyond these into enterprise data quality, where it is applied in many departments for many purposes. In fact, most trends result from or cause a broadened use of data quality tools and practices across an enterprise. Plus, data quality is now *de rigueur* as a component of various business initiatives and software solutions. While this broadening is good for the data, it's challenging for the organization, which must adjust its business processes and IT org chart to adapt to enterprise usage.

Data Governance and Enterprise Data Quality

TDWI data shows that many organizations are practicing enterprise data quality in some sense. The catch, however, is that practices from isolated areas (like data warehousing or marketing campaigns) aren't automatically successful on an enterprise scale. Accomplishing anything at the enterprise level requires close cooperation among IT and business professionals who understand the data and its business purpose and have a mandate for change. To achieve this, an organization can establish a data-governance committee according to the following definition:

*When an organization views data as an enterprise asset (transcending the data warehouse and spanning the whole organization), it establishes an executive-level **data-governance** committee that oversees data stewardship across the organization. Depending on the scope of a data-governance initiative, it may guide related initiatives, like data quality, data architecture, data integration, data warehousing, metadata management, master data management, and so on.²*

Distinctions between stewardship and governance are thin in some cases. But TDWI sees data stewardship as a local task that protects and nourishes specific data collections for specific purposes (like a data warehouse for business intelligence or marketing databases for direct mail). Data governance is a larger undertaking that exerts control over multiple business initiatives and technology implementations to unify these through consistent data definitions and gain greater reuse for IT projects and business efforts. The two can work together, in that a data-governance committee can be a management level that coordinates multiple data-stewardship teams. In a few companies, data governance is a subset of an even larger corporate governance initiative.

The most critical success factor with governance is *mandate*. Governance bodies and stewards must exert change on business and technical people—who own the data and its processes—when opportunities for improvement arise. The most effective mandates come from a high-level executive. Without a strong mandate for change and an attentive executive sponsor, stewardship and governance deteriorate into academic data profiling exercises with little or no practical application.

Ergon Energy is a large utility company in Australia that provides service to an area almost three times larger than Texas. It can take all day to drive or fly to a utility pole or substation, so it's critical to have high-quality geo-coded data describing these assets. As if that weren't challenging enough, Ergon Energy decided 18 months ago to consolidate 45 legacy applications into a packaged application for ERP and improve its data and information quality. "We couldn't have achieved this much progress in data quality this quickly without strong sponsorship from the executive management team," said Nigel Hey, the JET data quality project manager at Ergon Energy. "After the team finishes the short-term work consolidating and cleansing application data, it will shift focus onto developing a long-term stewardship and governance strategy for Ergon Energy."

The State of Data-Governance Initiatives

TDWI's data-quality surveys asked: "Who is responsible for data quality in your organization?" In both 2001 and 2005, the data warehouse team and IT bubbled to the top of the list (47% and 43% in Figure 19). This makes sense from a technology viewpoint, in that these are the technical people long involved in data quality. But it gives technology priority over business, whereas the two must collaborate in a stewardship or governance program. Respondents ranked business analysts

Without a strong mandate, governance deteriorates into an academic exercise.

USER STORY

Executive sponsorship is key to quick progress and long-term success.

Sharing data-quality responsibilities through stewardship and governance is still rare.

² Some form of governance is key to keeping BI projects (and others) aligned with business goals. See the section titled "BI Project Governance" in TDWI's report *Smart Companies in the 21st Century*, 2003, www.tdwi.org/research.

and power users in third place (30%), followed by the “cross-functional team from business and IT” (28%), a description that includes both sides, as in our definition of data governance. So, respondents recognize that responsibility for the quality of data must be shared by some kind of cross-functional team. But the fact that “data-quality analysts” and “data stewards” ranked even lower than “front-line workers” (whose data entry is the leading cause of garbage data) indicates that sharing responsibility through governance and stewardship is still rare.

Who is responsible for data quality in your organization?

