Managing multiple enterprise data sources is painful, especially when they contain conflicting or incomplete information. Steps must be taken to reconcile the data sources and establish a master data set for reporting and analytics. To gain control over enterprise data, implement a master data management (MDM) strategy to create a single authoritative source.
Executive Summary

Without a clear cut strategy for reconciling disparate data sources and normalizing databases, conflicting data can lead to problems for reporting and analysis, which will inevitably affect decision making. Many IT departments are turning to a master data management (MDM) strategy in order to establish a single version of the truth for enterprise reporting and analytics.

This note provides an overview of MDM. The topics to be covered include:

» What is MDM?
» What is master data?
» Why do enterprises need MDM?
» Types of MDM strategies.
» MDM architectures.

Deeper analysis of the above topics will be followed up in subsequent notes.
Technology Point

As businesses become increasingly dependent on data for management and operational decisions, they must ensure that there is agreement amongst the data sources. Moreover, protocols must be set that define which source(s) will act as the central reference data for reporting and analysis, or steps must be taken to reconcile and normalize disparate databases. The use of multiple transactional systems, such as Enterprise Resource Planning (ERP) and Customer Relationship Management (CRM) applications, has made the task of managing disparate data sources a challenge.

The main issue is the fact that transactional systems are typically set up only to collect and access data within their limited proprietary domain. This causes problems when two systems contain records that disagree about the very same business entities. For example, both ERP and CRM applications make records of customer information for billing purposes and account management respectively. Each system will collect data unique to its own interests but there will also be overlapping data, such as address information. Inconsistencies can arise if, for example, address information is updated in one system (e.g. CRM) but not the other (e.g. ERP). This could lead to the client failing to receive bills on time.

In an effort to combat the problems associated with multiple records, many enterprises are opting to institute a Master Data Management (MDM) strategy. By creating coordinating data sources and/or creating a single authoritative source of high quality enterprise data, organizations can avoid the problems associated with inconsistent and incomplete copies of the same data and ensure greater accuracy.

Two other technology factors that have prompted greater interest in MDM are Software-as-a-Service (SaaS) and Service Oriented Architecture (SOA). The inclusion of SaaS into the Enterprise Architecture (EA) of many organizations has required much more thorough data management practices. This is due to the fact that SaaS requires enterprises to have well organized data sets to take advantage of on demand software services. For similar reasons, adequate SOA projects require the same control over enterprise data.

What It Is & How It Works

What Is Master Data Management?

Depending on enterprise data needs, infrastructure, and suite of applications, actual implementations of MDM will differ from one IT environment to the next. For example, MDM for a BI data warehousing project can require a different approach than a data integration project for ERP and CRM. MDM strategies can also differ with regards to the particular domain of data they focus on (e.g. customer information vs. financial information). For this reason, it is difficult to provide a precise definition of MDM. However, a more general account of MDM can be given.
Broadly speaking, MDM is a multidisciplinary project that uses both business processes and data services technology to cleanse, reconcile, consolidate, and maintain disparate data sources across the organization. More specifically, the goal of MDM is to set up rules and standards for managing enterprise data in order to ensure that applications responsible for reporting and analysis access the most complete, accurate, and up-to-date data available.

Under the most ideal conditions, organizations can create a single authoritative repository of enterprise data, stored in a single physical location, which can be shared with the appropriate applications. However, since most transactional systems are set up to store their data in separate locations, an alternative approach must be taken. A number of standard architectural solutions are summarized below. At minimum, there needs to be a common definition for overlapping data entries and a common set of rules surrounding the data. In the absence of a single physical source, the MDM strategy will need to address which system will be the prime owner of the data and implement a procedure for keeping replicated data harmonized.

**What Is Master Data?**

Master data is a collection of information that describes core business entities – information that can be used to model (or describe) the business entity or provide a particular view of it. All business entities can be classified under one of four categories: people, places, things, and abstract objects. Each category can be further subdivided into domain areas (e.g. customer is a domain area for the category people). Domain areas can themselves be further divided (e.g. customer can be further segmented into individual clients vs. corporate clients) to make distinctions as granular as necessary. However, the degree of detail should be specified on the basis of business information needs – one should not have more categories than are useful for decision makers.
Master data does not necessarily refer to a single source of enterprise data. It is common to conceive of master data in terms of much more narrow categories. For example, in a typical retail company ERP system, it is common for there to be a master-data list corresponding to multiple business entities (e.g. a customer master, a product master, and an account master). The goal of MDM is to ensure coordinate master data across the same categories: for example, the ERP and CRM customer masters.

**Why Is Master Data Management Needed?**

The primary pain points that motivate enterprises to adopt an MDM strategy are the consequences of data redundancy and inconsistencies in duplicate copies of data.

**Redundant Data**

A classic example of redundant data causing problems is when an organization has two copies of address information for the same individual in client lists. Consider an enterprise that has decided to run a mail-out marketing program using the customer data from both the ERP and CRM systems. If steps have not been taken to account for the presence of redundant data across systems, the company will send out multiple copies of the marketing material to the same customer. The net consequence of this is that the enterprise will incur additional costs for every duplicate record and potentially annoy the customer.

**Inconsistent Data**

The problem gets even worse if a company has inconsistent address information across transactional systems. A typical example of this is when a customer moves and a client representative only updates the CRM record. In this case, there are still duplicate copies but only one copy is correct. Since the ERP system will only send bills to the address it has stored, the client will fail to receive bills. The result is that the client could go into default on their account with the potential result that the enterprise loses that client and others due to poor public relations.
Other Reasons to Adopt MDM

In addition to the above pain points, MDM strategies are also commonly implemented for the following purposes:

» **Complying with regulatory constraints.** Some regulatory bodies demand consistent reporting from enterprises (e.g. Sarbanes-Oxley). MDM provides better control over enterprise data to enable consistent reporting for regulatory purposes.

