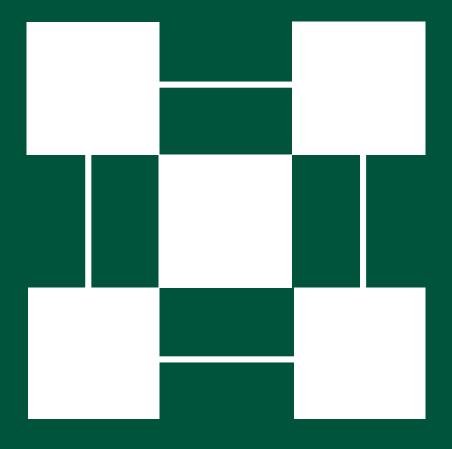
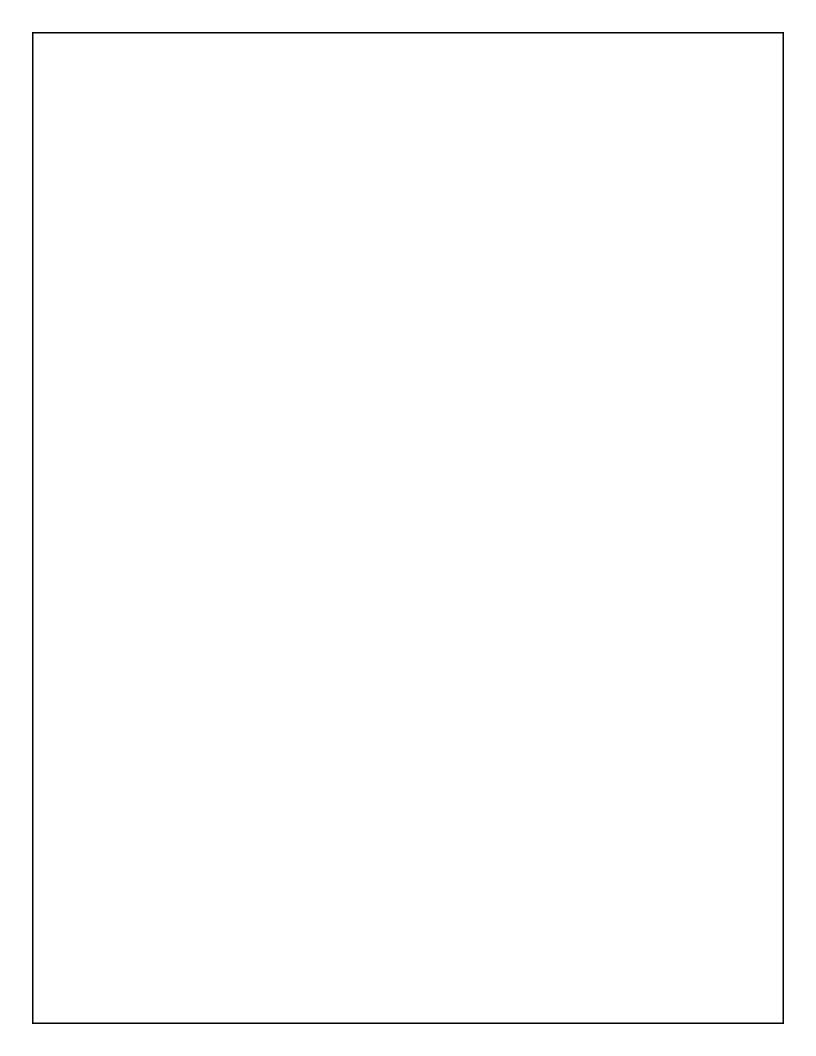
# Data Management Best Practices



MBA Residential Technology Steering Committee (ResTech)

August 2008





# Contents

Introduction	Error! Bookmark not defined.
initi oddction	Enoi: Bookmark not defined.
Purpose	4
Foundations of a Data Management program	4
Data Governance	6
Data Stewardship	9
Data Quality	10
Data Architecture	13
Data Standards	14
Data Security	15
Benchmarking your company's Data Management Maturity	16
Conclusion	18
References	19

# Introduction

We are all aware of the challenges faced and lessons learned over the past year and there's no doubt that every component within the mortgage banking process has been impacted. That can especially be said for technology. Recently we've all come to realize the crucial role that technology plays in our industry—not only in the moment, but also in a proactive capacity. After all, it can be leveraged to improve processes and ultimately prevent, or mitigate, future issues.