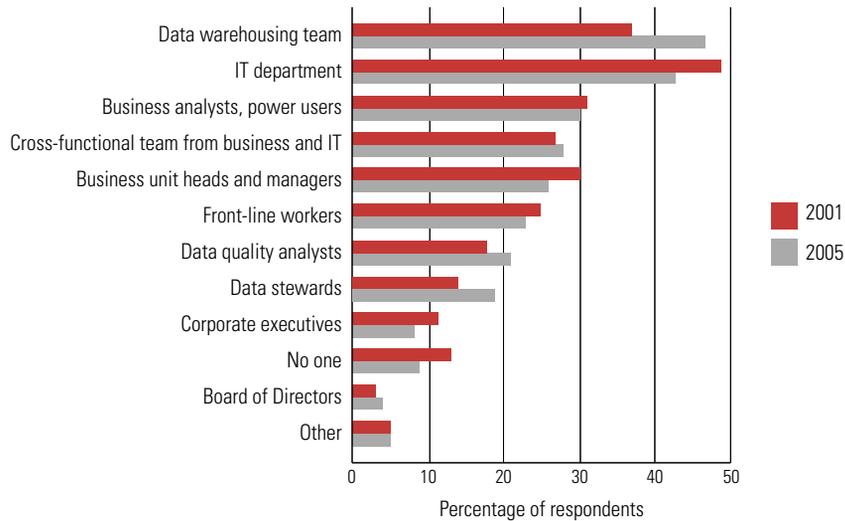


Figure 19. Based on 1,701 responses from 647 respondents in 2001; 1,957 responses from 750 respondents in 2005.

When asked about data-governance initiatives, a disappointing 8% reported having deployed one, while 42% have “no plans” (see Figure 20). TDWI then asked respondents (except those with “no plans”) to rank the effectiveness of the steering committees, degree of executive involvement, and usefulness of policies and processes in their data-governance initiative. The majority ranked all three areas as “moderate,” meaning there’s plenty of room for improvement.

What’s the status of your organization’s data-governance initiative?

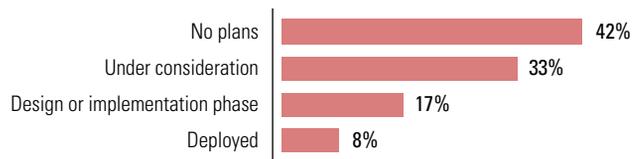


Figure 20. Based on 750 respondents in 2005.

These disappointing responses are most likely due to data governance being a relatively new approach, coupled with the fact that many organizations seem to be on the cusp—they've stretched data-quality practices over the enterprise in a disconnected way and now it's time to control them to ensure consistency and efficiency, whether the control is via stewardship, governance, or centralized IT services. Various forms of data governance will, no doubt, disseminate as more organizations come off the cusp.

“A few years ago, we worked out a high-level business process model guided by Six Sigma methods,” said Kevin McDearis, VP of information enablement at bill-payment outsourcer CheckFree Corporation, the winner of TDWI's 2005 Best Practices Award in Data Quality. “This led us to see that data needs ownership, change process, and continuous improvement, just like everything else in the business,” added Bedeke Cresci, CheckFree's director of information services. From this beginning, CheckFree developed an enterprise-scope data-stewardship program that today manages the ownership and improvement of data in a consensus-driven process, quantified by Six Sigma-style metrics. According to McDearis, “High-quality data and the services built atop it help differentiate CheckFree from its competitors.”

USER STORY

Data needs ownership, change process, and improvement, like everything else.

