

The Data Governance Maturity Model

Establishing the People, Policies and Technology That Manage Enterprise Data

Over the next two years, more than 25 percent of critical data in Fortune 1000 companies will continue to be flawed, that is, the information will be inaccurate, incomplete or duplicated...¹

Today, businesses are discovering that their success is increasingly tied to the quality of their information. Organizations rely on this data to make significant decisions that can affect customer retention, supply chain efficiency and regulatory compliance. As companies collect more and more information about their customers, products, suppliers, inventory and finances, it becomes more difficult to accurately maintain that information in a usable, logical framework.

The data management challenges facing today's business stem from the way that IT systems have evolved. Enterprise data is frequently held in disparate applications across multiple departments and geographies. The confusion caused by this disjointed network of applications leads to poor customer service, redundant marketing campaigns, inaccurate product shipments and, ultimately, a higher cost of doing business.

To address the spread of data - and eliminate silos of corporate information - many companies implement enterprisewide data governance programs, which attempt to codify and enforce best practices for data management across the organization. Although the goal is clear - the quality of information *must* improve to support core business initiatives - there is no definitive roadmap for starting these projects.

For any organization, the first step to address the quality and value of corporate data is to take an honest assessment of the data management infrastructure. Through the Data Governance Maturity Model, organizations can identify and quantify precisely where they are - and where they can go - to create an environment that can deliver and sustain high-quality information. This paper explores:

- The major issues of building better data across the enterprise
- Ways to utilize the existing people, business policies and technology to achieve more effective data quality policies across multiple departments
- How to determine the maturity of an organization's data management capabilities - and find a data governance strategy that fits the organization

¹ Gartner, Inc press release. "'Dirty Data' is a Business Problem, Not an IT Problem, Says Gartner," March 2, 2007.

The Impact of High-Quality Data

In the past, business units were only concerned with entering and tracking data to meet the needs of their specific departments. The result for the enterprise was a buildup of redundant, inconsistent, and often contradictory data, housed in isolated departmental applications from one end of the organization to another.

However, two significant forces complicate every company's data-driven projects. First, the amount of data is increasing every year; IDC estimates that the world will reach a zettabyte of data (1,000 exabytes or 1 million petabytes) in 2010.² Second, a significant portion of all corporate data is flawed.

The effect of this avalanche of bad data can be stunning. Larry English, author and information quality pioneer, writes, "Process failure and information scrap and rework caused by defective information costs the United States alone \$1.5 trillion or more."³

The impact of poor-quality data was initially felt in applications that focus on customer information: database marketing, data warehousing and customer relationship management (CRM). Data quality and data integration technologies, when applied locally or departmentally, could address some of these problems within each application. However, this only led to silos of consistent, accurate and reliable data. The goal is to move beyond those silos and find a way to manage this data across departments, applications, business units and divisions.

The benefits to a holistic approach are obvious; better data drives more effective decisions across every level of the organization. With a more unified view of the enterprise, managers and executives can create strategies that make the company more profitable. A successful enterprise strategy will encompass three main elements:

- **People** - Effective enterprise data governance requires executive sponsorship as well as a firm commitment from both business and IT staffs
- **Policies** - A data governance program must create - and enforce - what is considered "acceptable" data through the use of business policies that guide the collection and management of data.
- **Technology** - Beyond data quality and data integration functionality, an effective data governance program uses data synchronization technology, data models, collaboration tools and other components that help create a coherent enterprise view

² Mearian, Lucas. "A zettabyte by 2010: Corporate data grows fiftyfold in three years." *Computerworld*, March 6, 2007.

³ English, Larry. "Plain English about Information Quality: Information Quality Tipping Point." *DM Review*, July 2007.

The amount of data - and the prevalence of bad data - is growing steadily.

