July 2014

Governing master data

Maintain a single, trusted view of information with IBM InfoSphere MDM and IBM InfoSphere Information Server

Introduction

A successful master data management (MDM) initiative requires a solid governance foundation. Governance is a control discipline; it orchestrates technologies to meet the demands of certain business policies and rules. Without good governance, data quality, accuracy and access can suffer—leaving both systems and knowledge workers with outdated or incorrect information, and creating an inefficient, divided environment.

While MDM aggregates multiple information sources to create a "single version of the truth," it leverages the control and management principles of information governance to support an authoritative system of record.

In order to deliver a single, trusted view of master data, you must centralize governance. You need robust tools to extract, transform and load data; manage data quality; and govern data, including policy and metadata management. Data from across the organization must be acquired, analyzed, cleansed and reconciled as part of an MDM process aligned with governance policies around security, audit, lineage and so on. Once the data is mastered, governance should be maintained no matter how data is used, what business function it serves or how it is accessed (such as in real time or distributed in batches for analysis). A well-governed MDM solution delivers consistent, high-quality data throughout the enterprise on an ongoing basis.

Together, IBM® InfoSphere® MDM and IBM InfoSphere Information Server provide important components for a comprehensive governance solution. With InfoSphere MDM and InfoSphere Information Server, you can create and maintain a single, trusted view of your critical master data.

Integration challenges with MDM

There are several potential challenges with implementing an MDM system and maintaining it as it grows. For example:

- **Diversity of master data:** In many cases, organizations have multiple data sources, each with its own type of master data. There might be no clear management or ownership of this data across these sources.
- **Data quality issues:** A lack of standards across multiple systems, missing fields and redundant information can cause data quality issues that will persist throughout the lifecycle of an MDM implementation.
- **No single view:** There is no consistent, single view of master data across all sources. There might also be poor understanding of how each piece of data relates to others.
- **Inadequate synchronization:** Some organizations need better ways to keep master data in sync across the organization while providing the unique views of master data that consuming applications require.
- **Ongoing management:** Organizations need to better track changes to master data in source systems and monitor data quality over time so users can consistently trust data.

Without the right processes and tools for maintaining the integrity of master data, eventually its quality can decline. For example, if it is too difficult to apply existing governance rules and policies when integrating new applications, teams may just create a silo of data for that application—which negates the benefits of MDM and puts the environment at higher risk for data errors and inaccuracy. And if data quality degrades, the data—and subsequent business decisions derived from it—cannot be trusted, leading to delays and hesitation on the part of decision makers and developers alike.

InfoSphere MDM powered by InfoSphere Information Server

InfoSphere MDM and InfoSphere Information Server can help you overcome the major challenges of implementing and maintaining MDM through a set of well-defined phases: prepare, deploy and govern (see Figures 1 and 2).