When I spoke about this in March at the MBA National Technology in Mortgage Banking Conference, I specifically highlighted technology's business value, providing mortgage banking institutions with innovative systems and practices that are critical to the success of our entire industry.

While technology can aid us in preventing a reoccurrence of the issues we are facing today, we must not restrict future growth of the market and the broader industry. That's why MBA is committed to improving the quality of information that moves through the mortgage supply chain, ultimately creating more transparency for all stakeholders, especially investors and rating agencies.

One specific area that has been identified for improvement is the management of data. With multiple data exchanges between trading partners, we risk compromising data quality. Data management is one of the most important components in the mortgage process and MBA has published this latest white paper to help its member companies improve the quality of their data management practices in order to ultimately increase overall data quality throughout the industry.

The goal of this white paper is to share fundamental strategies and best practices for assessing and improving data management practices. It provides a high-level overview that drills down to the key elements of data management while also offering specific references for further insight and clarification.

I hope you find this information to be both applicable and tangible. I encourage you incorporate these suggested best practices into your own data management initiatives. By doing so, you will not only improve a critical component within your company, you will also positively impact a vital aspect of our industry.

Robert E. Story Jr.
Vice Chairman of the Board of Directors
Mortgage Bankers Association

# **Purpose**

The goal of this paper is to give the reader a basic understanding of how they can improve the way their company manages its data. It will review the basics of Data Management by touching on the salient points of such a program. It will also provide a means for the reader to benchmark their firm's maturity in the data management space and give references to where to find additional information on the various topics presented.

The intended audience for this paper is C-level executives. It seeks to provide them with the information as to why they should be interested in improving their data management. One reason can be seen by looking at examples in other industries where data has been utilized as a competitive weapon. In retail Wal-Mart has outpaced its competition by creating a renowned data management program. However, Wal-Mart and other retailers have a business model that involves the manufacture, shipment, and exchange of goods.

The mortgage industry has a distinct advantage over retail in that it does not have to deal with hard goods. The mortgage business deals with the exchange of property titles for promissory notes. Whereas Wal-Mart has to expend untold sums in moving actual "stuff" around, the mortgage business only involves the movement of data. If the mortgage industry gets data "right" there is no limit to the possibilities for the industry.

In order to get data right we need to engage the key leaders in the discussion of what needs to be done to improve the way data is managed in our industry. Hopefully this paper will be the start of that dialogue.

# Foundations of a Data Management program

The difficulty of effectively managing data is not unique to the mortgage industry. The proliferation of data throughout organizations is a common problem regardless of industry. Many firms struggle with managing their data strategically and transforming the data into actionable information.

Instead of having a comprehensive Data Management program these firms often rely on thousands of Microsoft Access databases, spreadsheets, and silos of data strewn throughout the organization. This leads to several problems; reports do not agree, additional time is required to provide executives with information, and confidential information is not effectively protected.

Many times these issues are perceived as solvable solely by technical means. The leaders think that an Enterprise Data Warehouse will cure all their problems. They then purchase tools, hire consultants, and create large programs to build a Data Warehouse. Many times these projects are less than successful. Gartner once estimated that 50% of all Data Warehousing projects fail, although this statistic has been debated.

Data needs to be treated as a strategic asset that can be used as a trusted source of information throughout the enterprise. Data needs to be managed throughout its entire lifecycle. This requires a comprehensive program rather than just a project to build a single data warehouse.

The practice of Data Management was created to ensure that users can have confidence in the data assets that they utilize to make business decisions. It seeks to create a comprehensive program for the management of enterprise data as opposed to just building a data warehouse.

Wikipedia defines Enterprise Data Management as "...focused on the creation of accurate, consistent and transparent data content. Enterprise Data Management emphasizes data precision, granularity and meaning and is concerned with how the content is integrated into business applications as well as how it is passed along from one business process to another." <sup>1</sup>

Data Management is comprised of multiple categories. Below is a model that shows how data management spans the practices of Data Quality, Data Architecture, Metadata Management, and Data Stewardship. Other models may differ but the basic concept remains the same.