The Data Governance Maturity Model

Best-practice data quality programs are not a one-shot measure (clean up and move on)... To achieve such results, successful programs identify the organizational processes behind data quality. Much like regular IT housekeeping, from virus scanning or performance monitoring to data backup, the data quality program becomes part of daily IT routine.⁴

The Data Governance Maturity Model helps organizations understand their current level of data management. More importantly, the model can identify a path for growth in the future. While achieving a single, unified enterprise view is an evolutionary process, an organization's growth toward this ultimate goal invariably follows an understood and established path. The four distinct stages are:

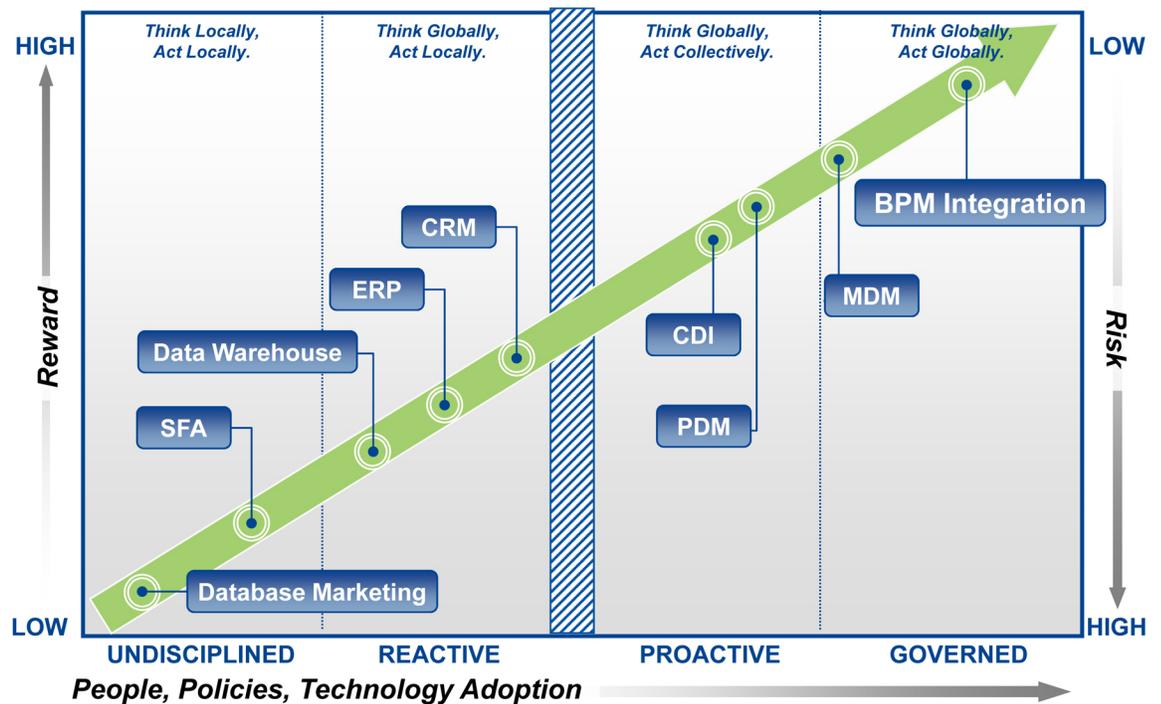
1. Undisciplined
2. Reactive
3. Proactive
4. Governed

It's important to identify the current stage of operation and understand why the organization is there. Companies that plan their evolution in a systematic fashion gain over those that are forced to change by external events. The Data Governance Maturity Model can help control that change by determining what stage is appropriate for the business - and how and when to move to the next stage.

Figure 1 shows the Data Governance Maturity Model and the typical use of enterprise applications common to each of its four distinct stages. Each stage requires certain investments, both in internal resources and from third-party technology. However, the rewards from a data governance program escalate while risks decrease as the organization progresses through each stage.

The model depicts the types of technologies where data consolidation and integration often occur. Companies typically try to drive value from data initially within smaller projects (database marketing, for example) and then move to larger projects. The stages of the model are a continuum, and movement from one level to the next is will not happen all at once. There is also a chasm between the second and third stage (described in more detail later), as organizations have found that the resources and commitment necessary to advance from Reactive to Proactive requires critical changes in executive support and corporate buy-in.

⁴ "Organizing for Data Quality." Research note from Gartner Inc., June 1, 2007.



© Copyright 2007 DataFlux Corporation LLC, Cary, NC USA. All Rights Reserved.

Figure 1 - The Data Governance Maturity Model

At the end of the process, companies have implemented a “single view of the enterprise” through master data management (MDM), which leads to opportunities to integrate high-quality data with business process management (BPM) systems. At this level, organizations are using high-quality data to support the automation of routine processes that do not, or should not, require human intervention.

