Global Headquarters: 5 Speen Street Framingham, MA 01701 USA P.508.872.8200 F.508.935.4015 www.idc.com

WHITE PAPER

Master Data Management: One Step at a Time

Sponsored by: IBM

Henry D. Morris Carl W. Olofson

February 2006

IN THIS WHITE PAPER

Master data management is a set of disciplines and strategies that can be combined with information management products and services in order to provide a single view of customers, products, or other business entities. Improved information governance for managing this shared data is required to achieve more profitable business processes both within and outside of the enterprise.

This paper considers the business drivers for the three major types of master data management projects of increasing scope and complexity. It recommends guidelines for organizations seeking to develop a coordinated response to this imperative for information governance. What is most effective is to approach master data management one step at a time, beginning with one type of master data, one division of the business, or one set of related applications. Identify the key processes where master data is updated, assign clear roles and responsibilities across business and IT, establish metrics to evaluate progress, refine, and improve.

SITUATION OVERVIEW

Industry Drivers, Definitions, and Market Assessment

The Information Governance Imperative

Customers, products, suppliers, and employees are key assets of an organization. Yet data about these business entities is not being maintained in a coherent and consistent way, via policies that define responsibilities for data ownership. Can any group or any region introduce a new product without a review to establish agreement? Who (and likewise, which process or application) is responsible for updating critical information about a customer, such as a credit score or a likelihood to churn? These questions are not about pure technology. They concern the alignment of individual responsibilities for the integrity of corporate information with the data models and applications that manage those assets. Achieving this alignment requires a re-examination of those business processes (such as new product introduction or account origination) where master data originates and is maintained. Inconsistent or redundant performance of such processes is frequently the root cause of poor data quality.

These are not new issues, but they have a particular urgency today for the following reasons:

- **Mergers and acquisitions.** The increased rate of M&A activity puts a strain on the systems that run the business and measure its performance. Organizations seek to eliminate redundant administrative systems to realize cost savings. But this entails careful mapping of financial account, customer, and product data across the separate applications in order to enable the system consolidation.

- **Multiple ERP instances.** In the 1990s the move to ERP represented an effort to consolidate many separate applications into a unified suite through which master customer, product, and account data were kept and updated via a single underlying database. But over time, these organizations established additional instances of the same ERP suite, either for geographic reasons or for implementing extensions to the core applications. With multiple suites, the strategy of master data integration via a single database was invalidated, and new approaches would need to be found.

- **Compliance.** New regulations such as Sarbanes-Oxley and Basel II require an enterprise view of business entities to enable the mandated rollup of information. And alternative rollup schemes may be needed for different types of management reporting. The management of these hierarchies and dimensions poses a special challenge.

- **Service oriented architecture (SOA).** New architectures recommend composing and recomposing applications based on underlying services making use of emerging Web and industry standards. This promises flexibility to adapt systems in response to business requirements. But a service-oriented architecture cannot work unless there is a common understanding of the data that must be shared across the separate components of such composite applications. This shared data is the master data, and, therefore, attention to master data management is a key prerequisite for moving to SOA.

For all of these reasons, organizations must have the ability to make a commitment to develop and enforce information governance procedures for sharing master data within an organization and among enterprises. These issues play out in different ways across a range of industries.

Industry Drivers for Master Data Management

For public companies that maintain enterprise reference data in multiple databases and other stores, master data management should serve as the key component of any strategy for information governance, which is a key component in ensuring compliance with Sarbanes-Oxley and other corporate governance regulations. In addition, specific industries and business areas would clearly benefit from master data management, including:

- The healthcare industry, especially in relation to patient information management (patient master data)

- The pharmaceutical industry, in tracking distributed drug facts including dosages, interaction facts, and contraindications (drug master data)
- The financial services industry, in managing investor and securities information (investor and securities master data)
- Manufacturing, in the management and distribution of new product and product update information (product master data)
- Insurance, in the management of policy holder and risk type (policy type and risk type master data)

The different ways in which master data management applies in these industries illustrate the value of industry-specific expertise and data models for success in master data management implementation.

