

Business Intelligence Network™
Research Report

**MASTER DATA MANAGEMENT:
Creating a Single View of the Business**

EXECUTIVE SUMMARY

MASTER DATA MANAGEMENT: Creating a Single View of the Business

Master data management (MDM) has become a very hot topic these days. Unfortunately it is also a source of confusion and frustration because of the disparate definitions, architectures and technologies available in this space. This report clarifies these problems and sets forth a concise, easily understood architecture for MDM into which you can map your MDM projects.

We begin by providing our definition of a mature, enterprise-wide MDM environment:

A set of disciplines, applications and technologies for harmonizing and managing the system of record and system of entry for the data and metadata associated with the key business entities of an organization.

We view the ultimate goal of MDM as much more than just a repository of integrated reference data. We feel that a critical part of MDM is the movement of both the system of entry and record from the operational systems to the MDM environment. While we recognize that not all master data may reach this lofty status in the MDM environment, it is still the goal that should be foremost in MDM implementers' projects.

The benefits of a solid MDM environment are many, but they boil down to three major areas: reduced master data redundancy, more consistent master data and a significantly more efficient business. These benefits translate into real dollar returns on investment as evidenced by the four case studies included in this report.

We also dispel some common misconceptions about MDM – that it is a data warehousing or business intelligence (BI) project, that MDM is solely for maintaining data consistency across business transaction applications, that MDM is simply another data integration project, and finally that MDM projects integrate and manage all enterprise-wide data. These misconceptions have plagued nascent MDM projects from their beginnings. Hopefully setting the record straight here will improve the success rate of these important initiatives.

Our research has generated a well thought-out architecture consisting of three major components for MDM:

- Master data store (MDS) and master metadata store (MMS) – one or more data repositories used for storing and maintaining master data and its corresponding metadata.

- Master data integration (MDI) services – the various technologies used to perform the integration of master data, such as EAI, ETL and EII.
- MDM applications – the processes that employ the MDI technologies to create and maintain an integrated set of master data.

Implemented correctly, MDM can provide significant business benefits in terms of improving productivity, reducing risk and increasing revenues. However, it must be realized that enterprise MDM is a multiyear strategic initiative that can evolve from smaller, more tactical projects provided an enterprise MDM plan is developed to support this evolution.

MASTER DATA MANAGEMENT: Creating a Single View of the Business

By Claudia Imhoff and Colin White

“The whole is more than the sum of the parts.”

Aristotle

This statement by Aristotle sums up the philosophy behind master data management or MDM quite nicely. MDM is about studying, cataloging, and accessing critical sets of data in your enterprise. That is, it means understanding everything you possibly can about your customers, products, locations, and other major subject areas. In this research paper, MDM is defined clearly, the business benefits and issues resulting from its creation are spelled out, techniques and technologies supporting its creation are elucidated, and practical steps for getting started in its implementation are listed.

To further illustrate MDM, its benefits and usage, a number of case studies are presented. Each study starts with the business need, works through the business and technological issues that arose, and concludes with the tips and techniques for avoiding pitfalls or increasing your success rate with the implementation.

WHY MASTER DATA MANAGEMENT?

Master Data Management: A Definition

Let's start by defining what is meant by MDM; there are two perspectives, and it depends on where you place the emphasis in pronouncing MDM:

- *Master data* management – or the management of master data. These are the applications for managing the creation and maintenance of master data. These include customer data integration (CDI), product information management (PIM), and other applications that deal with such data subjects.
- *Master* data management – or the management of the processes behind data management. This refers to the infrastructure and technologies for integrating enterprise data (including master data) such as ETL (extraction, transformation, and load) and EII (enterprise information integration) technologies.

For the purposes of this research paper, we consider the first definition or the management of all forms of master data to be MDM. However, the infrastructure and technologies referenced in the second definition are considered essential to providing the underlying integration platform for supporting MDM processing and applications.

Our official definition is that MDM is:

A set of disciplines, applications, and technologies for harmonizing and managing the system of record and system of entry for the data and metadata associated with the key business entities of an organization.

There are some key words used in this report that are critical to your understanding of the directions set forth. Therefore, these terms are defined in Appendix A. Please familiarize yourself with these terms before continuing to read this document.

In its simplest form, master data is reference data about an organization's core business entities (see Figure 1). The business entities include:

- People – For example, customers, employees (human capital), suppliers, partners.
- Things – These include products, finances (ledgers), assets.
- Places – Locations and geographic points of interest are included here.
- Other key entities – Those sets of cohesive reference data that are of interest to the enterprise. Each enterprise will have its own unique set of these.

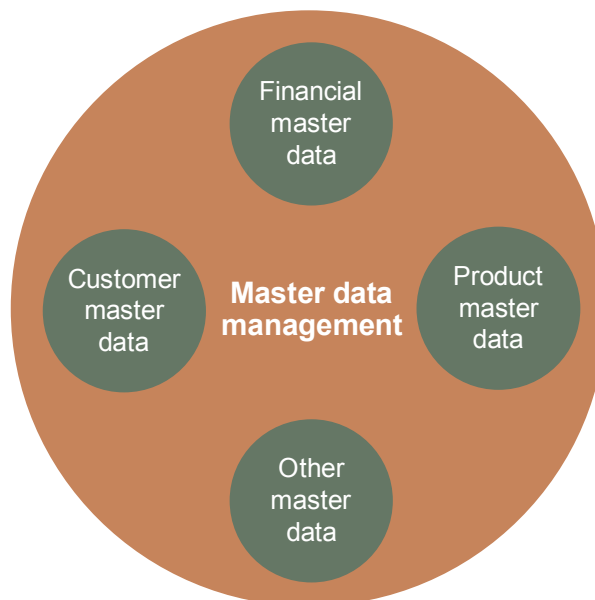


Figure 1: Master data consists of many business entities

MDM applications consist of processes that integrate, store and maintain specific instances of business entities such as a standardized set of geographic, product or customer data. Fully managed, master data becomes the *system of record* for these entities, that is, the “gold copy” that is used as the final word on each entry. This data is the ultimate version of the truth for reference data.

It is equally important to understand what data is *not* included in our definition of MDM. MDM applications do not track other business transaction (BTx) data associated with their specific business entities such as customer account deposits and withdrawals. This data is managed and distributed by business transaction applications and may be integrated in an operational data store (ODS). The historical versions of BTx data are maintained in business intelligence (BI) components such as the data warehouse and data marts.

It may not always be possible to impose a rule that states all master data must be managed by the MDM system. Some maintenance may have to remain in existing business transaction applications, or in outsourced applications such as front-office CRM, for example. This complicates the IT environment and has implications for both master data quality and accuracy. This topic will be discussed in more detail later in this report.

Business Purposes of MDM

Perhaps it is easier to understand the purposes of MDM by examining what the environment looks like in the absence of MDM. This can be summarized by discussing four major problems:

- 1. Data redundancy:** In the absence of an MDM function, each system, application, and even department within the enterprise collects its own version of key business entities. A good example of this comes from the collection of customer data. Key attributes such as customer name and address information are collected repeatedly throughout the enterprise. Unfortunately, it is rare that this gathering process produces the same or consistent data about customers. This leads to the critical difficulty (aside from the storage costs of it) with such redundant data – poor data quality. According to a report from The Data Warehousing Institute, “Data Quality and the Bottom Line” by Wayne Eckerson, corporations lose more than \$600 billion a year due to poor data quality, and most of that cost can be attributed to redundant and low quality master data. This leads to the second major problem.
- 2. Data inconsistency:** Because of the fractured nature of the master data, enterprises spend enormous resources (time, money, and people) doing a function that leads to minimal business benefit – reconciliation. Determining what a customer’s “real” address or name is does nothing to enhance the corporation’s business revenues. And unfortunately, the reconciliation process must be

frequently repeated because there is no mechanism to capture the data assets garnered from the first or succeeding reconciliations. Now we see the third problem.