Focus Area	Description	Why Important?	Artifacts
Data Quality	Defines and enforces quality and integrity standards and tolerances.	Proactively manage the quality and integrity of the data. Must be able to certify accuracy and completeness of data.	•DQ Strategy Document •DQ Implementation Plan •Data Profiling Metrics
Data Architecture	Establishes the future state vision and overall framework for development of business services that are aligned with strategic goals of the business.	Ensures that projects integrate with the overall data strategy. Eliminates silo-based solutions that meet tactical needs but lack strategic perspective. Over time, migrate from the current state "web" of interfaces and data feeds to the future state vision.	Data Interface model     Future state model     Data Architecture roadmap
Metadata Management	Defines how data is collected, labeled and stored through-out all applications and databases.	Provides data standardization, data inventory, and controls. Monitors the data flow as it relates to business processes. Manages and tracks critical data fields. Provides alignment to industry standards (MISMO).	Metadata repository     Business rules     Conceptual Model     Logical Model     Data Dictionary     Project Templates
Data Stewardship	Provides business definition and direction on the use and security of the data	Creates ownership by the business to define data and govern the use, security and dissemination of the data.	Domain Owner Matrix
Data Management			

<sup>&</sup>lt;sup>1</sup> Wikipedia: http://en.wikipedia.org/wiki/Enterprise Data Management

The key point is that Data Management is the overarching practice for these various components. This paper will briefly review the following areas of a Data Management program:

- 1. Data Governance
- 2. Data Stewardship
- 3. Data Quality
- 4. Data Architecture
- 5. Data Standards
- 6. Data Security

## **Data Governance**

Data Governance is a key component of any Data Management program. Data Management is usually perceived to be a technical problem that can be solved by purchasing tools and creating new data repositories. In fact data management is a people issue that requires the alignment of multiple constituents to solve issues where many times there is a lack of clear ownership.

Gartner estimates that organizations spend at least 70% percent of their Business Intelligence (BI) budgets to resolve issues related to people, process, and governance. Betsy Burton, VP and distinguished analyst at Gartner is quoted as saying that "Due to a lack of a cohesive strategy, many organizations have created multiple, uncoordinated and tactical BI implementations, which has resulted in silos of technology, skills, processes and people."

Many organizations establish a Data Governance committee to align the processes and technologies to manage data strategically. The goal is to bring together the people from across the enterprise to create a consistent view of a company's data. This will allow the firm to utilize this data to make more informed decisions, remove redundant data, and increase the accuracy and quality of its data.

The ideal way to initiate data governance is by having support from the most senior business leaders. Likely candidates to chair the board include the CFO or the COO. The CFO is concerned with financial reporting data, especially if the company has to adhere to Sarbanes-Oxley (SOX) regulations, and the COO has a need for operational data. Both these roles require access to timely and accurate data. Again, the best way to get sponsorship is by making the case for a strong data management program to the executive leadership of the firm.

Information Technology (IT) should have a seat on the board (generally the CIO) but should not run data governance. Since data is owned by the business, the business should be responsible for data governance. Members of the board should include senior members of the various business areas that are dependent on data. Examples include Lending, Servicing, Capital Markets, Finance, Sales, HR, Operations, Credit Risk, Legal,

and IT. This usually encompasses nearly every major business area in a large organization. For smaller organizations the board can be limited to two or three senior business leaders.

In larger organizations if senior executives serve on the Data Governance committee they may not be the most knowledgeable individuals regarding data in the organization. One way to address this is to split Data Governance into a Steering Committee represented by the executives and allow them to appoint a Working Group that will be comprised of data experts from their specific business units. Some larger organizations also create the role of Chief Data Officer whose primary responsibility is the management of corporate data.

The Data Governance committee should be responsible for all aspects of an organization's data management program. This means staffing a Data Stewardship program, sponsoring and overseeing the Data Quality program, approving Data Architecture and Data Security standards, and approving and prioritizing Data Warehouse and Data Mart project work.

The Data Governance committee is the key to ensure an organization's data is managed strategically and can be put to competitive advantage.

The Data Governance charter should be developed and approved by the Data Governance committee. It should document the purpose of Data Governance, the responsibilities of the committee, and the processes involved. The diagram below illustrates an example of how a Data Governance program could be structured.

# Data Governance

Definition: The exercise of authority and control (planning, monitoring and enforcement) over the management of data assets.

- To define, approve and communicate data strategies, policies, standards, architecture, procedures and metrics.
- 2. To track and enforce regulatory compliance and conformance to data policies, standards, architecture and procedures.
- 3. To sponsor, track and oversee the delivery of data management projects and services.
- 4. To manage and resolve data related issues.
- 5. To understand and promote the value of data assets.