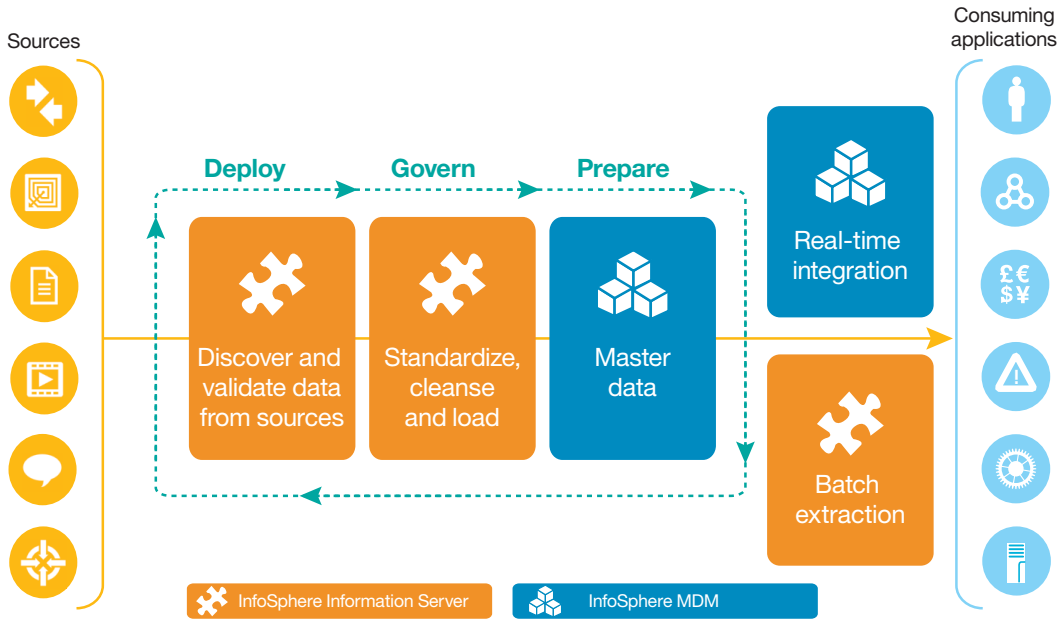

Figure 1. As information moves from data sources to consuming applications, InfoSphere Information Server and InfoSphere MDM help accelerate MDM-powered solutions.

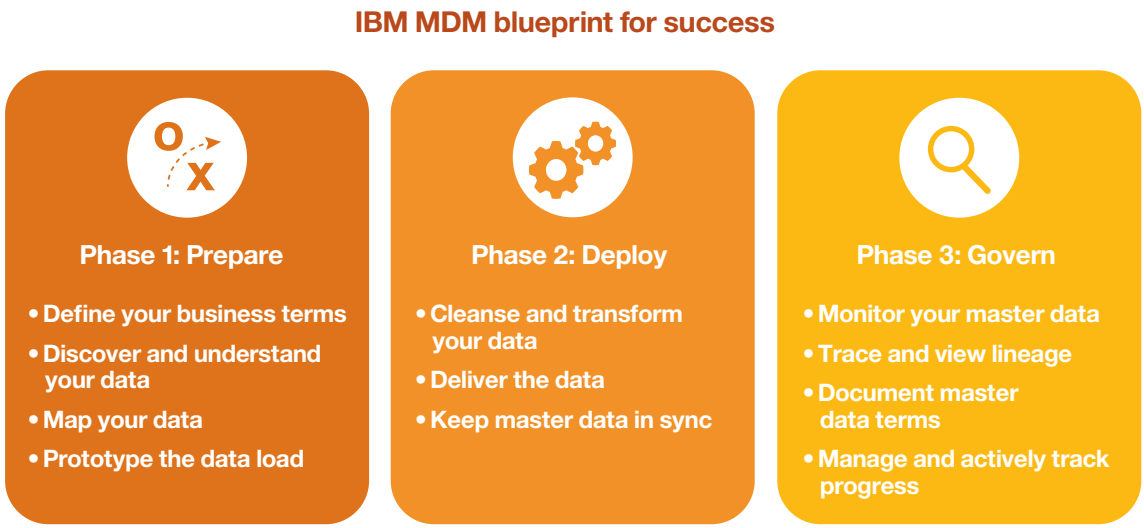

Figure 2. Address the challenges of implementing and maintaining MDM by following a three-phase process.

Prepare

Thorough preparation is critical for successful MDM implementation and master data governance.

- **Define:** Define your master data terms and communicate them across the business and IT.
- **Discover:** Analyze your source systems to understand relationships across your master data.
- **Map:** Use discovery results and profiling analysis to map sources to master data targets.
- **Prototype:** Generate a unified model to prototype your MDM hub.

InfoSphere MDM and InfoSphere Information Server offer integrated capabilities that can help streamline preparation. For example:

IBM InfoSphere Business Glossary and InfoSphere Reference Data Management Hub: InfoSphere Business Glossary maintains definitions for business terms related to information assets in an enterprise. InfoSphere Reference Data Management Hub (InfoSphere RDM Hub) maintains a consistent view of reference data across the organization. Together, the two technologies provide tools for users to govern, manage and make reference data easily accessible to the business. The integration allows InfoSphere RDM to maintain a valid set of business glossary terms, while InfoSphere Business Glossary keeps a set of RDM terms for editing.

