IBM

IBM Master Data Management: Effective data governance

Introduction

Gone are the days when doing business meant doing so only within the borders of the organization. What used to be single-source data is now multi-source. Today's business world is comprised not only of disparate systems and groups, but users as well; open networks with business partners, customers and suppliers; and diverse architectures and business functions, not to mention global outsourcing and off-shoring efforts. The net? Every link is an exposure and every data element is a risk. In short, business today means your network is the enterprise.

Like any dynamic entity, your organization — along with its data — changes daily. And as the data (whether structured or unstructured), is aggregated and consolidated, to successfully leverage the data across the organization, it must be treated as an enterprise asset. This is especially important for *master data*, the key business facts used across multiple business applications.

Yet, for many organizations with data-sharing environments, complex silos and isolated stovepipes of information and systems hinder business responsiveness and decision makers' ability to make informed decisions. Collaboration among users, functions and systems is often fraught with few clear-cut roles and responsibilities for protecting or enhancing data. Challenges like these illustrate why *data governance* has emerged as a strategic priority for organizations of all sizes.

This white paper will describe how engaging in a master data management (MDM) project enables effective governance of data — specifically master data — and achieves maturity in key categories of the IBM Data Governance Maturity Model.

Data governance and Master Data Management: A symbiotic relationship

In a world where data fuels the business, IT (the data custodian) has to deliver the data effectively and accurately to the entire business. For it to remain viable and relevant, there must be organization-wide support for transforming data, specifically master data into information.

For most organizations, this critical business information is replicated and fragmented across business units, geographic branches and applications. Enterprises now recognize that these symptoms indicate a lack of effective and complete management of master data. IT departments have attempted to gain control over this master data using a variety of methods. But few have demonstrated true success due to their reliance on existing, but repurposed, systems and applications.

What is master data management?

Master Data Management (MDM) is application infrastructure (not a data warehouse, enterprise application, data integration or middleware), designed to manage master data and provide it to applications via business services. MDM enables users to deliver a new class of information-rich applications based on business processes and accurate and complete master data. It supports, augments and leverages your investment in existing applications.

MDM offers:

- *An approach that decouples master information from individual applications and unifies it.*
- *A central, application- and process-neutral resource.*
- *Ensures consistent, up-to-date master information across business processes, transactional and analytical systems.*
- *Proactively addresses key data issues such as **governance**, quality and consistency.*
- *Simplifies ongoing integration tasks and new application development.*

A master data solution that manages master data domains (the high-value, business-critical information about customers, suppliers, products and accounts) and offers IT the capability to transform information into corporate knowledge is a good place to begin getting control over disorderly data.

In general, master data management solutions should:

- *Consolidate data locked within the native systems and applications.*
- *Manage common data and common data processes independently with functionality for use in business processes.*
- *Trigger business processes that originate from data change.*
- *Provide a single understanding of the domain – customer, product, account, location – for the enterprise.*

MDM products vary in their domain coverage, ranging from specializing in a single domain such as customer or product, to spanning multiple and integrated domains. Those that span multiple domains help to harness not only the value of the *domain*, but also the value *between* domains, also known as relationships. Relationships may include customers to their locations, to their accounts or to products they have purchased. This combination of multiple domains (customer, product, account, location, etc.), multiple functions (operational, collaborative authoring, and analytical) and the full set of capabilities in a transactional environment is known as *multiform master data management*.

Key characteristics and capabilities of multiform MDM solutions

Multiple Usage Styles/Functions	Characteristics	Capabilities
Collaborative MDM	Definition, creation, synchronization	Authoring, workflow, check in/out services to support collaboration on master data creation, management and quality control
Operational MDM	SOA management of master data	Business services ingest master data from range of sources, manage it and fulfill all consumer uses of master data; acts as system of record
Analytical MDM	Analysis and insight	Provides accurate, consistent, and up-to-date master data to data warehouses, as well as providing the ability to feed business intelligence insight data back into collaborative and operational MDM
Multiple domains		Support for multiple master data subject areas such as Party, Product, Account and Location
Enterprise business processes and SOA industry models		Integrate master data with data consumers (business applications, accelerators, and industry process and data models)

Applying data governance to MDM

It's at the intersection of people, business processes, information, systems and applications, and multiple MDM functions that data governance emerges. As a result, the logical starting point for data governance is to start by focusing on your master data. Moving to master data management can be the cornerstone of a data governance program. It is important to note that at the same time, moving to MDM *cannot* be successful without data governance.