This paper will examine each stage of the model and define the following components:

- **People** - Who is involved and what contributions must they make?
- **Policies** - What activities must be performed? What business rules must be available to properly govern data?
- **Technology** - What investments in technology are necessary?
- **Risk and reward** - What risks does the organization face at the current stage - and what could it gain from progressing forward?
- **Advancing to the next stage** - What actions are required to move from one stage to the next?

Improving the quality of enterprise data requires more than technology.

Undisciplined organizations have limited visibility into data quality problems.

Stage One - Undisciplined (Think Locally, Act Locally)

At the initial stage of the Data Governance Maturity Model, an organization has few defined rules and policies regarding data quality and data integration. The same data may exist in multiple applications, and redundant data is often found in different sources, formats and records. Companies in this stage have little or no executive-level insight into the costs of bad or poorly-integrated data. Not surprisingly, about one-third of all organizations are at the Undisciplined stage.

Table 1: Characteristics of an Undisciplined Organization

<p>People</p> <ul style="list-style-type: none"> • Success depends on the competence of a few individuals • Business analysts are removed from development of data quality rules • Organization relies on personnel who may follow different paths within each effort to reconcile and correct data • No management input or buy-in on data quality problems • Executives are unaware of data problems or blame IT entirely 	<p>Policies</p> <ul style="list-style-type: none"> • Data quality is non-existent or project-focused only, with no defined data quality processes • Data and data processing is siloed - systems operate independently • "Firefighting mode." Address problems as they occur through manually-driven processes • Resources are not optimized due to redundant, outdated data
<p>Technology</p> <ul style="list-style-type: none"> • No data profiling, analysis or auditing is used • Data cleansing and standardization occurs only in isolated data sources • Data improvement is focused on single applications, such as database marketing or sales force automation (SFA) 	<p>Risk and Reward</p> <ul style="list-style-type: none"> • Risk: Extremely high. Data problems result in lost customers or improper procedures. A few scapegoats receive the blame, although it is impossible to accurately assign culpability • Reward: Low. Outside of the success of a single employee, companies reap few benefits from data quality

Advancing to the next stage

An organization's exposure to risk in the first stage often leads to a single event or series of events that show the impact of poor data quality, such as an increase in customer churn, supply chain disruptions, or other events. At this point, companies recognize problems with data integrity (usually at the departmental or business unit level) and begin to quantify the effects of poor data quality in the organization. When this recognition spurs change, the organization can reach a higher level of maturity.

Reactive companies begin to understand the role of data governance.

To move to the Reactive stage, a company must establish objectives for data governance, starting with an initial assessment that establishes a baseline for data maturity across the enterprise. Transitioning requires organizations to identify the size and scope of data governance efforts (Is it a grassroots effort or is there executive sponsorship?). Also, before moving to the next level, organizations should identify the critical data assets (customer, product, etc.) that will be involved.

The technology components that support this growth must be able to handle data quality and data integration tasks for cross-functional teams. More sophisticated data profiling, standardization and verification capabilities provide a way to refine information across departmental boundaries. In addition, the ability to centralize business rules for core data quality functions in a single repository - and use those same rules across applications - is a critical element that facilitates growth.

Stage Two - Reactive (Think Globally, Act Locally)

A Reactive organization locates and confronts data-centric problems only after they occur. Enterprise resource planning (ERP) or CRM applications perform specific tasks, and organizations experience varied levels of data quality. While certain employees understand the importance of high-quality information, corporate management support is lacking. Studies show that the largest share of all organizations - 45 to 50 percent - fall into this stage.

Table 2: Characteristics of a Reactive Organization

<p>People</p> <ul style="list-style-type: none"> • Success depends on a group of database administrators or other employees • Individuals create useful processes for data quality initiatives, but no standard procedures exist across functional areas • Little corporate management buy-in to the value of data or to an enterprisewide approach to data quality or data integration 	<p>Policies</p> <ul style="list-style-type: none"> • Rules for data governance emerge, but the emphasis remains on correcting data issues as they occur • Most data management processes are short-range and focus on recently-discovered problems • Within individual groups and departments, tasks and roles are standardized
<p>Technology</p> <ul style="list-style-type: none"> • Tactical data quality tools are often available, such as solutions for data profiling or data quality • Applications like CRM or ERP utilize data quality technology • Most data is not integrated across business units; some departments attempt isolated integration efforts • Database administration tactics emerge (e.g., reactive performance monitoring) 	<p>Risk and Reward</p> <ul style="list-style-type: none"> • Risk: High, due to a lack of data integration and overall inconsistency of data throughout the enterprise. While data is analyzed and corrected sporadically, data failures can still occur on a cross-functional basis • Reward: Limited and mostly anecdotal. Most ROI arrives via individual processes or individuals, and there is limited corporate-wide recognition of data quality benefits