#### Suppliers:

- Business Executives
- IT Executives
- Data Stewards
- · Regulatory Bodies

#### Inputs:

- Business Goals
- Business Strategies
- · IT Objectives
- IT Strategies
- Data Needs
- Data Issues
- · Regulatory Requirements

#### Activities:

# 1. Data Management Planning

- Understand Strategic Enterprise Data Needs
- 2. Develop & Maintain the Data Strategy
- Establish Data Professional Roles & Organizations
- Identify & Appoint Data Stewards
- 5. Establish Data Governance & Stewardship Organizations
- 6. Develop & Approve Data Policies, Standards and Procedures
- Review & Approve Data Architecture
- Plan & Sponsor Data Management Projects and Services 8.
- 9. Estimate Data Asset Value and Associated Costs

#### 2. Data Management Control

- Supervise Data Professional Organizations and Staff
- 2. Coordinate Data Governance Activities
- 3. Manage & Resolve Data Related Issues
- 4. Monitor and Ensure Regulatory Compliance
- 5. Communicate, Monitor and Enforce Conformance With Data Policies, Standards, Procedures & Architecture
- 6. Oversee Data Management Projects and Services
- 7. Communicate and Promote the Value of Data Assets

#### Participants:

- Executive Data Stewards
- Coordinating Data Stewards
- Business Data Stewards
- Data Professionals
- DM Leader
- CIO

#### Tools:

- E-Mail
- · Personal Productivity Tools
- Intranet Website

#### Outputs:

- Data Policies
- Data Standards
- · Resolved Issues
- · Data Mgmt Projects & Services
- · Quality Data & Information
- Recognized Data Value

#### Consumers:

- Data Producers
- Knowledge Workers
- Managers & Executives
- Data Professionals
- Customers

#### Metrics

- Data Value
- Data Management Cost
- · Achievement of Objectives
- # of Meetings Held
- # of Decisions Made
- Steward Representation / Coverage
- · Data Professional Headcount
- Data Mgmt Process Maturity

Figure 1 - Sample Data Governance context diagram from DAMA: http://www.dama.org/files/public/DI DAMA DMBOK Guide Presentation 2007.pdf

# **Data Stewardship**

Data is a valuable corporate asset and needs to be actively managed. Through data stewardship, leadership and accountability is provided to establish and enforce data standards, improve data quality and improve data management capabilities within an organization. With effective stewardship, data quality will be improved, data redundancy will be reduced and the appropriate security and controls will be established to safeguard the data.

The Business Data Steward's role within an organization is to provide overall guidance, governance and oversight to the data management activities within an organization. The role of the steward focuses on the following:

- Business definitions & business rules
- Identification of critical data elements
- Data quality monitoring, issue identification & resolution
- Identification of trusted sources of data
- Support in the simplification of the data environments

The Business Data Steward works primarily as a liaison between the business, operations and technology teams, focusing on ensuring that data is treated as a corporate asset. In order for the stewardship function to be successful in an organization, the role must be embraced and enforced within the business. The importance of business ownership of the data cannot be overemphasized; the business must be accountable for data definitions, data quality and the overall definition of critical data.

# **Business Data Steward Responsibilities:**

- Facilitate the use and collection of common business definitions
- Ensure the proper data quality checks are established for critical data used in critical business functions
- Identify, prioritize and track data quality issues
- Provide link between business, operations and technology regarding data
- Align data activities to targeted projects
- Drive cultural transformation and provide thought leadership
- Promote data policies, standards, tools and best practices
- Partner with Enterprise Architecture, Enterprise Data Quality and other Corporate Stakeholders
- To validate that the business meta data is correct

The role of the business data steward is to promote data management practices and principles within an organization. In addition to Business Data Stewards there are also technical data stewards who are more focused on how data is stored and shared between computer systems.

# **Guiding Principles:**

- Use common business definitions
- Move data quality checks as far upstream as possible
- Put data quality checks on critical data elements
- Understand the quality of the data
- Execute common business rules once and centralize
- Ensure that metadata is captured in a central source
- Identify trusted sources of data
- Share data when possible

# Tips for Success:

- Focus on specific business issues-gain alignment with the business
- Start small build when you have successes
- Focus on metrics show real value
- Market the program at all times
- Ensure the right people are identified as stewards