IBM InfoSphere Information Analyzer: InfoSphere Information Analyzer is used to assess data quality and facilitate ongoing data quality monitoring and exception management. A pre-built integration sample for InfoSphere Information Analyzer with MDM accelerates deployment and helps ensure a robust MDM design.

Deploy

Accelerate deployment and help ensure governance by establishing processes for cleaning, transforming and delivering master data.

- **Clean:** Analyze the quality of your master data and apply a variety of cleansing techniques at whichever point in the process you prefer.
- **Transform:** Transform the master data coming from multiple sources into an MDM-supported format.
- **Deliver:** Keep master data in sync across the enterprise and move data either into or out of an MDM system for further analytics.

IBM InfoSphere DataStage®: InfoSphere DataStage integrates data across multiple systems using a high-performance parallel framework. Pre-built integration with InfoSphere MDM, implemented as a specialized stage, leverages MDM metadata to automate initial and delta loads of master data to the MDM hub, and to extract data from the MDM hub for transfer to target systems. The integration of InfoSphere MDM with InfoSphere DataStage enables InfoSphere MDM to take full advantage of the parallelism and scalability of InfoSphere DataStage.

IBM InfoSphere QualityStage®: InfoSphere QualityStage delivers comprehensive and customizable data cleansing capabilities in batch and real time. Pre-built integration with InfoSphere QualityStage enables verification of attributes such as names and addresses before the data is loaded into InfoSphere MDM. A common integration interface across all deployment styles of MDM, such as virtual and physical MDM, allows consistent results back from InfoSphere QualityStage, depending on which style of implementation is used. As a result, InfoSphere MDM can maintain enriched and well-cleansed data throughout the governance lifecycle.

IBM InfoSphere Data Click: A feature of InfoSphere Information Server, InfoSphere Data Click helps simplify information provisioning—the process of moving data into new data stores. It also enables you to populate private data marts used for data analytics without jeopardizing the integrity of business information in production environments. Architects can use InfoSphere Data Click to generate both design and operational metadata to support data lineage and impact analysis. They can also use InfoSphere Data Click to move data from InfoSphere MDM into a data warehouse for analysis.

Govern

Effective governance of master data requires a strategy for monitoring, tracing, documenting and managing the most important data assets in your organization.

- **Monitor:** Monitor and report on your master data quality over time, and manage exceptions.
- **Trace:** View lineage of your master data from source to consuming application.
- **Document:** Document your master data terms, mapping specifications and data flows.
- **Manage:** Leverage solution blueprints and actively track end-to-end movement of data and project progress; at a high level, manage rules around the data.

InfoSphere MDM and InfoSphere Information Server can help you successfully govern master data through a common metadata foundation, data quality assessments and other powerful management capabilities.

IBM InfoSphere Information Analyzer: After an MDM solution is implemented, the Master Data Policy hub and InfoSphere Information Analyzer can provide metrics for the master data against a set of rules and policies. InfoSphere Information Analyzer uses either predefined rules or customized rules for validation. These metrics can be reviewed to redefine or change policies for a business user. By assessing data quality at every step of the process, organizations can leverage InfoSphere Information Analyzer as an ongoing quality solution.

IBM InfoSphere Metadata Workbench: InfoSphere Metadata Workbench provides a unified framework for multiple user roles and functions. A role is generally a user persona, but it can also be multiple users working together on different components who want to share the same metadata, or a group of users working on a single component at different points in time. As each new role is created, metadata about it is shared among all the components. This common metadata foundation allows each role to actively share metadata with other roles, without forcing those other roles to use the same technologies or re-map metadata. Each role uses technologies with which it is comfortable, but each can see the flow of information through the entire system.

Capabilities from InfoSphere MDM workbench tooling allow export of services, objects and table models into InfoSphere Metadata Workbench to track data lineage. A user can then view the complete lineage graph, including MDM services and object definitions and MDM table definitions.