3. **Business inefficiency:** Low productivity, inefficient supply chain management, inconsistent customer treatment, customer dissatisfaction and wasted marketing efforts are examples of the types of business inefficiency generated from fractured master data. A customer service representative who must scramble through multiple operational systems to determine the status of a customer is not only inefficient, but also risks causing dissatisfaction or alienation of the customer by being perceived as incompetent. All along the work flows within an enterprise, fractured master data causes massive ineffectiveness and inefficiencies, rendering much of the manual effort useless or wasted.
4. **Business change:** Organizations are constantly changing as new products and services are introduced and withdrawn, companies are acquired and sold, and new technologies appear and reach maturity. These disruptive events cause a constant stream of changes to master data; and without a way to manage these changes, the issues of data redundancy, data inconsistency and business inefficiency are exacerbated.

Without MDM, organizations lack a complete view of their customers, products, and other critical business entities, which can have far-reaching consequences. Typically, the two main objectives for MDM are: improving the productivity of the enterprise and increasing the revenues of the company. The case study analyzed for this research paper had a very definitive business purpose for its MDM project.

Case Study: Honeywell

Honeywell deployed a customer MDM solution to provide better service to customers and to look for new growth opportunities. The company identified situations where having the ability to interactively access and analyze purchase and sales data across its many business units would improve customer satisfaction and increase sales revenues. In addition to strong executive backing, the project also had strong support from sales account teams who saw the benefit of being able to use the system to better manage customer relationships.

Detailed descriptions of the MDM application implemented by Honeywell and an overview of the IBM products used can be found in subsequent sections of this report.

Misconceptions About MDM

Because MDM is a relatively new initiative, there are several differing opinions about what it is and how it should be viewed throughout the enterprise. Here are a few of the misconceptions we have encountered in our research:

- *MDM is a data warehousing or BI project.* Untrue! MDM is neither a BI / data warehousing project nor is it an operational project. It stands on its own in terms of benefiting both types of environments. It may use the operational environment as its source of data, but once the data is cleaned, integrated and loaded into the MDM repository, it becomes the source or the *system of record* for systems requiring such data – including the data warehouse, ODS or data marts.
- *MDM is solely for maintaining data consistency across business transaction applications.* Again untrue. Certainly master data can help improve this consistency, but it has more of a role than simply supplying data consistency. As the system of record, the master data repository becomes the standard source for all systems, applications and environments.
- *MDM is simply another data integration project.* No – MDM projects involve business users and MDM disciplines and policies in addition to data integration technologies and applications. Because of the enterprise nature of the MDM function, all corners of the organization may become involved in the design, deployment and utilization of master data. Data stewardship and data administration functionality may need to be heavily involved to ensure adherence to the enterprise view as well as to resolve the difficult issues surrounding this enterprise view.
- *MDM integrates and manages all enterprise-wide data.* Definitely not true. MDM provides system of record for key business entities only. Other transactional data must be maintained elsewhere, perhaps in an operational system, ODS or mixed workload data warehouse environment.

MASTER DATA MANAGEMENT CONCEPTS & TECHNIQUES

Master Data in a Traditional IT Environment: The Problem

In most corporations today, master data is not maintained in a single MDM environment, but by dispersed line-of-business transaction (BTx) applications, each of which has its own business models, rules and definitions (see Figure 2). The data, business models and rules in these applications often overlap and conflict with each other, and this why obtaining a consistent and accurate view of an organization's operational master data is so difficult.

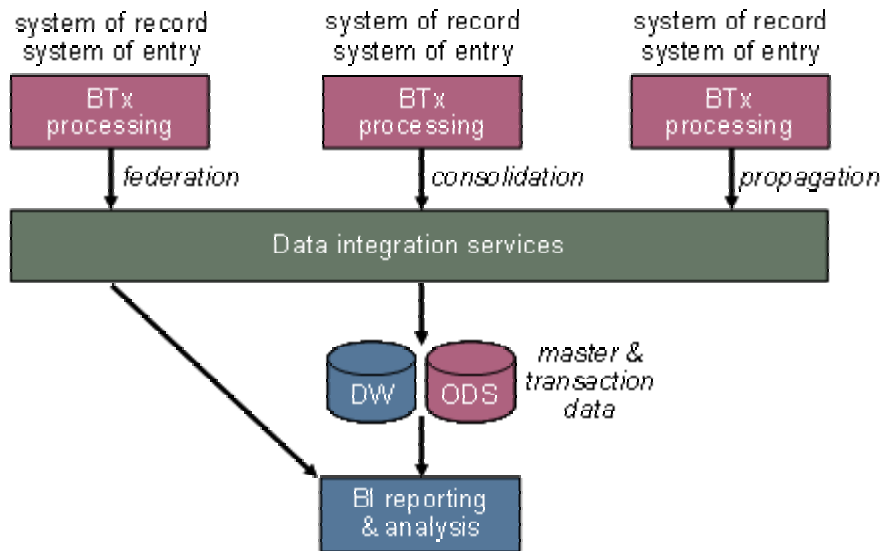


Figure 2: Master data processing in a traditional IT environment

To overcome master data consistency issues in a business intelligence environment, many companies integrate and maintain a historic record of master data and its associated business transaction data in a data warehousing system. There are different ways of doing data warehousing; but, in general, data consolidation and data propagation techniques are used to integrate current master and transaction data in a low-latency operational data store (ODS), and historic detailed and summarized data in an enterprise data warehouse and its underlying data marts.

When it is not possible, for cost or security reasons, to integrate operational transaction data in a data warehouse, a data federation approach is sometimes used to provide a single business view of current dispersed operational data. When this *virtual* data view is referenced by application queries, the data federation software dynamically gathers and integrates the required operational data as each query is run. Data federation can be used to overcome some data quality and consistency issues, but it is not suitable for solving complex data problems.

Building an Integrated MDM Environment

There are many different approaches to integrating and managing master data. Before discussing these approaches in detail, let's first identify the ideal architecture of an MDM system. As we shall see later, this architectural goal is achieved through an iterative and evolutionary application development process.

In an enterprise MDM system, all master data is maintained and published to business users and other IT systems using MDM applications (see Figure 3). These applications handle master data and metadata changes, and maintain a historical record of those

changes. An MDM application could, for example, manage and track customer account data such as account identifiers, customer names and addresses, credit ratings, etc.

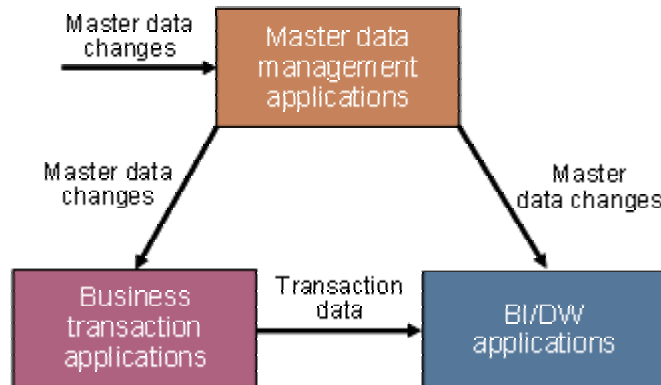


Figure 3: Data flow in an MDM system

The MDM system propagates master data to other internal and external IT systems as required. It also provides business views of master data that can be used by business users and applications to directly access the MDM system. MDM applications do not handle or manage other types of business transaction data such as customer account deposits and withdrawals. This data is managed and distributed by business transaction applications.

Figure 4 shows the main components of an enterprise MDM system. These components include:

- MDM applications for managing and publishing master data and metadata.
- A master data store (MDS) containing consolidated master data.
- A master metadata store (MMS) containing the master data business model, and master data rules and definitions. The master data business model documents master data entities, attributes, relationships and their business meaning.
- A set of master data integration (MDI) services for consolidating, federating and propagating master data.