Both IT and the business have to work together to change business processes, reach agreements on metadata/data models and data sources, as well as institute quality and security mechanisms. In short, a successful master data project requires that data governance play a role. The two exist symbiotically.

Data governance defined

Data governance is the orchestration of people, process and technology to enable an organization to leverage information as an enterprise asset. Data governance manages, safeguards, improves and protects organizational information. Effective data governance can enhance the quality, availability and integrity of your data by enabling cross-organizational collaboration and structured policy-making.

Data governance overcomes silos of organizational self-interest to benefit the overall organization, helping companies leverage the full breadth of their corporate information, directly impacting five factors critical to any organization:

- *Increasing revenue*
- *Lowering costs*
- *Reducing risks*
- *Increasing data confidence*
- *Improving compliance*

As organizations increase their dependency on information, it becomes critical to manage, control and measure the value of data that resides within the organization. Yet, in many cases, companies don't know how to examine their data practices, determine who should be involved and what kind of structure it takes to govern effectively.

With a solid data governance program in place, you can enact a transformational change throughout and beyond the organization that addresses the following data governance issues:

- *What the governance process looks like and who is responsible for governing*
- *What policies are in place, who writes the policies, and how they get approved and changed*
- *Which data should be prioritized, the location and value of the data*
- *What vulnerabilities exist, how risks are classified and which risks to accept, mitigate or transfer*
- *What controls are in place, who pays for the controls and their location*
- *How progress is measured, audit results and who receives this information*

A forum for data governance

With high-profile data breaches and incidents skyrocketing, the challenge to protect and manage data has become a universal concern for organizations. To help better understand this emerging space, IBM created a leadership forum in November 2004 for chief data, security, risk, compliance and privacy officers concerned with data governance issues. Since then, the IBM Data Governance Council has grown to nearly 55 leading companies, universities and IBM Business Partners, including large financial institutions, telecommunications organizations, retailers and public-sector governments.

With a common forum for data governance practitioners to explore challenges and solutions, the Council has developed benchmarks, best practices and guides to successful data governance. Working together, the members of the Council have identified the top governance challenges facing organizations, including:

- *A lack of cross-organizational data governance structures, policy-making, risk calculation or data asset appreciation, causing a disconnect between business goals and IT programs.*
- *Governance policies are not linked to structured requirements gathering, forecasting and reporting.*
- *Risks are not addressed from a lifecycle perspective with common data repositories, policies, standards and calculation processes.*
- *Metadata and business glossaries are not used as to track data quality, bridge semantic differences and demonstrate the business value of data.*
- *Few technologies exist today to assess data values, calculate risk and support the human process of governing data usage in an enterprise.*
- *Controls, compliance and architecture are deployed before long-term consequences are modeled.*

Maturity Model

The Council members also collaborated to define a common benchmark of observable and desired behaviors that every organization can use to evaluate and design their own data governance programs. What emerged was the Data Governance Council Maturity Model. Based on insights and benchmarks from their own practices, the Maturity Model helps define the scope of who needs to be involved in governing and measuring the way businesses govern data uses, such as sensitive customer information or financial details. It enables organizations to:

- *Assess where they currently are in terms of governance, where they want to be and the steps they need to take to get there.*
- *Gain an informed, objective, documented assessment of the maturity of their organization.*
- *Objectively identify, uncover, highlight and detail the strengths and weaknesses of their data management capabilities.*
- *Gain knowledge of existing capabilities and levels of understanding around these elements.*
- *Challenge internal assumptions and normalize methods for continuously examining business processes and IT practices.*
- *Benchmark future levels of organizational performance and develop a roadmap to get there.*
- *Document and centralize reference information that should reside across the organization to govern more effectively.*

Why data governance is critical in a Master Data Management environment

Two primarily rationales for pursuing a Master Data Management are reflected in the value creation and data risk management categories of the Data Governance Maturity Model. Value creation and business insights are much more readily achieved in an MDM environment where the enterprise has a consistent and complete view of the master data.