# **Data Quality**

Data Quality is a main reason that organizations initiate data management programs. Data quality issues are a pervasive problem in most organizations. There are many examples of how poor data quality has led to major problems for corporations. In 2001 Finextra conducted a survey that showed that 30% of all trades that need repair are due to poor quality reference data. When respondents were asked how they thought this situation had changed over the past five years, answers were mixed. Over a quarter (27%) thought the situation had got worse - due mainly to growing trade volumes in more sophisticated instruments such as credit derivatives. While reference data initiatives have helped to decrease trade failures in vanilla products such as equities, they believe this is being overtaken by manual processing and poor data standardization for complex instruments. 3

Have things changed since then? Apparently not since in a recent article in the *MBA Tech Newslink*<sup>4</sup>, Jeanne Capachin, Vice President of Global Banking and Insurance Practices at Financial Insights said: "From a risk management perspective, organizations are developing consistency and bringing it in as a discipline for growth. But that can't be achieved without a strong data management layer. Risk management at financial institutions has many different flavors. In credit risk management, current credit models

 $<sup>2\</sup> A\ good\ site\ with\ many\ examples\ is\ Larry\ English's\ InfoImpact\ at:\ http://www.infoimpact.com/publiclyexposediqproblems.cfm$ 

<sup>3 &</sup>quot;The Key to Quality STP and T+1", October 15, 2001, TowerGroup report sponsored by Reuters and Capco.

<sup>&</sup>lt;sup>4</sup> "MBA Tech Newslink": http://www.mortgagebankers.org/NewsandMedia/MBATechNewslink

are insufficient. Because the credit environment has changed, we need to accumulate good data now to repopulate models to better manage risk."<sup>5</sup>

Gartner has reported that 25% of critical data within Fortune 1000 companies will continue to be inaccurate through 2007. Another Gartner study estimates that poor quality customer data costs U.S. business \$611 billion dollars per year in postage, printing, and staff overhead.

These examples all point to the need for improved Data Quality. The U.S. Census Bureau defined data quality as data that is "fit for use." They go on to define data fit for use as having these three attributes:

- 1. Utility the usefulness of the information for its intended users.
- 2. Objectivity whether information is accurate, reliable, and unbiased, and is presented in an accurate, clear, and unbiased manner.
- 3. Integrity the security or protection of information from unauthorized access or revision.

To help apply limited resources to best achieve fitness for use, the Census Bureau further defines utility, objectivity, and integrity in terms of six dimensions of data quality:

- 1. Relevance the degree to which our data products provide information that meets our customers' needs.
- 2. Accuracy the difference between an estimate of a parameter and its true value.
- 3. Timeliness the length of time between the reference period of the information and when we deliver the data product to our customers.
- 4. Accessibility the ease with which customers can identify, obtain, and use the information in our data.
- 5. Interpretability the availability of documentation to aid customers in understanding and using our data.
- 6. Transparency providing documentation about the assumptions, methods, and limitations of a data product to allow qualified third parties to reproduce the information, unless prevented by confidentiality or other legal constraints.<sup>6</sup>

Further compounding the data quality issue is that in addition to incorrect data there is also the problem of inconsistent data. Organizations seem to generate redundant data

8/25/2008 11

\_

 $<sup>^5</sup>$  "Market Changes Shift Data Management Approaches" by Vijay Palaparty 6 http://www.census.gov/quality/P01-0\_v1.3\_Definition\_of\_Quality.pdf

systems at an amazing pace. Users will leverage tools like MS Access to bring data down to their desktops and manipulate it. This practice leads to reports that don't agree and can also lead to costly mistakes. One option to rectify this issue is to centralize data in a data warehouse. Data Warehouse projects tend to be large and expensive efforts. Before embarking on this type of effort it is best to consult experts in the field and interview experienced peers.

There are also methodologies that can be utilized to improve an organization's data quality. Many of these are based on more generic quality improvement methodologies, such as Six Sigma or Total Quality Management. For example, Larry English promotes a practice called Total Information Quality Management – TIQM.