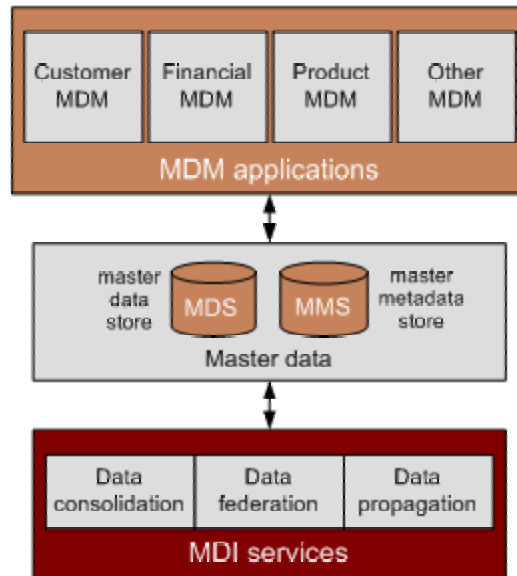


Figure 4: MDM system components

Business users employ custom-built and/or packaged MDM applications to access and maintain master data in the MDS. The MDS represents the *system of record* for enterprise-wide master data. Information about the system of record is documented and maintained in the MMS. As master data is created and maintained, MMS business rules ensure that the master data conforms to the business practices of the organization.

In a fully compliant MDM environment, all master data and metadata is managed by the MDM system. Of course, it may not always be possible to move all master data maintenance operations from existing business transaction applications to the MDM system. It is important, however, that even in situations where some master data is maintained outside of the MDM system, that the MDM system remains the system of record. To explain this point, we need to review in more detail the differences between the master data system of record and the master data system of entry.

System of Record and System of Entry

The system of record (SOR) is the application system responsible for publishing master data and metadata and ensuring its accuracy. The system of entry (SOE) is the application system responsible for creating and maintaining master data and its associated metadata. In a fully compliant MDM system, the SOR and SOE are the same system. Exceptions to full MDM compliance must be agreed upon and documented by IT and business users.

If the SOE is not the MDM system (see Figure 5), then any master data (or metadata) changed by the external SOE must be made visible to the MDM system so that the changes can be published and made available to other IT applications. This can be done

by propagating external SOE master data changes to the MDM system, or by using data federation techniques to make SOE master data visible to the MDM system and its master data business views. These data propagation and data federation facilities are provided by master data integration (MDI) services.

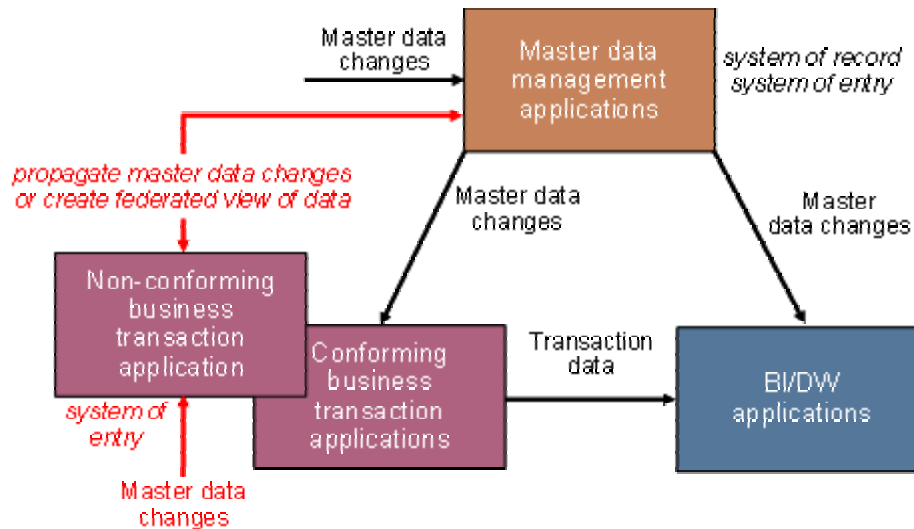


Figure 5: SOE may not always be the MDM system

One of biggest difficulties when the SOE is different from the SOR is maintaining data quality. In an ideal world, the SOE would employ the same data quality and master metadata services as the MDM system. With existing custom-built and packaged business transaction applications, this is unlikely to be the case. It is important whenever possible that MDM and SOE project groups maintain and share a common set of master data governance procedures and business models, and business rules to maximize master data accuracy. An integration competency center can play an important role in managing these procedures, models and rules.

The Role of MDM in the IT Infrastructure

An MDM system and its applications and services are often implemented as tactical extensions to existing business transaction and business intelligence projects. To be successful, however, a strategic MDM initiative should be approached as an independent enterprise-level project with strong executive backing. An MDM system should act as an intelligent source of master data that drives other IT systems. It should not simply consist of a set of adjunct IT applications that gather and integrate existing master data to overcome the problems caused by dispersed master data management.

Bottom-up tactical stealth projects driven by IT groups may be a way to get started in MDM, but an enterprise must develop a strategic MDM plan if it is to be successful in managing master data over the long term by avoiding solutions that simply alleviate the symptoms of master data problems, rather than cure them.

Techniques for Integrating and Managing Master Data

Now that we have reviewed key master data management concepts, we are in a position to discuss different techniques for implementing an MDM system.

There are three main techniques for integrating and managing master data: master data identity registry, master data integration hub and enterprise master data management. Some organizations use a combination of these techniques to build a *hybrid* solution.

The *master data identity registry* (see Figure 6) technique uses an identity management application to create and maintain a repository for interrelating master data across business transaction applications. This repository contains a global identifier that is used to interconnect the master data maintained by different business transaction applications. The global identifier when coupled with data federation software can be used to construct a *virtual* master data SOR.

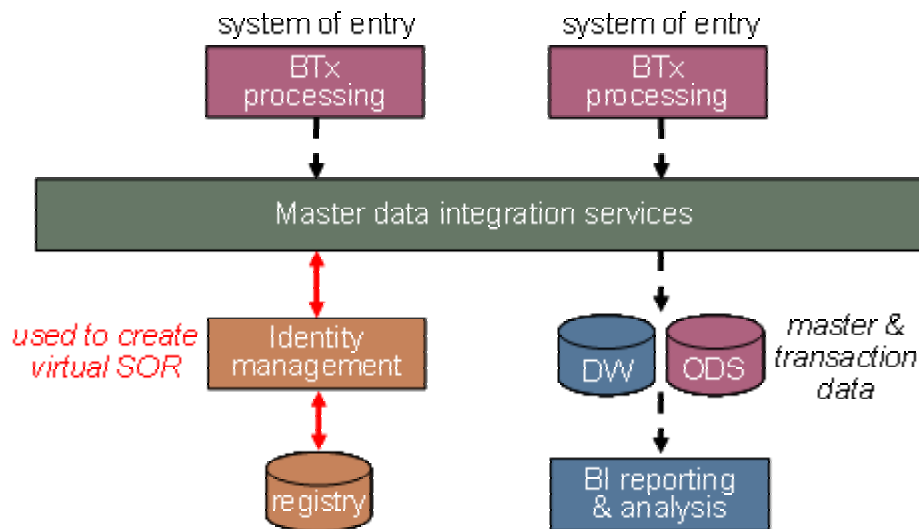


Figure 6: Master data identity registry

A *master data integration hub* (see Figure 7) propagates master data changes between disparate business transaction applications. Although the propagation process is normally done asynchronously, data delivery is guaranteed. Some integration hubs provide the ability to consolidate master data in a master data store (MDS). If all the required master

data is in the MDS, then this acts as the SOR. If only a subset of the master data is in the MDS, then data federation can be used to build a virtual SOR consisting of master data from both the MDS and business transaction applications. An integration hub often employs an associated data model for managing the business semantics of the data flowing through the hub.

