IBM Data Quality Management solutions: Building opportunity, managing risk

Introduction

Can you trust your information? A simple question—yet the implications are enormous.

With government regulations on the rise, from Sarbanes-Oxley in the United States to the equivalent European Sarbanes Oxley and the Japanese Financial Instruments and Exchange Law (J-SOX), poor governance of information can have an enormous impact including hefty fines, damaged reputations and declining market share.

At the same time, innovative competition and global expansion have created a critical need for trusted, relevant information. Now, more than ever, business decisions must be informed—and the quality of that information is a critical element in effective decision making.

Meanwhile, customer expectations are rising to unprecedented levels. Companies are expected to maintain complete, accurate information about a customer and their relationship history. They expect to be able to share in the management of their data, and breaches of trust are not forgiven. Reputations and revenues suffer the consequences.

In short, organizations must ensure uninterrupted availability, transparency and integrity of complete and accurate data. Doing so requires the ability to:

- *Understand the source, definition, and reliability of its enterprise information at an elemental level.*
- *Enforce standards about how the data is used and maintained.*
- *Trace information changes over time.*
- *Know who is accountable.*

The first step toward data quality is data governance.

IBM Data Quality Management solutions are designed to help you measure, improve and certify the quality of data within your organization.

This white paper explains how IBM can help your company assess and maintain the quality and transparency of your critical data in light of regulatory requirements, legal support and business needs. In addition, it outlines the specific resources IBM has developed to help implement a data quality strategy to facilitate anywhere/anytime access to information, to conduct business operations and quickly respond to compliance audits or legal discoveries.

Ensuring quality to ensure success

There's a compelling reason why organizations are increasingly focusing on data quality as a mission-critical competency: Astute business executives know that trusted data quality can increase revenue, reduce costs and mitigate risk. Data quality directly impacts the viability of the complex decisions corporate leaders must make regarding strategic and operational strategies. Poor data quality can quickly create a critical failure point. Most business challenges, in fact, require a view of information that represents all seven quality characteristics of data — correctness, integrity, precision, timeliness, completeness, validity and consistency — in order to be effective.

For example:

- *Business intelligence often provides the justification for revenue-enhancing initiatives like strategic sourcing, business transformation and product development, as well as reputation-enhancing strategies like customer-focused experience and behavior-driven marketing. Again, the data quality across corporate, supplier and customer systems is vital to these initiatives.*
- *Operational transformations will only deliver expected cost reductions if the data it was based upon was of high enough quality to produce accurate forecasts.*
- *Compliance standards and regulations often require the accuracy of data to be proven and traceable. The consequences of poor data quality in these initiatives can be dire for executives.*

In today's fierce, global market — regardless of industry or geography — the ability to depend upon data quality means survival. Organizations benefit tremendously when their strategic systems — including data warehouses, enterprise resource planning (ERP), customer relationship management (CRM) and legacy systems — use trusted, consistent information across every silo in the enterprise. By enabling a clear understanding of the data, its reliability at the point of capture, and its accuracy over time, organizations can help make better informed decisions, improve the bottom line, and meet with the expectations of both auditors and customers, neither of which are easy to please.

What's getting in the way?

The lack of clear-cut, ready-made answers is what stumps many companies. The problem is that "quality" is such a subjective word: what makes perfect sense for one company would be intolerable at another. A list of data quality metrics can be assembled from industry standards, but the daunting task of prioritizing those metrics, setting limits, designing enforcement mechanisms against such an elusive, technology-driven concept as electronic data is often too much. They stagnate for lack of strategy.

Another dynamic making data quality difficult is that many companies are facing the forest/trees situation: They have literarily lost their data with the complex maze of unintegrated stand-alone applications, hard-coded processes buried millions of lines deep, and silo-sourced systems. This underbrush infrastructure has grown for years as companies expanded through acquisition, and measured individual on siloed achievement regardless of duplication or efficiency. This complexity results in multiple views of critical business information across the multiple systems and an absence of a single, unified view of the truth. Most organizations are all too aware of this problem: In a recent IBM Data Governance Survey 2006, nearly 41 percent of CEOs surveyed felt they had redundant and multiple instances of data.[1] In short, they now realize the value of the data assets they own, but can't find them.