Anatomy of Data-Governance and Stewardship Programs

Staffing and management hierarchy for data-governance and stewardship programs will vary according to each organization's unique structure and needs. The following description—a composite drawn from multiple interviewees—illustrates the requisite parts:

- **A domain steward is assigned per business unit.** Domain stewards work directly with the line-of-business manager who owns the data and the IT manager who administers it. Each steward has a mandate to change the process and structure of any business, person, or IT system, if that's what it takes to improve data. Note that most changes proposed should be business oriented, with business value as a goal of any data-quality work that gets done. Without demonstrable value, it's unlikely the work will get approved or done.
- **A corporate steward manages a group of domain stewards.** This management hierarchy helps related domain stewards collaborate. And the corporate steward provides domain stewards with additional clout to help domain stewards enforce their mandates.
- **A governance committee consists of miscellaneous managers.** These include corporate stewards, corporate sponsors (CxOs and SVPs), and miscellaneous IT and line-of-business managers, as needed. This committee sets top-down strategic goals, coordinates efforts, and provides common definitions, rules, and standards, which apply to data structures, access, and use across the entire enterprise.

Figure 21 shows how the layers of stewardship and management may roll up into a data-governance committee. Dark arrows represent direct reports, while gray arrows represent significant interactions outside the reporting structure of the organization.

Possible Organizational and Report Structure for Data Governance

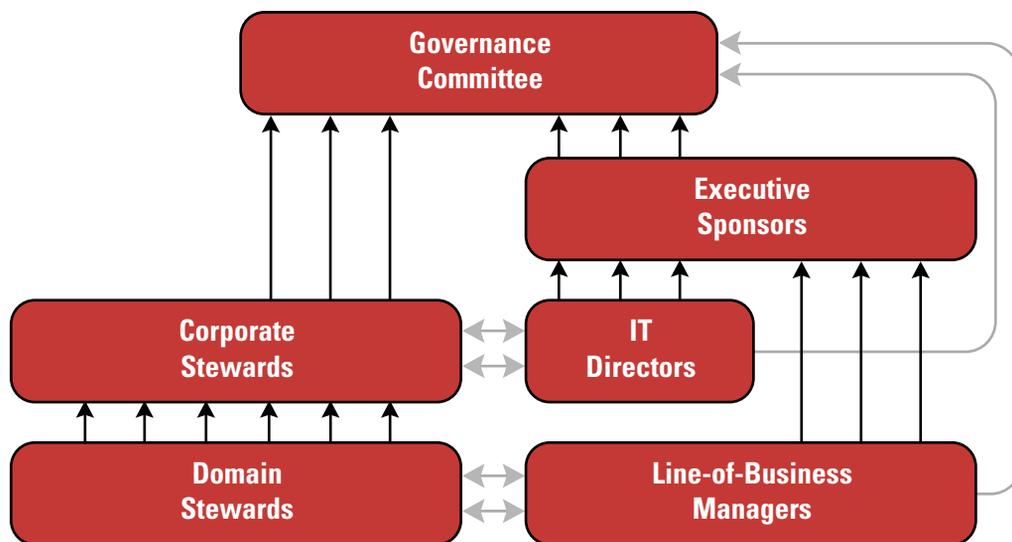


Figure 21. Stewardship and management roll up into data governance.

Summary. Accomplishing anything at the enterprise level requires close cooperation among IT and business professionals, who understand the data and its business purpose and have a mandate for change. To achieve this, an organization can establish a data-governance committee. Such committees are rare today, but will proliferate as companies take data quality into broader enterprise use and move beyond mere stewardship. TDWI recommends governance strongly, because it gives data quality and other data-management practices consistency, efficiency, and mandate as they reach for enterprise scale.

Recommendations

Profile data more deeply and more often. Up front, this yields better planning and fewer gotchas down the road. When possible, use profiling tools (instead of manual documentation) for better productivity. In a deployed data-quality software solution, profiling should be automated as a job that runs repeatedly to monitor the quality of data and ensures continuous improvement.

Find solutions in software tools, plus adjustments to business processes. Look to data-quality tools for invaluable automation of defect correction, matching, standardization, and appending. But also look for improvements in how end users handle data via applications, since this is where most data defects arise.

Establish both proactive and reactive processes for data quality. You need staff to deal with inevitable data-quality problems as they arise. But these chaotic fire drills can be reduced by also having staff and software that proactively seek out problems and opportunities.