Finally, there are many tools that are offered to assist in improving data quality. These include the following:

- 1. Data profiling initially assessing the data to understand its quality challenges
- 2. Data standardization a business rules engine that ensures that data conforms to quality rules
- 3. Master Data correcting data by comparing it against a master database of standards (for example Geo Coding, which corrects name and address data to US and Worldwide postal standards)
- 4. Matching or Linking a way to compare data so that similar, but slightly different records can be aligned. Matching may use "fuzzy logic" to find duplicates in the data. It often recognizes that 'Bob' and 'Robert' may be the same individual. It might be able to manage 'householding', or finding links between husband and wife at the same address, for example. Finally, it often can build a 'best of breed' record, taking the best components from multiple data sources and building a single super-record.
- 5. Monitoring keeping track of data quality over time and reporting variations in the quality of data. Software can also auto-correct the variations based on pre-defined business rules.
- 6. Batch and Real time Once the data is initially cleansed (batch), companies often want to build the processes into enterprise applications to keep it clean.<sup>7</sup>

The MBA is utilizing the data quality practices documented here to improve the state of the industry. The MBA Residential Technology Providers Forum is currently working to identify the key data elements that are utilized to make an execution decision in the secondary market. They will then trace these key data elements through the mortgage lifecycle to identify how many times changes occur. By doing this they hope to identify those key data elements that have the highest probability of having data quality issues introduced. The goal is to focus on improving the checks for these elements to ensure

<sup>&</sup>lt;sup>7</sup> Wikipedia: http://en.wikipedia.org/wiki/Data\_quality

execution decisions are made with reliable data. This group plans to publish a paper related to this research later this year.

# **Data Architecture**

Data Architecture is responsible for defining the target state for an Enterprise's data and defining the plan for how to achieve it. Data Architecture is one of the four foundational elements of Enterprise Architecture. These elements are:

- 1. Business Architecture the business processes that support the firm's basic purpose
- 2. Application Architecture the applications that support the business processes
- 3. Data Architecture the data produced, stored, and moved between applications
- 4. Technology (or Infrastructure) Architecture the hardware and technology utilized to support the other architectural elements

DAMA defines data architecture roles as "...identifying the strategic data requirements and the components of data management solutions, usually at an enterprise level." DAMA identifies the components of an Enterprise Data Architecture as follows:

- 1) Creating an enterprise data model which logically represents the business and shows the relationships between various subject areas
- 2) Creating diagrams that show the lifecycle of how data moves across major business entities
- 3) Identifying data stakeholder roles, organizations, processes and applications
- 4) Defining how data will flow between applications and databases.
- 5) Assigning data stewards to subject areas, data entities, and application code sets.

Data architecture is also usually involved in defining standards related to data. These standards can range from data naming conventions, to defining the logical data models that will represent the enterprise data, to defining the tools that will be used to manipulate, manage, and store the data. The next section reviews data standards in more depth and explores how MISMO can be leveraged to enhance any mortgage data standardization effort.

Data architecture is also involved in defining metadata. Metadata is "data about data." It gives the data context. For example, for a specific data element such as "Loan Amount" the metadata could include the field name, the field length, the data type (e.g. character, number, date), and the date the data element was last updated. Many times metadata is stored in a centralized tool such as a data dictionary or Enterprise Data repository that allows users to query the metadata to determine the context of a data element and its

8/25/2008 13

\_

 $<sup>8\,</sup>$  "The DAMA Dictionary of Data Management", Mark Mosley Editor

relationship to other data. The definition and management of metadata is a major effort that usually requires the involvement of Data Architectects and the Data Stewards.

The goal of a data architecture program is to strive for simplicity, integration, consolidation, and consistency in design. If different tools are used to solve a common problem this is probably adding to the solution cost. If different methods are being used to build data marts then they will be more difficult to manage and integrate.

Many rapidly changing or immature businesses don't consider data architecture and instead create siloed repositories of data. These businesses feel that data architecture is bureaucratic, will slow down delivery, is too theoretical, or can be done later. If the data clean-up ever does occur it usually requires enormous effort to integrate and consolidate all the disparate, duplicative, low quality data that was created.

It is imperative for an organization to focus on their data architecture and give it the sponsorship required to ensure data is consistently managed throughout the enterprise. A well run data architecture will allow a firm to reliably process information and make better informed decisions regarding its business.

# **Data Standards**

Standardization is a process by which all elements in a business data set are driven to conform to a common understanding of the terms, definitions and values. A data standard may be established by means of authority, custom, or general consent. There must be an agreement among all interacting parties that their data values will conform to that standard.

There are many benefits to data standardization. The first benefit is conformity for comparison. When aggregating data, we achieve more accurate results. The business facts are not invalidated by erroneous data values or misunderstanding of the data. Another benefit is the ability to manage data quality. Records that do not conform to the standard can be flagged as incorrect or directed to a more rigorous review process.