The situation only gets worse over time. As data is added to new systems, its relationship to other systems goes undocumented, uncataloged; over time the information is lost. Executives cannot understand how two numbers on two reports can be so different when they are supposed to say the same thing; analysts cannot trace the transformations and manipulations that have occurred. Questions either become too difficult or too time consuming to answer based on facts, so they fall back to running the business on hunches and gut instinct.

Add to this complexity the fact that the volume and variety of information continues to explode. Transactional databases, content repositories, workgroup documents, mobile devices and RFID tags have all contributed to a continuing explosion of digitized information. Finally, both organizational parochialism and lack of accountability can contribute to poor data quality.

> *"Although an increasing number of organizations recognize the negative impact of poor data quality, most address the issue in a tactical, reactive fashion. Traditionally a minority of organizations had established a formal program and few such initiatives reached a mature, highly effective state."*[2]

Data quality challenges need to be addressed through formalized business-driven strategies at the enterprise level managed through document governance standards and policies. Data governance establishes guidelines and standards to clarify what constitutes data quality within the organization, and delivers processes to manage that quality over time through innovation, enforcement and accountability across the entire organization. The core objectives of an effective data quality governance program are to:

- *Create a reliable, trustworthy enterprise asset.*
- *Establish business-driven quality standards.*
- *Ensure data is consistently defined and well understood.*
- *Increase the use and trust of data as an enterprise asset.*

11 Disciplines of Data Governance Maturity

Organization
Addresses Data Governance organizational structure, alignment, and culture. Data Governance maturity is derived by the level of teamwork between the business and IT among other elements.

Stewardship
Stewardship is a quality control discipline designed to ensure custodial care of data for both asset enhancement, risk mitigation, and organizational control.

Policy
Policy is the written articulation of desired organizational behavior.

Value Creation
The process by which data assets are qualified and quantified to enable the business to maximize the value created by data assets.

Data Risk Management
The methodology by which data risks are identified, qualified, quantified, avoided, accepted, mitigated, or transferred out.

Security / Privacy / Compliance
Describes the policies, practices and controls used by an organization to mitigate risk and protect data assets.

Data Architecture
The architectural design of structured and unstructured data systems and applications that enable data availability and distribution to appropriate users.

Data Quality
Methods to measure, improve, and certify the quality and integrity of production, test, and archival data.

Business Glossary / Metadata
The methods and tools used to create common semantic definitions for business and IT terms, data models, types, and repositories. Metadata that bridge human and computer understanding.

Information Lifecycle Management
A systemic policy-based approach to information collection, use, retention, and deletion.

Audit & Reporting
The organizational processes for monitoring and measuring the Data value, risks, and efficacy of Governance.

Ongoing data quality assessment and remediation is a vital component of a successful data governance program, promoting:

- Timely, integrated data delivered to support strategic opportunities.
- Relevant information shared across channels and LOBs.
- Trusted information to support regulatory compliance.
- Reliable information for business intelligence and executive dashboards.
- Complete view of a customer and their relationship history.
- Insightful data to accelerates time to market for new products and services.

Data Quality Management is an important part of an overall data governance strategy. It offers a methodology by which organizations can define, prioritize and apply quantifiable metrics to data quality, assess the current state of the data, and create a path that will enable business leaders to value the data they receive because it is relevant, accurate, timely and reliable. Data Quality Management can help your company answer questions about your data, including:

- ***Accuracy:*** *Is the data true?*
- ***Integrity:*** *Does this data mean what we believe it means?*
- ***Uniqueness:*** *Does each record represent a unique element?*
- ***Consistency:*** *Is a data element the same regardless of where it is located?*
- ***Relevance:*** *Is this the data I need to answer this question?*
- ***Necessity:*** *Do we need to spend the resources to maintain this data?*
- ***Standardization:*** *Does data adhere to common standards?*
- ***Fitness for purpose:*** *Is the data reliable enough to use for a given purpose?*
- ***Traceability:*** *Where has the data been, what has happen to it, and why?*
- ***Insight:*** *Is data quality getting better or worse as time goes on?*

Implementing a data quality management strategy

There are several critical elements to an effective data quality management strategy.