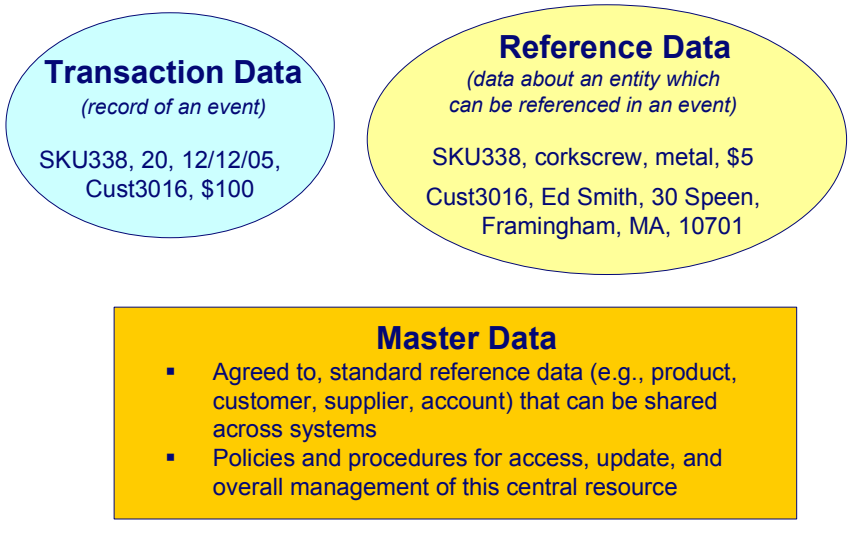
Let's look more carefully at what master data is and how it can be managed.

What Is Master Data Management?

To explain master data management, one must first distinguish reference data, transaction data, and master data (see Figure 1).

FIGURE 1

Reference Data, Transaction Data, and Master Data Distinctions



Source: IDC, 2006

Master Data Management Definitions

Reference data is data for key entities that represent persistent objects that are tracked in the data model, such as people, assets, places, etc. A principal characteristic of reference data entities is that they are referred to by other data entities that depend on the reference data for their definition, typically via a key (in the example illustrated in Figure 1, a sales record refers to a product by its product ID of SKU338; the product is master data). Reference data entities, however, do not depend on other data entities for their definitions. In business terms, the same reference data participates by reference in a variety of business activities, yet is not defined by them. For instance, customer data, which is reference data, participates in sales, service, delivery, billing, and payment activities, but is not defined by them, though the outcomes of these activities might affect the attributes of some customers (for instance, a customer who is chronically late in paying invoices may be so flagged). Examples of reference data entities include customers, products, employees, suppliers, etc.

Transactional data, by contrast, is data that exists to record the state of a transaction or business activity. It is created at the beginning of a business flow and destroyed or archived at the end of a business flow (although it may be preserved in detail or summary form in a data warehouse or mart for business intelligence purposes). Transactional data entities are always dependent on one or more reference data entities for their definitions (for instance, orders depend on customers and products for their definition; in Figure 1, the sales record illustrated is an example of transactional data). Examples of transactional data entities include orders, help desk trouble tickets, payments, shipments, etc.

Reference data is maintained discretely (i.e., one at a time or in select groups), tends to change infrequently, and in contrast to transactional data, does not have a fixed life cycle (meaning that unlike transaction data, it is not data that is created in relation to a specific transaction sequence or business flow and destroyed or archived at the end). Rather, it is maintained indefinitely and archived or destroyed only when it has become irrelevant, obsolete, or unused.

Master data is reference data that is shared across multiple systems and for which there is a common or consensus view. This requirement may arise from an enterprise's information governance rules; overarching business governance rules; or legally imposed constraints, such as statutes and regulations regarding the availability, protection, management, and reporting of business information.

Master data management is the set of processes to create and maintain such a single view through a physical or logical hub, including policies and procedures for access, update, and overall management of this central resource and its coordination with other participating systems across the enterprise or, in some cases, across enterprises.

The Market for Master Data Management

The market for master data management includes software to manage this shared data and the services to enable organizations to successfully implement a program of information governance.

Master Data Management Software

Older terms (such as product information management, customer data integration, and financial consolidation) reflected efforts to provide a single view of one type of master data. IDC has been tracking these **applied MDM** software markets for some time, as software applications came on the market in the 1980s and 1990s that could be sold to individuals with functional responsibility for specific types of shared data.

What is new is the emergence of **MDM infrastructure** software that is purpose-built to support a master data management process for any and all types of master data, and the relationships among these domains. This will be the fastest-growing category in the MDM marketplace, as organizations recognize that information governance requires policies and procedures across master data categories.

IDC forecasts master data management software revenue (including both applied MDM and MDM infrastructure categories) to grow from $2.6 billion to $4.6 billion over the forecast period (2004-2009), a compound annual growth rate of 11.8%.

Applied MDM software includes the following categories of purpose-built applications for managing specific classes of master data:

- **Product.** Software used to coordinate, manage, and share product data throughout the product life cycle.
- **Customer.** Software used in the definition and maintenance of a single view of the customer across multiple systems, otherwise known as customer data integration.
- **Location.** Software used in the definition and maintenance of base/reference geospatial data for entities that need to be shared across multiple systems. These include objects such as streets, buildings, rivers, legal/administrative constructs like state boundaries, oil fields, or stores.
- **Financial accounts.** Specialized software in support of a consolidation process, developing and maintaining a single, unified chart of accounts mapped to multiple existing financial systems for integrated reporting and performance management.