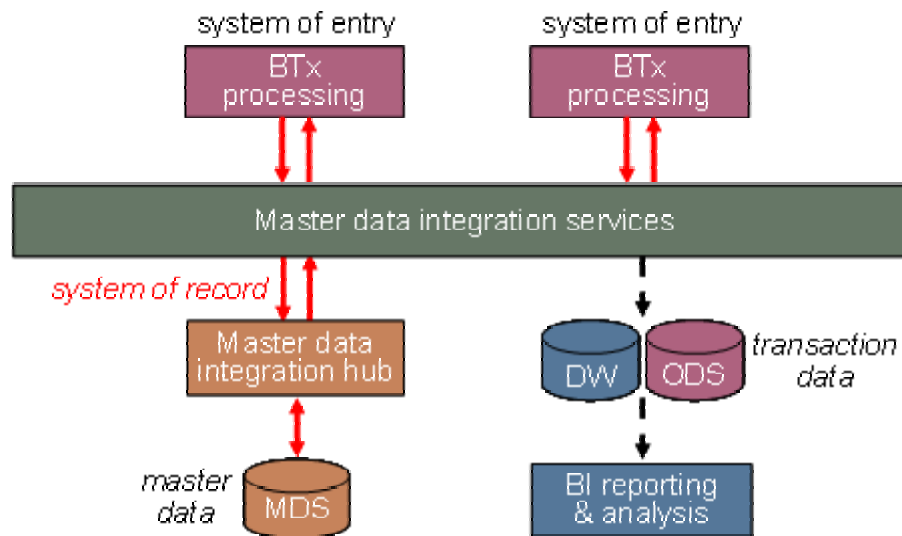


Figure 7: Master data integration hub

Enterprise master data management (see Figure 8) meets all of the requirements of an MDM system outlined earlier in this paper. Except in agreed situations, the MDM system is both the SOE and SOR for master data. When the SOE remains in a business transaction application, data propagation is used to copy data from the application to the MDS – this data is read-only in the MDS, however. Data propagation is also used to copy master data to downstream applications.

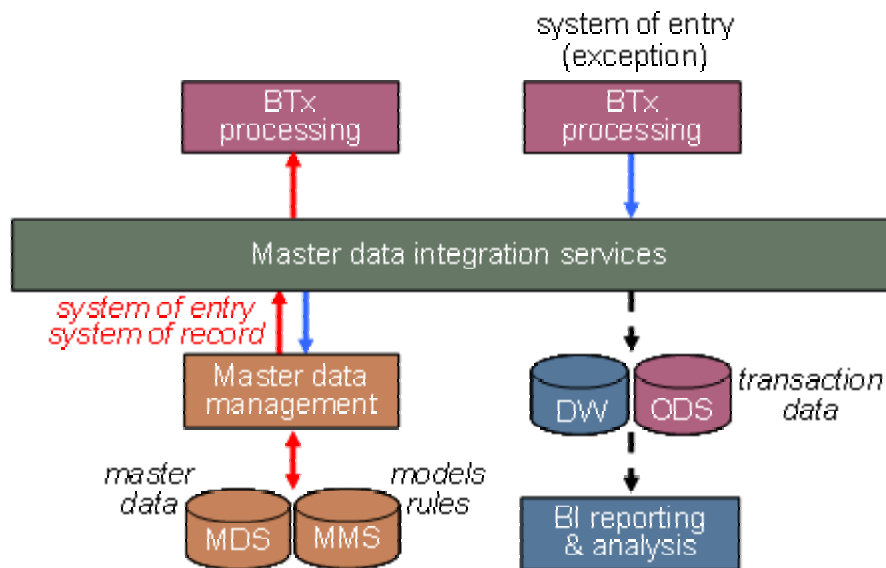


Figure 8: Enterprise MDM

Companies can implement MDM in a phased approach starting with an identity registry, then moving to an integration hub and finally to a full enterprise MDM solution. This evolutionary approach enables the SOR and SOE to be migrated to the MDM system over a period of time.

The Impact of MDM on the Traditional IT Environment

In a traditional IT environment, master data is maintained by multiple disparate business transaction applications. A virtual view of this disparate data can be generated using data federation techniques. Current master data can be also consolidated into an operational data store (ODS) and into a data warehouse for analysis.

Adding a master data identity registry to the traditional IT environment does not affect the way master data flows through the system. The global identifiers in the registry can be used in data federation scenarios to interrelate disparate operational master data. These global identifiers may also be employed in an operational data store and a data warehouse to simplify access to master data.

Building a master data integration hub enhances traditional processing by providing the ability to propagate master data changes between operational business applications. A master data hub also causes changes to ODS design. Master data is stored in the hub, and the remaining transaction data managed in the ODS. Both the master data hub and the ODS can then be used to propagate historical data to a data warehouse.

Enterprise MDM has a dramatic effect on master data flow because it maintains both current and historical detailed master data in a master data store (MDS). This means that master data is no longer maintained in an ODS or a data warehouse. The MDS can, however, be used to supply dimensional data for accessing and processing data warehouse information. The master data history in the MDS may also be employed to restate BI results to obtain valid historical comparisons. An example here would be the ability to compare July 2006 sales with July 2005 sales using the sales territories that existed in July 2005, even though the sales territories were changed in January 2006. This capability is particularly useful for financial reporting.

Enterprise MDM also provides the ability to model master data changes against data warehouse information to predict the impact of those changes on business operations and performance.

Comparing the Three Master Data Techniques

Figure 9 compares, at a high level, some of the key differences between the three MDM techniques. As the figure shows, an integration hub extends an identity registry by adding support for master data integration services, which are used to build and maintain a central master data store and propagate master data between applications. The master data store managed by the hub becomes the system of record.

Enterprise MDM provides additional capabilities in areas such as master data modeling and master data management. It also moves the systems of entry to the MDM system. Full enterprise MDM tracks master data and metadata changes and, unlike the two other approaches, is typically used to handle and relate multiple business entities. These additional capabilities add considerable functionality and business benefit to master data management but, of course, require more resources and development effort to implement. This is why full enterprise MDM should be viewed as a strategic and multiyear initiative.

Requirement	Registry	Hub	EMDM
Identity and Data Quality Management	Y	Y	Y
Master Data Modeling	N	Y	Y+
Master Metadata Store and Management	N	Y	Y+
Supports Master Data Integration Services	N	Y	Y
SOE is MDM System	N	N	Y
SOR is Master Data Store	N	Y	Y
Master Data and Metadata Change History	N	N	Y
Supports Multiple Master Data Entities	N	N	Y

Figure 9: Comparing the three MDM approaches

Business Area MDM versus Enterprise MDM

Many MDM initiatives are targeted at addressing a specific business need such as creating a single view of the customer or a single parts catalog. The issue here is that although these projects are faster and less costly to implement than full enterprise MDM, there is a risk of multiple master data silos being deployed in an organization. This is analogous to the data mart and enterprise data warehouse issues that occur in business intelligence.

A recent blog entry on the Business Intelligence Network had this to say:

"... we are managing hundreds of different categories of master data, only two of which are product and customer. The same issues permeate large corporations, whether the data is HR related, supply chain, asset, brand, etc. Forward-looking companies are taking an integrated approach to the problem rather than a siloed approach."

Although most organizations realize that it is better in the long term to implement an enterprise data warehouse, they nevertheless build independent data marts because they are faster and cheaper to deploy. After building several data marts, companies realize they are creating data silos and then embark on an expensive project to consolidate their data marts. The same situation is beginning to occur with master data. The solution is to plan top down and implement bottom up.

The long-term objective should be to develop a consolidated master data business model and to have integrated master data and metadata stores. All tactical master data project managers should keep the long-term MDM objective in mind when designing and deploying master data applications.

The best approach to balancing the needs of short- and long-term master data requirements is to have a master data practices group that is responsible for helping support the strategic master data objectives of the organization.

MASTER DATA MANAGEMENT REQUIREMENTS

Figure 10 outlines the main capabilities required by an enterprise MDM system. These requirements can be broken down into four main areas: application design, metadata management, data management and integration services.