In the Mortgage Industry MISMO was established by the MBA (Mortgage Bankers Association) to coordinate the development and maintenance of business data standards. The organization's mission also includes promoting the voluntary adoption of the standards within the industry. The organization is made up of representatives from every aspects of the mortgage industry.

MISMO has developed an industry business dictionary that serves as a central repository for agreed upon terms and definitions. Each business fact is represented by a single entry in the logical data dictionary (LDD). Data points in the logical data dictionary represent mutually exclusive loan characteristics and type lists. This eliminates the redundant and contradictory data terms and definitions that are currently used throughout the mortgage process. It also eliminates proprietary terms and values that are often poorly defined or

misunderstood. It provides a clearer understanding of the data. The use of the logical data dictionary results in improved data quality, accuracy and reliability as well as a high level of data consistency across the industry.

Sarbanes-Oxley has made the need for rigorous data management more critical than ever before. The improvement in data quality resulting from MISMO adoption enables companies to be more transparent and accurate in financial and regulatory reporting. Improved data quality feeds the need of industry players to clarify performance results to which management must attest. This, in turn, improves corporate governance accountability and strengthens investor confidence.

The MISMO standard enables the seamless exchange of mortgage-related data across multiple internal and external systems more efficiently and without the risk of data misinterpretation. Deployment of reusable MISMO interfaces allows organizations to use those interfaces, processes, and transactions throughout the organization, minimizing errors and increasing data reliability.

The level of granularity in the MISMO data standard enhances the flexibility for the implementation of new products and services. This serves as an enabler to process improvement and improved service levels.

In addition to MISMO the Mortgage Identification Number (MIN) was created by the MBA to facilitate registration of loans on the MERS® System. MERS was created by the MBA in 1995 as the industry's utility to eliminate the need to prepare and record assignments. The purpose of MIN is to provide an industry standard, universal loan number. The goal is that MIN will be the unifying number that all other numbering systems could be associated with until such a day that propriety loan numbers can be phased out.

The adoption of the MISMO standard and MIN should be an important business priority for all who seek greater operational efficiencies and the agility needed to compete effectively in a dynamic market.

# **Data Security**

One of the more critical areas of financial services and the mortgage industry is the protection of critical data. Data security, or Information Security, covers the confidentiality, integrity, and availability of data regardless of the form the data may take: electronic, print, or screen visible. It also can be regarded as the protection of data from unauthorized (accidental or intentional) modification, destruction, or disclosure.

Financial institutions, banks, and mortgage lenders amass a great deal of confidential information about their employees, customers, products, research, and financial status. Most of this information is now collected, processed and stored electronically and transmitted across networks to other organizations. Should confidential information about

customers, finances or a new product line fall into the hands of a competitor, such a breach of security could lead to lost business, law suits or even bankruptcy of the business. Protecting confidential information is a business requirement, and in most cases, an ethical and legal requirement. For the individual, information security has a significant effect on privacy, which is viewed very differently in different cultures.

More information regarding Data Security can be found at the MBA's Information Security Resource Center.<sup>9</sup>

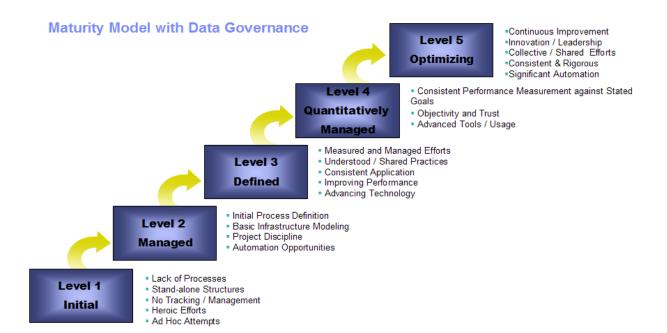
# Benchmarking your company's Data Management Maturity

Launching a data management program inside the company will involve driving changes in processes, technology and culture. Business, Operations and Information Technology must all play a role in realizing the changes required to manage data as a company asset. An important step in the data management program is base lining the current maturity level of the organization. Fortunately there are several assessment tools available from the industry and from leading data vendors. While these assessment tools are new and still evolving, they provide a repeatable, consistent mechanism for assessing and communicating where a company is on the path to improvement. All these tools employ a multi-stage maturity model (usually 4 to 6 stages) derived by answering a set of survey questions. The survey questions attempts to answer the following:

- 1) Strategic intent: How is data perceived across the firm and what level of senior management commitment exists?
- 2) Data Controls: What data policies, standards, governance and controls are in place and at what level of compliance?
- 3) Data Technology: What data warehouses and data tools exist at both the functional and enterprise level? Do data technologies meet the business needs?
- 4) Operations: What information lifecycle processes and performance management reports are in place?

