Master Data Management from a Business Perspective

Writer: Tyler Graham

Technical Reviewer: Roger Wolter

Published: April 2010

Applies to: SQL Server 2008 R2 Master Data Services

Summary: This article helps business stakeholders understand how to determine the master data management needs of their organization. After reading this article, business stakeholders should be empowered to define a high level vision for implementing a master data management solution within their organization.

This article is part of a series called "Organizational Approaches to Master Data Management." For more articles in this series, see http://go.microsoft.com/fwlink/?LinkId=187888.

Introduction

As any business grows, the management of master data is a critical driver to the organization's success. Every acquisition brings in new sets of master data to be merged with the existing business. Although master data management is important to organizations of all sizes, the size of an organization plays a crucial role in how to approach the implementation.

Organization Size	Characteristics	Central Challenge
Small	Small amounts of master data. Data integration is not a top priority.	Creating a plan that will scale with the business.
Mid-size	Data integration starts to become difficult for an organization. Data stewards can be clearly defined.	Implementing effective controls and data stewards.
Large	Huge amounts of master data and system integration. Mostly homogeneous data silos with relatively consistent attributes. Data stewards may now have a full time role.	Building consensus among large number of stakeholders, managing numerous integration points.
Conglomerate	Many disparate businesses that may create many groups of data (i.e., multiple product lines, general ledgers, and so on).	Determining the level at which master data lives.

Table 1: Challenges of different size organizations

Many small businesses do not consider their master data a problem to be concerned with. After all, the spreadsheets they currently use to manage their Product List work great and the accounting system is the only location that the chart of accounts needs to exist in. In my experience this is the easiest and cheapest time to handle the master data management problems. The number of stakeholders for each dataset is small. The number of systems that rely on each dataset is very small. This is the time to implement a master data management strategy that can grow with the business. Each new system implementation will benefit from the single source of all of master data.

Mid-size organizations have a number of dependent systems for each set of master data, so system integration starts to become important. The number of stakeholders for each silo of master data is relatively small. These groups may still work in small teams efficiently. Effective master data controls and an owner for each of the master data entities must be defined. These owners, usually called data stewards, are responsible for managing their domains as new systems are integrated into the organization.

Large organizations have a number of challenges when implementing a comprehensive master data management solution. They are large enough that there are several stakeholders for each silo of master data. Many systems rely on these same models of master data. At this size, coordination of data is a central concern and requires the input of many different stakeholders.

Conglomerates are the most complex master data management challenges. While they may be smaller in overall size in term of employees, assets, or revenue than their large organization brethren, the distinguishing characteristic of these organizations is the breadth of their products offerings or the

diversity of their businesses. Typically these organizations have a diverse offering that makes tracking their master data very challenging. With a significantly diverse product offering, being mindful of how the different businesses interact is extremely important. Also, many industries have specific regulatory hurdles with regards to customers and products that may not be readily known by the organization as a whole.

Why current systems are ineffective master data applications

In many applications, especially ERP systems, master data is created and stored as a requirement of these process systems. Some companies may even call the teams that manage the central data for the ERP systems the master data management group. Although this group is a great place to start to source the new roles of true master data management, these systems do not provide many of the features required to properly manage master data for the entire organization.

Some common limitations of these master data management strategies are:

- Limited ability to version this master data
- Inefficient methods of exporting this data into other applications
- Master data is specific to the functions of the system that manages it and doesn't readily satisfy the requirements of other applications that need to consume it
- Inability to properly store hierarchies or change hierarchies as business requirements change
- Limited or no ability to model relationships between different data groups

Limited ability to version master data

Functional systems require master data to run their specific operations for the organization. Their chief consideration is the most current general ledger, cost centers, organizational entities, and products. Companies spend thousands of hours and hundreds of thousands of dollars to reorganize their sales team. Invariably, a large portion of this time and money is spent mapping the old business units to the new structure.

Inefficient methods of exporting data to other applications

Large, business-wide applications are heavily customized for each organization. These systems provide limited ability to transfer data out of the master data systems. Most systems have some export mechanism that resembles a query language with output of text files. The ability to transform this data with system tools during the export process is very limited if it exists at all. It is also very difficult to export changes within a specified period of time. Due to the lack of versioning, it is unlikely that master data transactions will be available.

Master data is skewed to the functions of the system it is in

These systems have been customized to provide tailored processes to the organization. In the process of customizing these systems, many of the strategies used to customize these processes revolve around making modifications to the master data stored. These changes may work well for their intended

4

function, but as we will see in the next section, storing data in a function-dependent manner makes it less usable to the rest of the organization's systems.

Inability to properly store hierarchies

Some ERP solutions tout the ability to store master data hierarchically. In actuality this is usually managed by placing multiple identifiers into an attached attribute. By giving each character in this attribute special meaning, a surrogate-derived hierarchy can be formed in any subscribing reporting engines. This is a messy solution that tends to scale poorly. As a company grows, each of these character sets can become overextended, creating complex interim solutions. Changing the hierarchy requires changing the identifiers of all records, which can be prohibitively expensive.

Proper hierarchy storage should allow for both derived-data hierarchy relationships and arbitrary parent-child relationships.

Limited or no ability to model relationships between different data groups

Many solutions are not designed to allow relationships to be made between two disparate data groups. Managing products per customer or products per salesperson can be difficult, if not impossible, as the systems may be working with a small subset of the overall corporate data set.