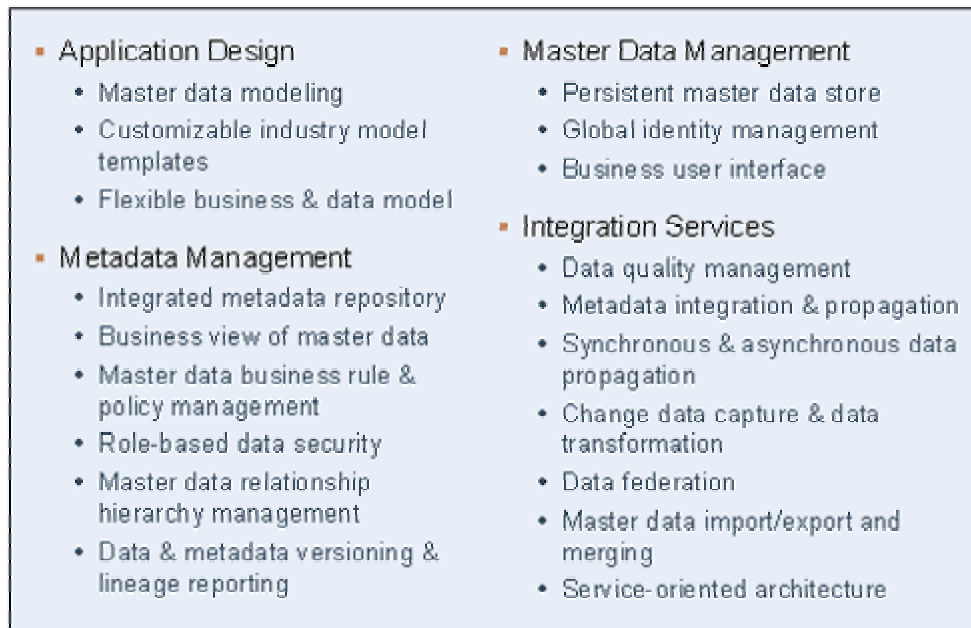


Figure 10: MDM requirements

During design, both IT and business users need an intuitive and flexible modeling facility that allows them to easily document, visualize, define and modify the business model (entities, attributes and relationships) and rules of the MDM application. For organizations with less complex master data, horizontal and vertical industry templates are a valuable starting point for this modeling process.

Master metadata management capabilities should include an integrated metadata repository that documents and handles all of the information associated with the MDM project, including the MDM business model, business views, business rules and policies, and business roles with associated security profiles. Requirements in the data management area include a common master data store and global identity management.

One of the biggest differentiators in vendor MDM solutions is how the metadata and data management facilities manage and track master data, master data relationships and master data hierarchy changes over a period of time, and whether the product provides versioning and master data and metadata lineage reporting. These capabilities are vital to enterprise MDM and key to supporting compliance regulations.

The fourth, and last, group of requirements is concerned with services for integrating MDM applications into the overall enterprise IT environment. These master data integration (MDI) services are discussed in the next section.

Master Data Integration Services

MDM commercial products often provide proprietary integration tools. Unless these tools can be easily incorporated into the existing enterprise integration environment, they should be avoided. Proprietary tools encourage metadata duplication and make it difficult to evolve MDM applications to a full enterprise MDM approach. Key requirements here include:

- Data quality management
- Metadata integration and propagation
- Synchronous and asynchronous data propagation with guaranteed delivery
- Change data capture and data transformation
- Data federation
- Service-oriented architecture

These capabilities form the underlying master data integration (MDI) services (see Figure 11) that support MDM application processing. These services should be a key component of an organization's enterprise integration architecture.

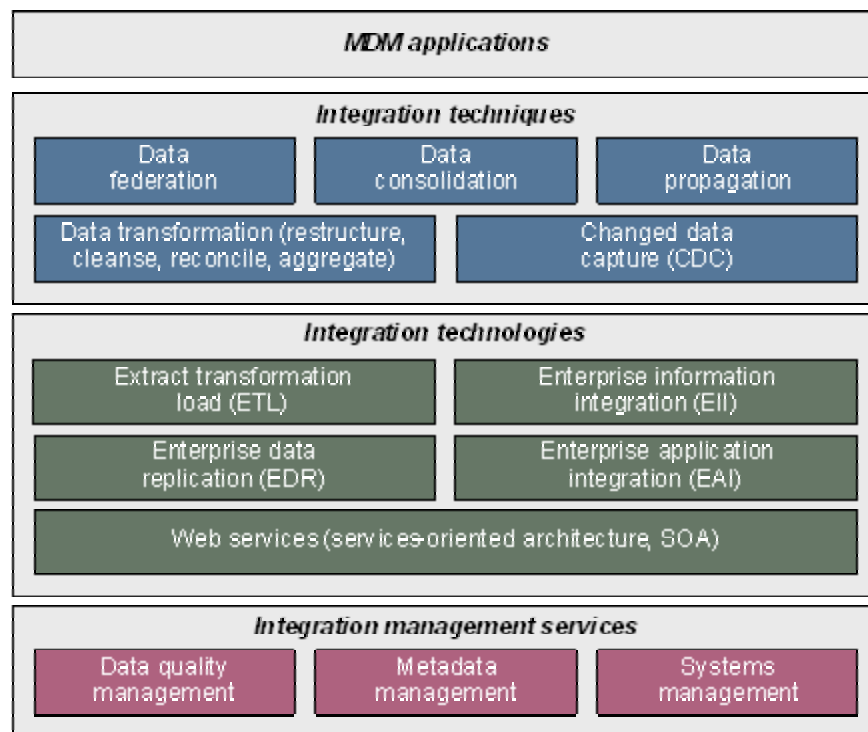


Figure 11: MDI services architecture

Master Data Applications: Build, Buy or Outsource?

To date, most MDM applications have been custom built. An example here is the use of a low-latency operational data store (ODS) to create a single view of customer data. Whereas these custom-built applications may provide the equivalent of a master data registry, or a master data integration hub, they rarely support full enterprise master data management. They are focused primarily on supporting master data integration, rather than master data management.

Like any build versus buy decision, the trade-off is between the cost of in-house development and maintenance versus the cost of vendor licensing and maintenance fees. There is no question, however, that most MDM vendor solutions provide additional capabilities compared with custom-built solutions, especially for enterprise MDM. They are also usually better integrated with the operational business transaction (BTx) packages. This, of course, is particularly the case if the MDM and BTx packages come from the same vendor.

MDM packaged solutions are often specific to a business area (customer, financial or product, for example) and/or industry (retail, banking, telecommunications, healthcare, etc.). For certain business areas and industries, this can be an advantage because built-in business models and templates offer a quick-start to MDM application development. Customer MDM is a good example here. In other areas, such as products, complexity and

lack of industry standardization make the use of predefined templates and models less beneficial. When buying industry-focused solutions, organizations should be careful not to create multiple MDM silos.

Master data is crucial to an organization, and it is highly unlikely that MDM processing would be outsourced. If an organization has outsourced parts of its operational BTx application processing, this can make the implementation of in-house MDM applications difficult. If a CRM system is outsourced, for example, it will be necessary in a customer MDM application to collect data from the outsourced CRM application to create the SOR. It will be difficult, however, to move the SOE from the CRM system to the MDM system, and full enterprise MDM may not be possible.

When building MDM applications, some companies use external information-providers such as Acxiom and D&B to validate and extend their internal corporate master data.

MDM CHALLENGES AND SUCCESS FACTORS

MDM initiatives, like other enterprise-wide ones such as CRM, BPM, BI, data stewardship and so on, have their share of challenges surrounding them – not the least of which is to get cooperation from personnel across multiple departments and lines of business. Following are a few to areas to think about before beginning your first MDM project:

- **Responsibility:** Currently, there is no single business unit responsible for master data in most organizations. Bits and pieces of master data are created and used in multiple lines of business, divisions, departments and even within individual employees' databases. Once captured in an operational system, a business unit responsible for that system washes its hands of further responsibility. It is passed from one application to another, changed, updated and even deleted with little regard to its other instances (upstream or down) in the enterprise. Until a function is created with full responsibility to create master data, an MDM initiative cannot succeed. Unfortunately, this can also be a very political and difficult problem to resolve. High level executives must be involved to resolve the more difficult political storms that may develop.
- **Discipline:** Closely tied to responsibility is the need for the corporation to establish the authority of those responsible for the MDM initiative. This includes defining roles and assigning personnel to the MDM function. The discipline also includes creating formal procedures that must be adhered to by the entire enterprise.