Take data quality upstream. For many years, data quality has been part and parcel of managing data in downstream data stores, like data warehouses or direct mail databases. In recent years, it swam upstream to the source of most problems—operational applications. This should be your goal, since adding value upstream improves all the processes that touch data along its flow.

Use compliance like a Trojan horse. Many data owners who've resisted data improvements have reversed their stand recently to allow data quality for compliance. After all, the accuracy and completeness of regulatory reports or subpoenaed datasets is critical. Use this precedence to expand data-quality efforts with applications and databases previously closed to you.

Recognize that any project touching data will reveal data-quality issues and opportunities.

This includes business and technology initiatives like business intelligence, customer relationship management, supply chain management, and data integration. When designing or revising these, include staff and project time devoted to assessing and acting on data quality.

Consider master data management as an emerging data-quality practice. TDWI data shows that inconsistent data definitions are a leading cause of problems. Master data management can be defined different ways, one being a metadata-driven correction for conflicting data definitions, akin to the data standardization goal of many data-quality initiatives.

Embrace the diversity of data-quality practices. Many organizations need to move beyond name-and-address cleansing, data warehouse enhancement, and product catalog record matching. These are useful applications, but are narrow in scope. The lessons learned and skills developed for these can be leveraged in other data-quality applications across the enterprise.

Address enterprise data quality. The data-quality initiatives of 39% of survey respondents already address the “whole enterprise.” Follow their lead into enterprise usage, but resist the urge to deploy data quality in isolated pockets of software tools and IT personnel. Some kind of centralization can improve personnel allocation, project reuse, and data consistency.

Give EDQ required organizational structure through data governance. EDQ's chances of large-scale, long-term success are limited without a support organization, whether its form is a data governance committee, a data stewardship program, or a data quality center of excellence. Another key requirement is a strong mandate supported by a prominent executive sponsor.

Design a successful data-quality solution by including lots of data profiling, a combination of software automation and business adjustment, and both reactive and proactive tasks.

Look for any opening that gets data quality into more systems, more departments, and more initiatives. They all have issues, whether their owners admit it or not.

Get ready for enterprise data quality. It will improve many business and technical processes, if you're open to its diversity and give it necessary organizational structure.



Business Objects
3030 Orchard Parkway
San Jose, CA 95134
408.953.6000
Fax: 408.953.6001
www.businessobjects.com

Business Objects is the world's leading BI software company. Business Objects helps organizations gain better insight into their business, improve decision making, and optimize enterprise performance. The company's business intelligence platform, BusinessObjects™ XI, offers the BI industry's most advanced and complete platform for performance management, reporting, query and analysis, and enterprise information management. BusinessObjects XI includes Crystal Reports®, the industry standard for enterprise reporting. Business Objects has built the industry's strongest and most diverse partner community, and also offers consulting and education services to help customers effectively deploy their business intelligence projects.



DataFlux (a SAS Company)
940 NW Cary Parkway, Suite 201
Cary, NC 27513
919.447.3000 (US)
+44(0) 1753 272 020 (UK)
www.dataflux.com
sales@dataflux.com

DataFlux enables organizations to analyze, improve and control their data through an integrated technology platform. Through its enterprise data quality integration solutions, companies can build a solid information foundation that delivers a unified view of customer, product or supplier data. A wholly owned subsidiary of SAS (www.sas.com), DataFlux helps customers enhance the effectiveness of their data-driven initiatives, including customer data integration (CDI), enterprise resource planning (ERP), legacy data migration and compliance. To learn more about DataFlux, visit www.dataflux.com.



Collaborative Consulting
877.376.9900
info@collaborativeconsulting.com
www.collaborativeconsulting.com

Collaborative Consulting is a leading professional services organization dedicated to helping clients optimize their business and technology capabilities. We combine a powerful blend of business knowledge and market-leading technology expertise with an effective partnership approach, allowing us to understand even the most complex business problems. By first assessing an organization's existing technology abilities and business processes with exceptional accuracy, we determine the best solutions and execute flawlessly. Aligning business and technology initiatives enables our clients to achieve superior, cost-effective business solutions. Founded in 1999, the organization serves clients across the U.S.