First, you must take control of your data at an enterprise level. This requires a governance structure that has authority from the highest executive to establish standards and guidelines, to create requirements that the organization must drive to meet through innovation, collaboration and focus. Business leaders must break down traditional boundaries claim ownership, and take responsibility for the data.

Next, you need a foundation platform that provides a centralized place to maintain and measure data quality rules over time, while allowing these rules to be universally deployed. The design environment for these rules needs to be business-driven—weighting and defining the rules according to the requirements of the business.

Beyond the platform and tools for data quality, you also need an ongoing process for measuring and managing data quality. This includes the ability to measure quality improvement or degradation over time, so that rules can be adjusted accordingly. Data quality rules need to be universally deployed across all points of entry—and measured and reinforced in all places where data is processed.

But most important, company leaders must recognize the value of their data assets—for its potential to grow revenues, cut costs and mitigate risks—and take responsibility to ensure it is developed, cultivated, managed and used for the company's strategic advantage.

Using proven best practices, technologies and methodologies as a guideline to establish a Data Quality Management program, organizations can take the first step toward:

- *Taking ownership and responsibility for the quality of its data.*
- *Understanding enterprise data, its meaning, its sources and its history.*
- *Ensuring that data in various repositories, including data warehouses, data marts, master data, operational data stores and transactional applications, is accurate and trusted.*
- *Cleansing data, remove duplicates and rectify misinformation to facilitate accurate information accessible for reporting, analysis and other purposes.*
- *Resolving inconsistencies in disparate data sources.*
- *Maintaining the accuracy of information over time at all cycles: as it is acquired, when it is moved or while it is stationary.*

Driving value through end-to-end data governance solutions

IBM offers end-to-end data governance solutions that span a range of business and IT processes that support successful Data Quality Management objectives. These solutions seamlessly integrate with each other and other IT processes to help you achieve an end-to-end view of your data quality management measures. Using a highly modular approach, you can implement the process areas that help generate the greatest value today, and then build out more as your needs change.

Data Quality Management Governance

IBM has a global team of experienced professionals to assist your company in creating a governance structure tailored to your specific corporate culture and strategic vision. Techniques like the business-driven Data Governance Maturity Model and the Readiness Assessment will help you chart your course while templates and methodologies help you navigate the rough waters of governance transformation. And the experience and guidance offered by IBM consultants will provide you the confidence you need to achieve your goals.

IBM Information Server

At the core of any data quality program is a platform that is capable of managing and measuring data quality, and universally deploying auditable data quality processing logic. Using a metadata-driven approach, IBM Information Server provides the ability to understand and measure any kind of information from any kind of source systems, capture business rules and definitions and apply this understanding to the design of data quality rules. It ensures rules are driven from the baseline of understanding, leveraging metadata to accelerate and automate the development of rules, and providing traceability and auditability throughout the process.

Summary

Data quality has become increasingly critical to business operations. As the type and volume of data continues to explode, organizations are scrambling to ensure accurate, trusted information is available when and where it is needed. Without this assurance, the burden of compliance regulations becomes that much greater—and riskier. Gone too, is the opportunity to leverage information more effectively to improve decision-making, gain insights into customer needs, and create innovative products and services.

Yet while many companies are ready to address data quality improvement, it has been difficult to know what metrics and standards to apply. Now, through proven business technologies, collaborative methods and best practices, IBM Data Quality Management solutions can help organizations of all sizes, across all industries, measure, improve and certify the quality and integrity of data.

For more information

To learn how IBM can help you get started on the path to Data Quality Management, visit http://www-306.**ibm.com**/software/data/information/trust-governance.html or call a representative at 1 877-426-3774.

IBM Software Group
Route 100
Somers, NY 10589

Produced in the United States of America
08-07

[1]IBM Data Governance Survey 2006

[2]Ted Friedman, Gartner, 2006

Each IBM customer is responsible for ensuring its own compliance with legal requirements. It is the customer's sole responsibility to obtain advice of competent legal counsel as to the identification and interpretation of any relevant laws and regulatory requirements that may affect the customer's business and any actions the customer may need to take to comply with such laws. IBM does not provide legal advice or represent or warrant that its services or products will ensure that the customer is in compliance with any law.

LIW11882-USEN-00