One example is in the Party domain of master data—having a 360-degree view of your customer can directly lead to upsell/cross-sell opportunities, evaluation of the success of marketing campaigns, and deeper insight into customer concerns. Decoupling master data from applications speeds the development of new applications to exploit the master data. In the Product domain of master data, accelerators such as New Product Introduction for WebSphere® Product Center can significantly reduce the time to market for new products, creating value by cutting cycle time and costs.

Data Risk Management and Compliance can become markedly easier in an MDM environment. Having a single repository of master data makes that data easier to protect, and easier to ensure that only the right people have access to the right parts of the master data at the right time. Compliance is simplified by tracing the provenance (original source) and history of changes. Accesses of the master data are centralized when the data is also moved into the MDM environment, and further enhanced when master data is fed into the well-defined compliance business processes that are part of the IBM Industry Models.

For MDM, the focus from a business perspective are the higher level processes around value creation and data risk and compliance management. However, unless the other categories of data governance are addressed by MDM technology and deployed properly, the benefits of value creation and enhanced data risk and compliance management will be significantly reduced. One key concern is Data Quality (Category 8 in the Maturity Model). To maintain ongoing quality, the master data must be cleansed, removed of duplicates and subject to ongoing quality checks beyond the initial loading of the master data repository. The MDM environment must ensure accurate, complete information with documented methods to measure, improve, and certify the quality and integrity of production, test, and archival data.

**IBM Data Governance Maturity
Model Categories**

Category		Description
1	**Organizational Structures & Awareness**	Describes the level of mutual responsibility between business and IT, and recognition of the fiduciary responsibility to govern data at different levels of management.
2	**Stewardship**	Stewardship is a quality control discipline designed to ensure custodial care of data for asset enhancement, risk mitigation, and organizational control.
3	**Policy**	Policy is the written articulation of desired organizational behavior.
4	**Value Creation**	The process by which data assets are qualified and quantified to enable the business to maximize the value created by data assets.
5	**Data Risk Management & Compliance**	The methodology by which risks are identified, qualified, quantified, avoided, accepted, mitigated, or transferred out.
6	**Information Security & Privacy**	Describes the policies, practices and controls used by an organization to mitigate risk and protect data assets.
7	**Data Architecture**	The architectural design of structured and unstructured data systems and applications that enable data availability and distribution to appropriate users.
8	**Data Quality Management**	Methods to measure, improve, and certify the quality and integrity of production, test, and archival data.
9	**Classification & 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.
10	**Information Lifecycle Management**	Management A systemic policy-based approach to information collection, use, retention, and deletion.
11	**Audit Information, Logging & Reporting**	The organizational processes for monitoring and measuring the data value, risks, and efficacy of governance.

IBM MDM technology is designed and implemented to provide strong support for the enabling and baseline categories in the data governance Maturity Model, ensuring that MDM deployments extract the full value out of the master data. IBM WebSphere Customer Center and WebSphere Product Center, combined with IBM Information Server, effectively addresses MDM for customers. It is with Data Quality and the other enabling and underlying data governance categories (stewardship, security and compliance, etc.) that overall MDM is enabled for customers.

It is important to note that while MDM addresses all 11 categories, IBM MDM technology enables overall master data governance. For the purposes of this paper, the following four categories play a major role in providing the foundation for business value in an MDM undertaking:

2. **Data stewardship** defines who is the real business owner of the information based on the role it plays. Stewardship is not a monolithic role. It is the degree to which an organization manages its information as business assets, and implements executive and management *roles*, *supporting structures*, and *processes* to establish and sustain information accountabilities within the business.