*Moving from
Reactive to
Proactive
requires an
executive-level
commitment to
ensure success.*

Advancing to the next stage

In the Reactive stage, applications are still non-integrated, disparate point solutions. The impetus for progressing to a Proactive stage requires managers and executives to create a new, strategic vision that will ensure that processes are in place to correct and consolidate data, leading to tangible business results.

The move to Proactive is not an easy one (hence, the chasm depicted in Figure 1). After years of investing time and resources in complex ERP or CRM systems, moving to a more unified enterprise view via customer data integration (CDI) or product data management (PDM) solutions takes a concerted effort across departments and divisions. Business units that are accustomed to maintaining the own applications and data structures may find it difficult to embrace a more corporate view of the data governance strategy. As a result, progressing to the next stage requires a high degree of executive support - and a resulting culture shift - to create a more unified view of the organization.

Once a vision and strategy has been established, the move to Proactive requires the creation and codification of a data governance team (sponsors, stakeholders, domain experts and data stewards). This team - particularly the data stewards responsible for day-to-day oversight of data quality procedures - establish cross-functional business rules that correspond to identified levels of data integrity. These rules are often based on established best practices that were used effectively during ERP or CRM implementations.

From the technology side, data quality and data integration capabilities become a core component of the cross-enterprise IT platform. The organization is more reliant on SOA to tie data management processes to operational applications, making data quality a critical feature of any system. Finally, companies moving to Proactive use data monitoring technologies to uncover sub-standard data before causes problems.

Stage Three - Proactive (Think Globally, Act Collectively)

Reaching the Proactive stage of the maturity model gives companies the ability to avoid risk and reduce uncertainty. At this stage, data goes from an undervalued commodity to an asset that can be used to help organizations make more informed decisions.

A Proactive organization implements and uses CDI or PDM solutions - taking a domain-specific approach to MDM efforts. The choice of CDI or PDM depends on the importance of each data set to the overall business. A retail or financial services company has obvious reasons to centralize customer data. Manufacturers or distributors would take product-centric approaches. And although the CDI and PDM marketplace has been growing in recent years, less than 10 percent of all companies have reached this level.

Table 3: Characteristics of a Proactive Organization

<p>People</p> <ul style="list-style-type: none"> • Management understands and appreciates the role of data governance - and commits personnel and resources • Executive-level decision-makers begin to view data as a strategic asset • Data stewards emerge as the primary implementers of data management strategy and work directly with cross-functional teams to enact data quality standards 	<p>Policies</p> <ul style="list-style-type: none"> • Real-time activities and preventive data quality rules and processes emerge • Data governance processes are built into the foundation of CDI, PDM and other solutions • Data metrics are sometimes measured against industry standards to provide insight into areas needing improvement • Goals shift from problem correction to prevention
<p>Technology</p> <ul style="list-style-type: none"> • A data stewardship group maintains corporate data definitions and business rules • Service-oriented architecture becomes the enterprise standard • Ongoing data monitoring helps the company maintain data integrity • More real-time processing is available and data quality functionality is shared across different operation modes 	<p>Risk and Reward</p> <ul style="list-style-type: none"> • Risks: Medium to low. Risks are reduced by providing better information to increase the reliability of sound decision-making • Reward: Medium to high. Data quality improves, often in certain functional areas and then in broader realms as more employees join the early adopters

Advancing to the next stage

At the Proactive stage, organizations begin to unify the corporate view of a specific domain (typically customers or products). The next phase creates a unified approach for all corporate information, ultimately leading to the quality of information that can support the automation of business processes.

To progress to the final stage - Governed - a company needs to assemble and integrate many of the pieces already in place. A "Center of Excellence" (or similar framework) emerges to organize the work of multiple data stewards within the enterprise. Business analysts start to control the data management process, with IT playing a supporting role. And the master data efforts provided by CDI and PDM initiatives provide the foundation for business process automation, as the data is now robust and reliable enough to support high-end process management.