- **Investment:** An MDM initiative requires significant commitment from the enterprise, not just in the assignment of people and their time, but also in terms of overall funding for the technology needed to support and maintain the environment. Both hardware and software will be needed as well as ongoing purchases of MDM applications. While this investment may be small to begin with, it will grow over time. Therefore, it is necessary to be sure the commitment to MDM has full enterprise support at the executive level.
- **Ongoing Effort:** Business processes and master data are constantly changing, and MDM must be seen as not just a one-time project to clean some reference data. The enterprise must understand that an MDM function is, in fact, an ongoing program consisting of multiple coordinated and prioritized projects. There may be several MDM initiatives – one for customer, another for product, etc. – occurring at the same time. These projects must be coordinated so that they share technologies where possible, use standard nomenclature, formats and definitions, and build upon the preceding projects' progress toward the ultimate goal of an MDM environment.
- **Return on Investment:** Finally, for each MDM project, a specific and measurable return on investment should be determined and calculated. While not an easy metric to determine, the overall value of the MDM environment must be demonstrated to the enterprise or future funding may be jeopardized.

These issues can be daunting at first. Therefore, it is necessary for a new MDM initiative to start small and tackle these problems a little at a time. Not all issues can be resolved with the first project, so we recommend that you build upon the success of each project, taking on more of these issues with each successive project.

In addition to these issues, there are also a number of other concerns to be considered as well. For example, security and privacy concerns, particularly around the customer master data, are of great importance. Creating a set of integrated, well-documented, easily accessible master data may be good and bad news for the organization unless proper governance policies and procedures are put in place.

- **MDM Policies:** It must be made clear through documented policies how, when and by whom this critical data is manipulated and used. These policies must include not only what should be done, but also what will happen if a policy is not followed. For example, a policy should be established for determining who owns and makes decisions how master data is used and secured. Exceptions should also be documented, and mitigating actions suggested for these exceptions.

- **MDM Procedures:** Similar to the MDM policies, the procedures or process for creating a piece of MDM data should be documented. For example, what is the procedure for changing the definition of a customer, for adding an element to the customer record or for deleting a customer record from the MDM repository? What should happen when a security breach is detected? Again, exception procedures may have to be set up if a work-around is needed temporarily.

One final consideration that should be discussed involves the rise of data integration centers of excellence or competency centers. No matter what you call them, they have the goals of creating an integrated, enterprise-wide, trusted, sustainable data environment that delivers information to business communities for better decision making. Their purpose is to create repositories of high quality, integrated, current and historical data. This data combines both master data and its corresponding transactional data, which is where these centers may create disjointed projects with an MDM initiative.

There are two ways to resolve this possible conflict. One resolution may be to maintain the MDM function as a separate one from the data integration center. In this case, the MDM function creates and controls the master data within its function and then shares information about this master data with the data integration center. The downside to this is that the two groups may not share technology or enterprise data standards, causing yet another integration problem to occur.

Alternatively, the MDM function could be a sub-process within the overall data integration center, providing the other center functions with the access to information about the master data for their utilization. These functions would share the same technology, standards and even personnel, making the integration of master data into all projects easier.

Depending on your political and funding situation, one or the other of these scenarios may work best for you. By starting small, you can test one organizational structure to see if it is indeed the right one for your situation.

Get Started on an MDM Project

There are some hurdles you may have to overcome when starting your MDM initiative. A few are cultural in nature and may constitute your biggest challenges. The rest are technical ones.

Cultural Hurdles

The first cultural hurdle is generating a business case for MDM. The absence of master data results in many visible difficulties such as the inability to reconcile data across multiple business units, poor information quality, long decision-making cycles, loss of revenue opportunities or the reduction in profits. The disintegration of master data leads to difficulties in identifying customers, improper management of inventory levels, inefficient supply and demand chain execution, inconsistent operational and financial reporting, and poorly run operations. Certainly, without quality master data, regulatory compliance and even mergers and acquisitions become problematic. A sound business case for MDM can be built around any of these business problems. In creating this case, you should focus on the tangible benefits such as improved operations, streamlined IT and business processes, and improved profits. These have real dollars associated with them. The intangible benefits are also useful to mention such as higher customer satisfaction, improved supplier relations and higher quality data.

The second cultural consideration is obtaining and sustaining executive backing. In any enterprise-focused initiative, there are bound to be turf wars, disagreements between divisions or departments, or arguments over the definitions of business entities. An MDM initiative certainly brings all of these conflicts to the fore. An executive steering committee fully committed to the ideals of the MDM initiative should be considered if these problems exist to ensure that the initiative can move forward with full executive support.

The executives themselves may have to be actively involved in dispute resolution and in assuring compliance to the new master data after it is implemented. Overcoming these political barriers may be the most difficult hurdle of all. Moving from a line-of-business view of master data to the enterprise view of it requires that the entire organization accept the MDM repository as the system of record for master data. This means that there must be full agreement or at least acceptance of the common business definitions for these critical business entities. An integration competency center can play an important role in overcoming these barriers. Such a center can also be responsible for defining and managing the governance policies and procedures for master data. A comprehensive governance scheme is an essential MDM success factor.

A third challenge is to ensure sufficient funding to see the initiative through to its completion. Although there are significant benefits to implementing a full scale enterprise-wide MDM system, such a project is an expensive undertaking. MDM should be viewed as a multiyear program that is implemented as a series of incremental projects. In many cases, tactical projects may be needed to meet short-term business needs. The overall long-term objective, however, is to build a full enterprise MDM system, and organizations should plan with this strategic objective in mind, even if they deploy MDM applications bottom up in small incremental steps.

Technical Hurdles

Many of the technology issues associated with MDM have already been discussed. However, a summary of the key technology issues follows.

The first one is the creation of a flexible MDM business and data model. This model is your road map to success. You will not – cannot – know all the master data that your organization will ever need. This need evolves over time, just as the business changes and grows with market shifts and cultural changes. Therefore, the designers responsible for the master data model must take extra care to create a model that is flexible and can accept additions without causing significant perturbations to the rest of the model. Some MDM vendors provide customizable industry templates and data models that can give you a quick start to an MDM project.

The next technical issue involves the management of master data quality. Data quality involves identifying and defining all master data attributes, assessing the quality of those attributes and correcting problems where possible. This can be a significant task, particularly when it comes to reconciling definitions across sources and gaining agreement from the business units where definitions or usage conflict. Data quality activities are done either on all sources of data when the MDM solution is first implemented or on a source-by-source basis as sources are brought into the MDM initiative. These activities are also applied to the data in each source as it changes on an ongoing basis. There are a variety of different data quality and data profiling tools that can aid in this process. For specific types of master data entity, there are specialist tools that can assist with address pattern matching, or with analyzing the business semantics of products and parts documentation.

A third challenge revolves around constantly changing master data. Just as the master data model must remain flexible to accommodate new master data elements, the MDM applications must support the ability to handle complex changes, not only to the content of the data, but also to the hierarchies, relationships, and business rules established within and between the data. Change is inevitable; therefore, when choosing an MDM technology, make sure that it is not too difficult to accommodate changes.

Migrating the system of entry from the operational or line-of-business systems to the MDM system occurs as both a cultural and technical challenge. The technical challenge is to ensure that this transition is smooth in terms of the switchover from operational system to MDM application. As more and more of the system of entry is converted to the MDM environment, each operational system with entry capability for that data must be either altered to eliminate potential dual entry of that data or redirected to the MDM application for entry and update capability. If it is not possible to migrate a system of record to the MDM system, then facilities are required to blend the external master data into the MDM system so that its system of record is kept current.