The example below is the 5 stage IBM maturity model:

<sup>9</sup> MBA's Information Security Resource Center: http://www.mortgagebankers.org/InformationSecurityResourceCenter.htm



There are additional assessment models from other organizations. Two examples are:

- The Data Warehousing Institute (TDWI) has an online BI Maturity assessment tool. 10
- DataFlux published a white paper that measures the Data Governance maturity of an organization. <sup>11</sup>

The assessments can be completed as a self assessment or a data management software vendor can be employed to perform the assessment. Surveying a broad spectrum of company employees from executives, data stewards to data providers and data consumers will ensure a comprehensive view of the state of data maturity in your company. There are pros and cons to each approach. The self assessment can be done quickly and is more "cost friendly". The vendor approach provides recommendations for how to move from the current state to the next as well as provide information on how the company compares to other organization in the industry.

If the self assessment method is used, internal resources will still be needed to administer the survey. The internal resource will perform the following tasks:

- 1. tailor the survey questions to the company
- 2. select the surveyed employees
- 3. collect the survey answers (either online or through face to face interviews)
- 4. tabulate the results

<sup>&</sup>lt;sup>10</sup> http://www.tdwi.org/display.aspx?id=8500.

<sup>&</sup>quot;The Data Governance Maturity Model Establishing the People, Policies and Technology That Manage Enterprise Data" Copyright 2008 DataFlux Corporation LLC. All Rights Reserved: <a href="http://www.dataflux.com/resources/resource.asp?rid=184">http://www.dataflux.com/resources/resource.asp?rid=184</a>

# 5. present result findings

As a communication tool, the assessment is an objective, measurable way to communicate the current state of the company's data management maturity but more importantly to communicate progress in advancing the data process, technologies and culture as perceived by important data stakeholders.

# Conclusion

This paper provides an overview into the emerging enterprise data management field. Perhaps some of this information will be assist you in improving the way your company manages its data. We recommend that you utilize one of the benchmarking tools to do an assessment of your current state and then use the references provided here to improve the areas where your organization needs to mature.

Our hope is that this paper acts as a catalyst to help improve the quality of data management practices throughout the mortgage industry. If we all focus on this common purpose it will benefit our industry and once again help people realize their home ownership dreams.

## References

The following organizations and articles were consulted in the formulation of this paper. These are excellent sources for more information regarding data management practices. The references below are for informational purposes only. The MBA in no way endorses these organizations.

Business Intelligence Network <a href="http://www.b-eye-network.com/home/index.php">http://www.b-eye-network.com/home/index.php</a>

Data Management Association (DAMA) <a href="http://dama.org/">http://dama.org/</a>

DAMA's Data Management Body of Knowledge (DMBOK)
<a href="http://www.dama.org/files/public/DI">http://www.dama.org/files/public/DI</a> DAMA DMBOK en v2 1.pdf

DM Review magazine http://www.dmreview.com/

MBA Information Security Resource Center http://www.mortgagebankers.org/InformationSecurityResourceCenter.htm

MBA Tech Newslink <a href="http://www.mortgagebankers.org/NewsandMedia/MBATechNewslink">http://www.mortgagebankers.org/NewsandMedia/MBATechNewslink</a>

MDM Institute http://www.tcdii.com/index.html

**MERS** 

http://www.mersinc.org/

MISMO (Mortgage Industry Standards and Maintenance Organization) website http://www.mismo.org/default.htm

The Data Administration Newsletter (TDAN) http://www.tdan.com/

The Data Warehousing Institute (TDWI) <a href="http://www.tdwi.org/index.aspx">http://www.tdwi.org/index.aspx</a>

"Data Warehouse Failures" by Larissa Moss & Sid Adelman <a href="http://www.tdan.com/view-articles/4876">http://www.tdan.com/view-articles/4876</a>.

"Developing a Data Warehouse Architecture" by Laura Hadley <a href="http://www.users.qwest.net/~lauramh/resume/thorn.htm">http://www.users.qwest.net/~lauramh/resume/thorn.htm</a>

Information Impact International – Larry E <a href="http://www.infoimpact.com/">http://www.infoimpact.com/</a>	inglish site devoted to Information Quality
9/25/2009	20
8/25/2008	20