The technology required to reach the final stage also centers on the ability to automate business processes. The core components of MDM are in place, and organizations typically need to concentrate on making master data a core component, regardless of the originating application or data type. Through a high degree of data quality, the foundation for supporting full BPM integration is now feasible.

The benefits from the Proactive stage establish the foundation for MDM efforts - and business process automation.

Stage Four - Governed (Think Globally, Act Globally)

At the Governed stage, an organization has a unified data governance strategy throughout the enterprise. Data quality, data integration and data synchronization are integral parts of all business processes, and the organization achieves impressive results from a single, unified view of the enterprise.

Table 4: Characteristics of a Governed Organization

<p>People</p> <ul style="list-style-type: none"> • Data governance has executive-level sponsorship with direct CEO support • Business users take an active role in data strategy and delivery • A data quality or data governance group works directly with data stewards, application developers and database administrators • Organization has “zero defect” policies for data collection and management 	<p>Policies</p> <ul style="list-style-type: none"> • New initiatives are only approved after careful consideration of how the initiatives will impact the existing data infrastructure • Automated policies are in place to ensure that data remains consistent, accurate and reliable throughout the enterprise • A service oriented architecture (SOA) encapsulates business rules for data quality and identity management
<p>Technology</p> <ul style="list-style-type: none"> • Data quality and data integration tools are standardized across the organization • All aspects of the organization use standard business rules created and maintained by designated data stewards • Data is continuously inspected - and any deviations from standards are resolved immediately • Data models capture the business meaning and technical details of all corporate data elements 	<p>Risk and Reward</p> <ul style="list-style-type: none"> • Risk: Low. Master data tightly controlled across the enterprise, allowing the organization to maintain high-quality information about its customers, prospects, inventory and products • Rewards: High. Corporate data practices can lead to a better understanding about an organization’s current business landscape - allowing management to have full confidence in all data-based decisions

At this final stage of the maturity model, a company has achieved a sophisticated data strategy and framework, and a major culture shift has occurred within the entire organization. Instead of treating issues of data quality and data integration as a series of tactical projects, these companies have a comprehensive program that elevates the process of managing business-critical data. With support from executive management and buy-in from all business functions, the program can flourish, creating more consistent, accurate and reliable information to support the entire organization.

More importantly, the company can automate processes that once required minimal (but necessary and time-consuming) human intervention. At this stage, BPM becomes a reality, and enterprise systems can work to meet the needs of employees, not vice versa.

For example, a company that achieves this stage can focus on providing superior customer service, as they can understand various facets of a customer's interactions due to a single repository of all relevant information. Companies can also use an MDM repository to fuel other initiatives, such as refining the supply chain by using better product and inventory data to leverage buying power with the supplier network.

Summary

The amount and the complexity of corporate data in every business is growing. Data is increasingly shared across corporate and geographical boundaries. And the success of any organization will ultimately hinge on the ability to maintain a coherent view of data, both now and in the future.

For any company that wants to improve the quality of its data, it is critical to understand that achieving the highest level of data management is an evolutionary process. A company that has created a disconnected network filled with poor-quality, disjointed data cannot expect to progress to the latter stages quickly. The infrastructure (both from an IT standpoint as well as from corporate leadership and data governance policies) is simply not in place to allow a company to move quickly from Undisciplined to Governed.

However, the Data Governance Maturity Model shows that issues such as data quality, data integration, CDI, PDM or MDM are not "all or nothing" efforts. For example, companies often assume that CDI or PDM is the panacea for their problematic data and that they should implement a new system immediately. But the lessons of large-scale ERP and CRM implementations (where a vast majority of implementations failed or underperformed) illustrate that the goals of CDI, PDM, MDM and BPM are not just a technology issue. The typical result of failure is the lack of support across all phases of the enterprise.

To improve the data health of the organization, organizations must adapt the culture - from how staff collects data to the technology that manages that information - to a data governance-focused approach. Although this sounds daunting, the successes enjoyed by an organization in earlier stages can be reapplied on a larger scale as the organization matures. This minimizes the risk of failure while leveraging the strategies that brought positive changes in the past. The result is an evolutionary approach to data governance that grows with the organization - and provides the best chance for a solid, enterprisewide data management initiative.