IBM InfoSphere Blueprint Director: InfoSphere Blueprint Director provides consistency and promotes communication across your data integration project by linking the solution overview and detailed design documents (see Figure 3). Creating a well-documented and complete blueprint of both the business vision and the technical vision helps your IT department align business requirements with your enterprise reference architecture. It also facilitates sharing of information and progress status.

Conclusion

A complete and continuous governance framework is imperative for a successful MDM implementation. IBM InfoSphere MDM and InfoSphere Information Server provide a rich set of well-integrated and complementary capabilities that can help your organization achieve a mature and well-governed MDM model. Whether you are new to MDM, are in the process of implementing an MDM system or are well along the path to information integration maturity, IBM solutions can help with your governance needs.

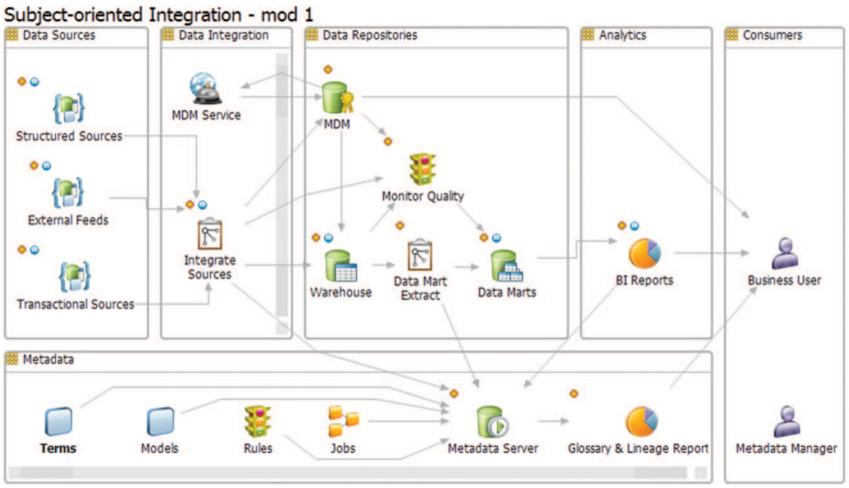

Figure 3. This example from InfoSphere Blueprint Director shows how users can view the entire data integration process through connected stages.

For more information

To learn more about information governance and the integrated solutions offered by InfoSphere MDM and InfoSphere Information Server, please contact your IBM representative or IBM Business Partner, or visit the following websites:

- **ibm.com**/software/data/integration/info_server
- **ibm.com**/software/products/en/infosphere-master-data-management

Additionally, IBM Global Financing can help you acquire the software capabilities that your business needs in the most cost-effective and strategic way possible. We'll partner with credit-qualified clients to customize a financing solution to suit your business and development goals, enable effective cash management, and improve your total cost of ownership. Fund your critical IT investment and propel your business forward with IBM Global Financing. For more information, visit: **ibm.com**/financing

© Copyright IBM Corporation 2014

IBM Corporation
Software Group
Route 100
Somers, NY 10589

Produced in the United States of America
July 2014

IBM, the IBM logo, ibm.com, DataStage, InfoSphere, and QualityStage are trademarks of International Business Machines Corp., registered in many jurisdictions worldwide. Other product and service names might be trademarks of IBM or other companies. A current list of IBM trademarks is available on the web at "Copyright and trademark information" at **ibm.com**/legal/copytrade.shtml

This document is current as of the initial date of publication and may be changed by IBM at any time. Not all offerings are available in every country in which IBM operates.

THE INFORMATION IN THIS DOCUMENT IS PROVIDED "AS IS" WITHOUT ANY WARRANTY, EXPRESS OR IMPLIED, INCLUDING WITHOUT ANY WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND ANY WARRANTY OR CONDITION OF NON-INFRINGEMENT. IBM products are warranted according to the terms and conditions of the agreements under which they are provided.

The client is responsible for ensuring compliance with laws and regulations applicable to it. IBM does not provide legal advice or represent or warrant that its services or products will ensure that the client is in compliance with any law or regulation.

Please Recycle

IMW14780-USEN-00