» **Moving to a SOA.** To create a single business service that functions appropriately, it is essential that the information the service accesses is accurate and up to date. For example, if customer information is stored in multiple databases, a decision will have to be made about which will be the record of source.

» **Moving to a SaaS vendor.** SaaS vendors can require a copy of the company database, as is the case when moving to a CRM system like salesforce.com. If there are multiple databases, a decision will have to be made about which one is the most accurate. With an MDM program in place, this task is made much easier.

» **Getting a single picture of an enterprise with multiple lines of business (LOB).** LOB typically have their own distinctive view of core business information, such as customers, accounts, and products. In order to get a clear and comprehensive view of the organization, these different perspectives, realized as disparate data sets, need to be reconciled.

» **Mergers and acquisitions.** The merging or acquiring of another business can cause serious problems with data replication. Each organization that is part of the process will have its own set of master data records. If efforts are not taken to reconcile master data, the enterprise may not be able to leverage the data effectively. There are also additional costs that can be incurred as a result.

Key Considerations

**Three Types of Master Data Management Strategies**

MDM strategies can vary depending over which data sets the enterprise aims to gain better control. According to the Data Warehousing Institute (TDWI), there are three broad categories of MDM programs:

1. **Operational MDM** aims to integrate the data for operational applications such as ERP, CRM, financial reporting, etc.

2. **Analytic MDM** is focused on the creation and support of a data warehouse and business intelligence (BI) platforms.
3. **Enterprise MDM** focuses on creating enterprise-wide strategy that encompasses both operational and analytic MDM.

Each of these approaches has the common goal of resolving data inconsistencies across transactional data sources by creating a single source of valid data. Organizations will typically start with operational or analytic MDM and develop the program over time into an enterprise MDM. Taking a phased approach to implementing an MDM strategy should be considered a best practice.

**Master Data Management Architectures**

MDM can be achieved by both technological means and by changing business processes. On the technology side, there are a large number of vendors that offer middleware enterprise application integration (EAI) solutions for MDM. They are available from platform vendors (e.g. IBM), application vendors (e.g. SAP and Oracle), and data integration vendors (e.g. Informatica). Some enterprises will also opt to develop a homegrown solution. Although EAs can be an important part of an MDM strategy, a detailed discussion of them is beyond the scope of this note.

The most important technological consideration is creation of an MDM architecture that fits with the existing infrastructure. The architectures are essentially the physical implementation of the MDM strategy types, some of which make use of EAI software. The list below gives a quick summary of the main architectural approaches.

- **Application specific master.** This approach is probably better understood as a non-architectural approach. It involves selecting a specific operational application (e.g. CRM or ERP) to act as the primary master data source (system of record). Other applications will then access data from this system.

- **Consolidation architecture.** An MDM hub is deployed as master data store, where changes are consolidated and maintained. This creates a single authoritative, albeit historical, record of master data that is updated regularly. The centralization of master data makes it much more accessible to analytical applications like BI. Analytical MDM is provided by most large multiplatform BI vendors (e.g. IBM Cognos, Oracle-Hyperion, SAS) but also by niche products (e.g. Kalido).

- **Registry-based architecture.** The MDM system contains a registry with sufficient information to be able to identify master data records when queried. The registry also contains cross-reference information that enables access to more detailed information regarding core data elements. Using this information, it acts as a virtual record for master data. It is a read-only system of reference for other applications.

- **Coexistence architecture (synchronization).** Using enterprise application integration (EAI) software (e.g. SAP, Siperian, Tibco, Purisma), application specific master data sets are synchronized with one another in real time, while remaining in their native storage location. This creates a logical enterprise master data source, rather than a physical one.
» **Transactional hub architecture.** Application specific master data sets are replaced by a neutral third-party application that creates a system of record (e.g. a read/write operational master file). This acts as the current system of record for all applications and can be used to resolve conflicts and reconcile and merge master data records.

A detailed evaluation of these architectures will follow in a note focusing on MDM architectures in more detail. It will outline criteria for deciding which architecture is most suitable for each organization.

**Key Takeaways**

1. **MDM is part technology, part process.** Buying data integration software in and of itself facilitates MDM. Decisions about the definitions for core business entities and defining rules for the storage and updating of data must be made. This will necessarily be an enterprise-wide multidisciplinary project.

2. **Failing to coordinate disparate data sources is an IT problem and a business problem.** Without consistent data, operational and strategic decision making can be negatively affected with potential tangible costs to the organization.

3. **Look at MDM applications to streamline data cleansing, reconciliation, and integration.** MDM applications are designed to help automate the technical process of creating a master data set. Using a commercial-off-the-shelf product can be cheaper on the labor front than building an application in-house.

**Bottom Line**

Managing multiple enterprise data sources is painful, especially when they contain conflicting or incomplete information. Steps must be taken to reconcile the data sources and establish a master data set for reporting and analytics. To gain control over enterprise data, implement an MDM strategy to create a single authoritative source.

Info-Tech’s products and services combine actionable insight and relevant advice with ready-to-use tools and templates that cover the full spectrum of IT concerns. Our practical approach is designed to have a clear and measurable positive impact on your organization’s bottom line.

We serve over 21,000 clients at 8,000 organizations around the world. Since 1998, we have focused on making the work of IT professionals easier - and on helping them achieve greater personal and corporate success.

[More About Info-Tech](#)