6. **Security, privacy and compliance** delineates the controls (policies, processes and technologies) an organization has put in place to protect its data from misuse (accidental or malicious) based on risk-driven data classification and regulatory requirements.

This category also addresses Operational MDM (Party Domain) where access is controlled at multiple levels. Based on groups of attributes and attribute values, operational MDM defines the operations a user in a specific role can perform (*transactional authorization*) and the data the user can access (*entitlements/ rules of visibility*). It can track data access for auditing and store and respect privacy policies.

Data quality leads to organizational maturity

Organizational maturity can be assessed whether or not formal information quality programs and initiatives are in place, and the degree to which information quality is managed from an end-to-end perspective vs. in silos. As maturity increases, organizations establish *books of record* for their information assets; can show information lineage or *pedigree*, and can demonstrate measurable business results from improved information quality.

This category also addresses Collaborative MDM (Hierarchical Domains) which shares the same concerns of Operational MDM, but also specifies access down the organization's hierarchy. Collaborative MDM also provides authorization for collaborative workflows (who can create, initiate, participate, etc.); and controls access to master data driven by role level.

In addition, this category also addresses privacy and compliance as it relates to security by protecting sensitive information with privacy policies that comply with regulatory requirements, particularly those that apply to customers, suppliers and vendors.

7. **Data architecture** defines in a consistent way the information that comprises master data by examining where the information resides, what the relationships are between different pieces of it and any limitations. The category addressed the real need to author this information when it is spread throughout the enterprise.

8. **Data quality** underscores the degree to which an organization ensures that its core information assets achieve and sustain an appropriate level of quality. Quality means that information is well defined and fit for use, i.e., it is timely, relevant, valid, accurate and consistent. As mentioned above, IBM MDM technology helps ensure accurate, complete information allowing the ability to conduct quality checks, standardize information, eliminate redundancies, deal with matching entries, enable automated merging and kick-back to appropriate data stewards.

10. **Information Lifecycle Management** is the discipline around the planning, collection, creation, distribution, archiving of information, up to retirement and deletion/destruction, based on business and regulatory requirements.

Levels of data governance maturity

Within each of the 11 categories, subcategories exist and are grouped into five levels of maturity. The five levels of maturity include:

- *Level 1 – Initial*
- *Level 2 – Managing*
- *Level 3 – Defined*
- *Level 4 – Quantitatively managed*
- *Level 5 – Optimizing*

Through these multiple levels, organizations can not only assess where they currently are, but can set concrete goals about where they want to be.

Using the Data Governance Maturity Model, organizations can identify and evaluate gaps in their data governance practices and master data management environment, while constantly driving value to:

- *Satisfy auditors and regulators.*
- *Create new opportunities for growth.*
- *Align IT with business.*
- *Improve process management.*
- *Create a more nimble and strategic organization.*

IBM Master Data Management solutions

IBM Master Data Management solutions manage master data domains (customers, accounts, products) that have a significant impact on your most important business processes. The following solutions offer proven technologies and collaborative methods to manage, and build consistency and quality control in master data and governance, helping your organization protect and leverage critical data.

IBM WebSphere Customer Center

WebSphere Customer Center provides real-time, transactional customer data integration (CDI). It helps organizations keep a single, complete and accurate record of their customers across the enterprise.

- *The integrated customer data yields a single version of customer "truth" to all customer-facing channels and front- and back-office systems through multiple interfaces.*
- *It is based on open standards and designed to implement within a Service-Oriented Architecture (SOA). It includes over 480 business services to manage and maintain customer data.*
- *It provides the infrastructure foundation to help companies move to a more customer-centric business model, improving customer service and cross/up-sell execution.*
- *Additional benefits include cost savings from the ability to recognize and process duplicate customer records.*

IBM WebSphere Product Center

IBM WebSphere Product Center is a product information management solution for building a consistent central repository. It links product, location, trading partner, organization and terms of trade information, data typically scattered throughout the enterprise.