Identity management is the fifth challenge for MDM initiatives, especially for customer master data. Because customers can be global in nature, have myriad relationships with the enterprise as well as with each other, and are constantly changing these relationships, it is imperative that the MDM environment be able to create universal identifiers that can handle these complex and evolving situations. In many ways, this is a cultural issue as well. The enterprise must agree to and adhere to the new global identifiers and make sure that the MDM function is aware of changes or updates to these identifiers and relationships.

The final technical hurdle involves the overall master data infrastructure itself. While the idea of MDM is not new, the technologies supporting it are. There are a great number of companies that offer partial or incomplete solutions. These may be considered best-of-breed vendors. They have focused a great deal of research and development on a very specific technological aspect of the overall MDM environment such as the ability to improve the quality of customer or part data. While they may have the best solution for a particular aspect of MDM, you will have to ensure that each piece of technology can integrate into the overall MDM infrastructure. Also, because of the number and complexity of the disparate data environments and systems involved in MDM processing, it is crucial that MDM solutions employ common master data integration architecture. This MDI architecture should be consistent with your enterprise data integration strategy and solutions.

Some vendors focus on supplying an end-to-end solution encompassing all or most of the aspects found in a mature MDM environment. In many instances, these companies have merged or acquired the needed technology to complete the MDM technological landscape. While these suite solutions may not have the depth found in the best-of-breed approaches, they do have the advantage of reducing the need to create custom adapters and bridges to integrate products from different vendors. You must determine what works best in your current IT environment.

MDM Success Factors

Enterprise MDM, as has already been pointed out, is a multiyear project that requires strong executive backing and a long-term strategic MDM plan. Success in MDM will come through coordinated incremental line-of-business (LOB) projects. Without a strategic plan and coordinated projects, MDM tactical approaches will result in MDM silos and LOB MDM models.

Sometimes MDM projects will be a part of, or extensions to, enterprise data integration and data warehousing projects. Every integration project team should consider the effect on corporate master data and the overall goal of building an enterprise-wide MDM solution.

Throughout the MDM development cycle, emphasis should be placed on building a consolidated master data business and data model, data quality management and the creation of comprehensive governance policies and procedures. MDM applications must be tightly integrated into the overall enterprise integration framework. The use of a service-oriented architecture will increasingly become a key aspect of this integration framework. A data integration competency center is an important element in helping to support and coordinate all aspects of MDM application development and deployment.

SUMMARY

Implemented correctly, master data management can provide significant business benefits in terms of improving productivity, reducing risk and increasing revenues. Many companies build MDM solutions by deploying a master data integration (MDI) application that is targeted at a specific business issue such as building a single view of the customer, or managing products and parts catalogs. It is important to realize that although MDI can be used to provide a single business view of disparate data, it does not solve the master data problems that exist in most organizations – it simply masks them.

An enterprise MDM solution involves business users in the MDM process, and extends MDI by adding capabilities to manage, track and audit constantly changing master data and metadata. Enterprise MDM is a separate component that manages the system of record and system of entry on behalf of other IT components. It works in conjunction with, and supplies the master data to, business transaction and business intelligence applications. It can also be used to drive the design of the data warehousing environment that forms the underpinning of business intelligence application processing.

Enterprise MDM is a multiyear strategic initiative that can evolve from tactical line-of-business master data registry and master data integration hub projects, provided an enterprise MDM plan is developed to support this evolution.

We would like to conclude with a quote from one vendor executive about MDM because it summarizes the current state of the art in MDM. “While a one size fits all MDM is what everybody expects and wants, the current technologies and solutions in the market are still a few years away from fulfilling that need. The jury is still out on whether a single MDM or multiple MDMs are the best approach to solve the master data problem. The variability and the broad range of key factors such as the type of data (transactional versus analytic), latency of data, level of aggregation, volume of data and the user (business versus IT) perspectives are the fundamental considerations that need to be evaluated before choosing an appropriate MDM solution.”

HONEYWELL CASE STUDY

Company Overview

Honeywell is a technology and manufacturing company providing aerospace products and services, control technologies for buildings, automotive products and specialty materials. Honeywell employs 116,000 people in 100 countries.

The Business Problem

Honeywell has a diverse business in terms of products, customers and go-to market strategies, and it wanted to leverage this diversity to provide better service to customers and to look for new growth opportunities. To meet this need, Honeywell deployed an IT-funded customer master data management (MDM) solution that helped its sales organization cross-sell and up-sell products across the complete Honeywell customer base. “We decided to focus initially on customer master data because it provided visibility to the best service and growth opportunities,” said Cathy Zyetz, program director and business project manager for the MDM project.

Although the customer MDM project focused almost exclusively on the use of integrated master data to drive analytical processing, it has provided some productivity gains in the operational area. “We have been able to identify productivity enhancements that can be realized from data quality improvements,” said Cathy Zyetz. “Customer master data from our ERP and CRM systems is reconciled in the corporate customer data hub and then used to enhance the quality of the data in those operational systems. We are now working on the requirements for extending the existing project to manage the customer identities in our operational application systems.”

Honeywell worked with a consulting group to develop a business case for the customer MDM solution. Cathy Zyetz noted that, “In terms of measuring actual project ROI, we have documented our success stories, but we have not defined these in terms of actual dollars.”

In addition to strong executive backing, the project also had strong support from sales account teams who worked with customers on a day-to-day basis. They saw the benefit of being able to better manage customer relationships. Some middle managers were less supportive. “To develop and demonstrate the business value, we were required to understand the perspective and objectives of the specific business group they are responsible for,” said Jane Booth, Technical IT lead for the MDM project. “We are working on a reporting project that will help the CFOs in the business groups better manage working capital, and this will provide an opportunity to better demonstrate to business unit managers the benefits of the information provided by our customer master data application.”

The MDM Solution

The solution chosen by Honeywell was IBM’s WebSphere Customer Center (WCC). The project team initially preferred a solution from one of its packaged enterprise application suppliers; but in a product *bake-off*, the IBM product offered the best approach for meeting the project’s needs. One of the attractions of WCC was its ability to handle

complex customer hierarchies. “Even though we have fewer customers than large consumer-facing organizations, these customers are large complex companies with many locations and branches, and we needed a product that could handle this complexity,” said Jane Booth. “We liked the flexibility of the WCC data model and the product’s support for Web services. The team members also felt an application agnostic solution might be a better long-term solution.”

Master data from operational systems is loaded in the customer MDM system using batch applications. These applications use a standard XML data format for transferring data. “This approach makes it very easy to switch from a batch environment to an online transaction environment as the project moves toward supporting business transaction processing,” said Jane Booth. “For example, our aerospace group is consolidating several Web applications, and they want to be able to use our system in real time to validate that a person logging onto the Honeywell system is an employee of one of our customers and is authorized to access requested services or products.”

Honeywell uses services from D&B for standardizing name and address data in the WCC application. The source customer data is matched with the DUNS listing for a customer, and the application provides the flexibility for business groups to extend the DUNS definition with their own business relationship hierarchies (name, address, contact information, etc.) and other views as required. However, all of the definitions are linked using the DUNS number as a global customer identifier. To improve the match rate to DUNS records, the project team created a master data governance team to define and implement data quality procedures and tools.

To support inter-business unit master data standards, a governance council, with representation from businesses and functions, was established. Cathy Zyetz noted that, “This is the first time such a group had been formed in Honeywell, and there was strong support for it from the four business units. The objective of the group is to bring an enterprise focus on master data and its stewardship processes in Honeywell. The group is working on standards for customer master data definitions, metrics for data quality and guidelines for management processes. There may always be differences in the handling of customer data in operational systems, but at least we will have a better cross reference capability.”

Lessons Learned and Success Factors

“Our key success factors for the project were to start off with a narrow scope for the project and to use as much out-of-the-box IBM WCC functionality as possible,” said Jane Booth. Cathy Zyetz added, “The project was scoped well in terms of the customer data to be captured. We focused on speed and the ability to produce information as quickly as possible. We gathered data across the complete organization, but we restricted ourselves to a limited set of data.”