Collaborative's Web site is www.collaborativeconsulting.com.



DataLever Corporation
1515 Walnut Street, Suite 200
Boulder, CO 80302
303.541.1516
Fax: 720.294.8344
info@datalever.com
www.datalever.com/edq

DataLever Corporation is a leading provider of next-generation data-quality management technology. The company's products enable organizations to quickly achieve a single integrated 360-degree view of their customer, partner, supplier and other data. Clients use DataLever's software solutions to understand and improve the quality of the data that drives their business.

DataLever Enterprise Server provides a full range of data management capabilities, including data access, data assessment, data cleansing and quality, data reengineering, spatial analysis, and data integration. By automating tasks and seamlessly integrating a broad range of data management functions, DataLever Enterprise Server accelerates the delivery of mission-critical business systems while mitigating risk.



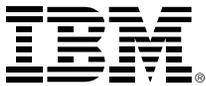
Firstlogic
Corporate Headquarters
100 Harborview Plaza
La Crosse, WI 54601-4071
608.782.5000
Fax: 608.788.1188
information@firstlogic.com
www.firstlogic.com

Firstlogic offers data quality software and services to help global organizations maximize the value of their data and make effective business decisions. Firstlogic's data profiling solution measures, analyzes, and reports on the quality of information and is designed to help users develop a repeatable process for monitoring data quality issues. Its data quality technology cleanses and standardizes customer and operational data, appends third-party information, and builds relationships and corporate households through matching and consolidating records. The software seamlessly integrates into leading enterprise applications. Firstlogic also offers strategic data quality consulting services that can help start, enhance, and continually monitor data quality initiatives.



Informatica Corporation
100 Cardinal Way
Redwood City, CA 94063
650.385.5000
Fax: 650.385.5500
www.informatica.com

Informatica Corporation delivers data integration software and services to solve the problem of data fragmentation across disparate systems, helping organizations gain greater business value from all their information assets. Informatica's open, platform-neutral software reduces costs, speeds time to results, and scales to handle data integration projects of any size or complexity.



**IBM Information
Integration Services**
50 Washington Street
Westborough, MA 01581
800.966.9875
[http://ibm.ascential.com/company/
contact_form.html](http://ibm.ascential.com/company/contact_form.html)
<http://ibm.ascential.com/>

Organizations must ensure they have access to authoritative, consistent, timely and complete information.

The IBM® WebSphere® Information Integration platform integrates and transforms any data and content to deliver information clients can trust for their critical business initiatives. It provides breakthrough productivity, flexibility and performance, so clients and their customers and partners have the right information for running and growing their businesses. It helps clients understand, cleanse and enhance information, while governing its quality to ultimately provide authoritative information. Integrated across the extended enterprise and delivered when the client needs it, this consistent, timely and complete information can enrich business processes, enable key contextual insights and inspire confident business decision-making.



**TRILLIUM
SOFTWARE**

Harte-Hanks Trillium Software
25 Linnell Circle
Billerica, MA 01821
978.436.8900
Fax: 978.670.5793
www.trilliumsoftware.com
trinfo@trilliumsoftware.com

Harte-Hanks Trillium Software® has been selected by companies worldwide, both large and small, to improve their operational and analytic business decisions through accurate and timely information. Trillium Software offers an integrated suite of Total Data Quality software and services architected to discover and correct today's data quality problems and establish a platform prepared for tomorrow's yet unknown data challenges. The Trillium Software System® is recognized as critical to the success of customer relationship management, master data management, customer data integration, data warehouse, business intelligence, enterprise resource planning, supply chain management, e-business, and other enterprise applications, and data integration, data migration, data stewardship, and data governance initiatives.



5200 Southcenter Blvd.
Suite 250
Seattle, WA 98188
Phone: 206.246.5059
Fax: 206.246.5952
E-mail: info@tdwi.org
www.tdwi.org