- *Based on open standards, WebSphere Product Center provides a middleware foundation for companies to address strategic initiatives and comply with industry standards.*
- *Delivers rich product information to Web sites and e-commerce applications, printed documents and marketing collateral, kiosks and mobile devices as well as directly to customers and trading partners through various access points.*
- *Helps companies efficiently deploy their product and service information across countless customer, partner and employee touch points.*
- *Leverages key information, including product attributes, price and location, making it more accurate and valuable to business processes.*
- *Links product information to product-related terms of trade information, such as pricing, establishing valuable linkages that can be leveraged.*
- *Synchronizes information internally with existing enterprise systems and externally with business and trading partners.*

Data Governance Category	WCC	WPC
Data Architecture	Predefined Model for Party/Role/Organization and more • Based on Industry Standard IFW Model • SOA interface for flexibility and consumability • Extensible in content (data) and behavior	Flexible, extensible model around specs Specs define shape and content of data Hierarchical attributes can be inherited
Security and Privacy	Integration with LDAP for authentication and group membership Transactional authorization Rules of visibility for data-driven authorization Stores and retrieves party privacy preferences	Integration with LDAP for authentication and group management Rich authorization model supporting role, attribute, workflow and hierarchy based authorization Comprehensive auditing
Data Quality	Automated validation of master data Automated duplicate detection Integration with external standardization and validation mechanisms like the IBM Information Server Quality Stage product Rich extensible rules engine for data validation	Built-in validation of master data based on data specifications (specs) Rich set of extension points for custom validations
Data Stewardship	Stewardship tools that support: • Party maintenance • Duplicate suspect processing • Hierarchy maintenance for organizational parties • Grouping of master data	• Role and hierarchy-based stewardship of products • Groups of related attributes
Information Lifecycle Management	• Master Data Lifecycle Event Management plug points for integration with external systems • Support for data lineage and data decay tracking	• Business process workflows around the lifecycle of product information

IBM Information Server

IBM Information Server is a revolutionary new software platform from IBM that helps organizations derive more value from the complex, heterogeneous information spread across their systems.

IBM Information Server, working hand-in-hand with the patented Iterations 2 data integration methodology, supports integration of all data types and multiple architectures, providing your organization with a solid basis for expediting transactions, streamlining operations, supporting customers and making sound decisions.

- *Offers a comprehensive, unified foundation for enterprise information architectures, scalable to any volume and processing requirement.*
- *Provides auditable data quality as a foundation for trusted information across the enterprise.*
- *Delivers metadata-driven integration, providing breakthrough productivity and flexibility for integrating and enriching information.*
- *Provides consistent, reusable information services—along with application services and process services, an enterprise essential.*
- *Accelerates time-to-value with proven, industry-aligned solutions and expertise.*
- *Broadest and deepest connectivity to information across diverse sources: structured, unstructured, mainframe, and applications.*
- *Integrated with IBM WebSphere Customer Center for delivering and standardizing master data, and for detecting duplicates in master data.*
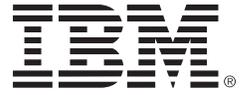
IBM Global Services consultants can help you engage in a master data management project for more effective master data governance by leveraging the Maturity Model into your business strategy. Our data governance and master data management technologies and services are designed to help you understand your organization's data so you can transform how your data is governed, valued, protected and leveraged.

Summary

Data governance is a reflection of your organization's behavior. To govern change, your business should be able to change its organizational controls to meet new demands. IBM solutions and Global Services, along with the IBM Data Governance Council and Maturity Model can help you measure your data governance maturity to take the first step in determining where you are today— and where you want to go tomorrow.

For more information

To learn how the IBM Data Governance solutions and Maturity Model can help you enter or advance in a data governance and master data management environment, visit: please **ibm.com**/software/data/ips/products/masterdata/.

IBM®