Master data solutions must be process-agnostic

Large ERP systems are designed to manage all of the master data tailored for their system's needs. In this regard they are highly effective, but master data needs to be stored separately from the processes that use the data. As systems evolve over time, one of the easiest ways to modify complex system functionality is to modify the data to solve the problems. Many systems will have multiple customer IDs that map to the same customer to meet some custom reporting needs.

Another issue with storing master data in a process-oriented system is the need to store transactional history. As transactions are created, each is tagged with a combination of account, customer, product, cost center, and so on. These tags must be kept to maintain referential integrity in the system. These histories are like shackles to your master data, requiring multiple custom fields to maintain open and closed statuses.

Master data systems should be agnostic to the uses of this data. This approach keeps these records clean of any attempt to circumvent the programming of a production system. By eliminating the need to maintain ancient accounts for transactional history, master data management systems can provide clean representations of each master data set.

Different methods of implementing master data management

A master data management solution can have many different looks. It is rare to see a large organization implement a corporate-wide master data management solution in one project. Most of these projects grow organically as different groups associated with the project spread the word of the cost and time

savings that master data management enables. There are a number of different factors that contribute to the style of implementation that is chosen. Some of these factors are:

- Level of the organization behind this initiative
- Structure of the current organization
- Structure of the current functional systems
- Complexity of the systems to be integrated
- Size of the organization
- Degree of internal pain attributed to master data issues

Level of the organization behind the initiative

A true enterprise-wide master data management implementation requires the highest level of an organization to underwrite the project. If data integration pains are felt at a lower functional level of the organization, single-dimension master data management solutions are a good fit. Once success is achieved for this one area, more centralized support for expanding the implementation may be found.

Size of the organization

How many people within the organization are dependent on this master data set? How many records need to be stored for each data set? These questions help an organization to determine the type of solution to implement. Once an organization reaches a certain size, implementing an enterprise-wide master data management solution in one project becomes unfeasible. A phased approach may be more prudent.

Structure of the current organization

Is the company large and centrally located? Will multiple organizational units need to be synchronized? Will the internal corporate culture create drag on the implementation of any new solution? As much as size matters, the current structure of an organization matters more.

Structure of the current functional system

When evaluating each system to integrate into a master data management solution, a number of structural elements can affect the decision-making process. Are all customers' records reflected in the system to be integrated? Do multiple records for the same customer provide some functional benefit to the system? Can these customer records be aggregated in the master data management solution? Freeing master data from process systems can allow for better data quality.

Complexity of the systems to be integrated

It is important to evaluate the complexity of each system to be integrated. How critical is the system to the business? What are the best methods to import/export master data from the system? Will the master data stored have a high correlation factor to other systems within the organization?

Degree of internal pain attributed to master data issues

As is the case with any IT project, how large a problem the current process creates for the organization is directly related to the amount of resources that are brought to bear on alleviating the issue. Without

the financial incentive to optimize master data management, many organizations will choose less costly methods of data integration.

Master data management applications

There are a number of different techniques for implementing a master data management solution for an organization. Many organizations begin master data management projects based on data integration issues associated with a dimension of their organization. During the initial research phases of master data management solutions, targeted solutions designed around their specific need will be very appealing. At first blush, these solutions will alleviate much of the data integration problems experienced by the organization. Two highly successful types of targeted master data management solutions are customer data integration (CDI) and product information management (PIM) solutions.

Customer data integration solutions center around the integration of an organization's customer data. These solutions are highly integrated with a company's CRM and ERP systems. Due to the nature of combining so many disparate sources of customer data, match merge and duplicate removal algorithms are critical areas of data integration within the CDI subtype. As this name implies, most of these solutions are focused on integration and require additional internal processes to maintain these systems.

Product information management solutions provide an organization with product-centric management. These solutions typically focus on managing, correlating, and merging product data as bills of materials and online catalogs.

Due to the limited scope of these solutions, many organizations can implement them in a relatively short time. The limited scope also keeps the number of stakeholders that must reach a consensus to a minimum. Quick wins and a narrow scope cause many companies to implement these solutions, ignoring the serious limitations these solutions provide from the perspective of organizational master data management.

Neither of these solutions is designed to be applied to all of the data sets within the organization. An organization can implement a number of separate solutions to create coverage of all their possible master data sets. Implementing these multiple solutions reduces the number of possible integration points but maintains data silos. Cross-dimensional relationships are difficult, if not impossible, to manage.

Working with IT

It is critical for business owners taking on the challenge of master data management to work well with the organization's Information Technology group. While many of the master data management tasks required for sustained success rely heavily on the business users themselves, management of the technologies associated with the data integration routines will need to work well with the IT current processes.

7

Conclusion

Many of the current data management solutions that have been implemented by organizations have grown organically and are unable to meet all of their requirements. A true master data management solution must be process-agnostic and customizable to the needs of the organization. With time and consideration, business stakeholders can determine how a master data management solution will help their business and determine which domains should be integrated first.

Feedback

Did this paper help you? Please give us your feedback. Tell us on a scale of 1 (poor) to 5 (excellent), how would you rate this paper and why have you given it this rating? For example:

- Are you rating it high due to having good examples, excellent screen shots, dear writing, or another reason?
- Are you rating it low due to poor examples, fuzzy screen shots, or unclear writing?

This feedback will help us improve the quality of white papers we release. [Send feedback](#).