“One of our biggest challenges was the complex infrastructure required to support the project. Some four different servers were involved in the implementation,” noted Jane

Booth. “We used not only IBM WCC, which had just been acquired from DWL, but also IBM WebSphere Portal and Websphere Application Server. Tightened systems security requirements added to this complexity. However, we had good support from IBM, and we expect the situation to improve as DWL is integrated with the WebSphere product set.”

Future Plans

“We have had a corporate-level vendor data warehouse for the past five years, and we marry our customer sales information with the vendor purchase information in the data warehouse to produce a complete trading picture,” said Jane Booth. “We have talked about merging the two solutions, but they currently address different business needs and users.”

“We have had some preliminary discussions about product master data management; but because of the diversity of our business units, we have not decided when we would tackle this. There is a lot of interest in extending information about how much we sell to add data about what products we sell. Our product hierarchies, however, have a lot of variation, even within a single business unit, and they are far more complex than even our customer hierarchies. Meanwhile, we are learning a lot from our current project. We want to see how our governance effort evolves. We may be able to apply the results of this effort to the product area before we start a product master data integration project.”

IBM's MASTER DATA MANAGEMENT SOLUTION

“Master data management [MDM] helps organizations manage a complete and accurate master record for critical business entities, such as products, customers, employees, partners and others. MDM helps organizations ensure that data is consistently described, used and stored. With one unified version of data, companies can provide better service to their customers and suppliers, personalize sales and marketing efforts, speed product introductions and streamline their supply chains.

An efficient master data management strategy must take into account a number of different customer or product characteristics such as specific product attributes or product SKU numbers. Since master data is used by different applications, it must also accommodate different use patterns and the data characteristics needed by each application. Such data can be used collaboratively among different applications and might be used for analysis purposes or as operational data. As a result, the master data management solution must manage not only the master data itself, but also how and when the master data is used among applications and processes.”

David Corrigan, WebSphere Customer Center Product Manager
IBM Software Group

Solution Overview

IBM has invested heavily in master data management over the past several years. Within the IBM organization, the master data management (MDM) group has the mandate to leverage all of IBM and partner technology components as well as industry expertise to create master data management solutions. At present, the MDM group offers MDM solutions in three areas:

- **IBM WebSphere Customer Center (WCC)** is a service-oriented business application built on technology acquired from DWL in August 2005. WCC provides functionality for managing customer master data and related business entities. It captures and consolidates multiple sources of customer master data to provide a single integrated store of party records (a party may be a customer, prospect, employee, etc.). Master data is created, maintained and propagated from source and to target systems using a hub-based architecture and a set of some 480+ Java business services. These services cover eleven subject areas including party demographics, roles, party relationships, events and customer insights, customer service and sales, locations, data stewardship, party financial profiles, party identification, accounts and products. WCC's business services are designed for integration within a service-oriented architecture.
- **IBM WebSphere Product Center (WPC)** provides capabilities for managing product, location and supplier master data. It captures and consolidates multiple sources of product master data into a single integrated store. The product can be

used to manage current as well as past product master data. WPC consolidates product information from ERP and other data sources and creates the additional information needed for key business processes such as new product introduction for Web and brick-and-mortar stores, adding new suppliers, price and promotion management, and global data synchronization with trading partners. Consistent and up-to-date product information is then synchronized with multiple customer facing systems such as Web store fronts, kiosks and business portals. Key technology features include an extensible data model, master data versioning, field level security and access control, workflow for modeling business processes and a service-oriented architecture designed for large-scale implementations.

- **IBM Industry Models** provide industry data, process and integration models for the banking (IBM Information Framework), insurance (IBM Insurance Application Architecture), telecommunications and retail industries. These industry models help accelerate strategic business initiatives such as process renewal, risk and compliance, new product introduction, financial management, and business intelligence and data warehousing. The models provide a business and technical *blueprint* expressed in the form of data warehouse, business solutions templates, object models, service definitions and business process models that foster business and IT collaboration. The models are designed to facilitate the deployment of a service-oriented architecture involving MDM, and business and data integration infrastructure platforms.

The MDM team works closely with product groups responsible for IBM's data integration technologies and products. These products provide the underlying master data integration infrastructure for IBM's MDM solutions. The following IBM product areas complement the MDM portfolio:

- **IBM WebSphere Data Integration Suite** (built on Ascential technology) provides functionality for profiling and extracting data from multiple sources of master data, cleansing and transforming the data to a standard format, and consolidating the clean master data into a central master data repository.
- **IBM WebSphere Information Integrator (WebSphere II)** supports the federation, propagation and searching of structured, semi-structured and unstructured master data. IBM WII enables applications to access and integrate diverse structured, semi-structured and unstructured master data via a single virtual view, regardless of where the data resides.
- **IBM WebSphere Business Integration** includes enterprise application integration functionality for propagating and synchronizing master data with a wide array of enterprise applications including SAP, Oracle, PeopleSoft, Siebel and Retek.

- **IBM DB2 Entity Analytics Solution** (formerly SRD) provides a solution for managing and relating master data associated with a person's identity. This technology is being used by various banks, insurance companies, retailers and the federal government for mission-critical applications, ranging from fraud detection to homeland security.

Solution Analysis

IBM's MDM solutions are represented by WebSphere Customer Center (WCC) and WebSphere Product Center (WPC). These products provide the ability to store master data for both products and customers, and the functionality to maintain and access that master data across the enterprise. Product capabilities include a persistent master data store, master data history and versioning, and master data enrichment. These solutions can be connected to existing IT systems using a service-oriented architecture and the capabilities of the IBM WebSphere Data Integration Suite, IBM WebSphere Information Integrator and IBM WebSphere Business Integration. These latter three products support the concept of the master data integration (MDI) infrastructure outlined in this paper.

Appendix A: Definitions

Master data: This is reference data about the core business entities of an organization, such as people (customers, employees and suppliers), things (products, finances/ledgers and assets) and places (locations). Many people think of these as major data subject areas (in enterprise data modeling terms) or as dimensions (in multidimensional analytics terms).

Master data business model: This is a model that documents (in an easy-to-understand format) master data entities, attributes, and relationships and corresponding technical metadata. This model is stored in the master metadata store (MMS).

Master data identity registry, master data integration hub and enterprise MDM: These are the three main approaches to master data management.

Master data integration (MDI) services: These are the various underlying data services and associated technologies used by MDM applications to integrate master data.

Master data management (MDM): A set of disciplines, applications and technologies for harmonizing and managing the system of record and system of entry for the data and metadata associated with the key business entities of an organization.

Master metadata store (MMS): A single repository containing metadata and business rules associated with master data.

MDM applications: These employ master data integration (MDI) services to create and maintain master data.

System of entry: The application system responsible for creating and maintaining master data and its associated metadata.

System of record: The application system responsible for publishing the master copy of any given piece of master data and its metadata.

Appendix B: MDM Survey

As a part of this research project on master data management (MDM), the authors conducted a short survey on the Business Intelligence Network about MDM usage in companies. Of the 60 companies responding to the survey, 30 were planning MDM projects, 7 had implemented an MDM solution and 23 had no MDM plans. Of the 37 companies working on MDM, 26 were focused on product data, 23 on customer data and 9 on financial data. It was interesting to note that the focus on product master data was higher than on customer master data, which we found surprising. The expectation was that customer master data integration would be higher.

One key focus of our report is on the difference between master data integration and master data management. Many MDM projects focus on integration, i.e., they try to solve the symptoms of master data problems, rather than attack the root cause. Management of master data, on the other hand, focuses not only creating a single system of record, but also focuses on the systems of entry where master data problems are created. True enterprise MDM also handles the management of past, present and planned master data hierarchies. Such a system is extremely valuable for driving data warehouse design.

In the survey, we provided two definitions for MDM. One was oriented toward master data integration and technology, whereas the other was oriented toward true enterprise MDM and MDM disciplines. Nineteen respondents opted for the technology definition, whereas 39 chose the enterprise MDM definition.