MDM infrastructure software supports the processes for establishing and maintaining a policy hub for master data. The primary domain concerns structured data managed by database management systems. But there is a growing awareness of the importance of unifying reference information across collections of documents or content.

- **MDM/data integration.** In the database world, these products may be called "enterprise MDM" or "enterprise dimension managers" reflecting the intent to tackle this issue beyond a single type of master data and beyond a single department, business unit, or application. This category includes generic data integration tools (such as ETL or data quality) when they are primarily used for maintaining a single view. Data integration software also includes enterprise information integration (EII) software, which applies data movement, data mapping, and data connectivity software to the problem of providing a virtualized or federated data access and management environment for a set of key enterprise data sources such as operational application databases. Master data management can play a key role in an EII strategy by providing a means of ensuring that key queries will not return inconsistent or ambiguous results. Also, the advanced tools and technologies of EII are well suited for MDM, since they can use managed metadata to preserve and extend the mapping of data between application sources and master data hubs even as schemas and processing rules change.

- **MDM/content.** The content world is also intent on unifying its repositories of information, but refers to these software products as "categorizers, controlled vocabularies, tools for taxonomy building, or machine-aided indexing. Master data for content is a set of knowledge frameworks that define the concepts, terminology, or document structures for a given collection of content. The goal is to map content — documents or rich media — to the reference set of established knowledge frameworks in order to bring together similar types of information, no matter how that information is stored or expressed. For example, the taxonomy of legal categories is a knowledge framework to which cases in collections of legal documents are mapped. Metadata for a document or image is derived from the master data, but each metadata instance is not master data itself. Though most of these techniques are being applied to content, there are also promising applications of content technologies (such as fuzzy matching) to improve MDM processes for structured data.

The Role of Metadata in Master Data Management

As may be seen from the descriptions of MDM/data integration and MDM/content software, metadata can play a key role in enabling, accelerating, and managing the building and maintenance of master data management solutions. Metadata, sometimes called "data about data," is really data that describes IT assets, especially data and processes, for the purpose of defining, maintaining, and extending them.

Metadata is used to assign meaning to structured data, and to do so in such a way that it becomes possible to easily determine or discover data in different data sources that mean the same, or similar, things. By applying semantics (structures that indicate meaning) to data, tools can guide, or even automate, such tasks as mapping data from one database to another, or even govern the way in which data is updated in one system based upon the rules that govern corresponding data in other systems. Because metadata can include rules that pertain to correct data formats and usages of data, metadata can be used in a data quality role. Metadata can also be used to determine the impact that a change to data definitions in one database will have on applications and databases to which that data is mapped. Clearly, such functions

have an important role to play in the managing of master data, where mapping the master data to and from application databases is a key function. In general, metadata greatly eases the task of moving or blending data between source databases and a master data hub by applying semantics, or "semanticizing" the data definitions.

Content, especially text, has inherent meaning, including the semantics of language. Yet unlike structured data, it is not governed by a formal structure or schema; rather each element of content is self-defining and self-structured. Although tools exist to discover the structures inherent in pieces of content, organizing them into a schema that can govern structured queries and enable the blending of elements of content with structured data requires (as indicated in the MDM/Content section above) metadata to record that content structure or schema. Once this has been done, then definitive links may be established between elements of content and related structured master data, and automated rules may be applied to ensure consistency and uniformity in their retrieval together. In other words, metadata can be used to build a structure, or schema, by which master data in the form of content may be ordered for the purpose of coordinating its use with corresponding structured master data. This could be referred to as "schematizing" the content.

Taken together, the use of metadata to semanticize structured master data and schematize unstructured or content master data can enable the development of a unified master data management environment that embraces both structured and unstructured master data.

Master Data Management Services

MDM services includes consulting, systems integration, custom development, and training services in support of the establishment or extension of a master data management process. The implementation can either be specific to a type or domain of master data (where applied MDM software is frequently used) or can range across multiple domains of master data. Services include initial business consulting, implementation, and rollout/training. Using recent reports on MDM projects, organizations should expect to pay a higher percentage of their budget on services as compared with software.

IDC forecasts master data management services revenue to grow from $2.8B to $5.7B over the forecast period (2004-2009), a compound annual growth rate of 15.6%. This is a faster growth rate than software, reflecting the increasing complexity and scope of MDM implementations over the forecast period.

Three Scenarios for Master Data Management

There are three different types of projects being done today under the subject of master data management, reflecting a range of organizational objectives. The scenarios are management reporting, data synchronization, and single point of origination.

- **Management reporting.** The objective here is to marshal master data from multiple systems to a common hub for reconciliation and rationalization, enabling the roll up of business data (such as revenue or profit) according to business

dimensions such as customer or product. The business driver is reporting for compliance or for business performance management (BPM).

- **Data synchronization.** In this second scenario, the main objective shifts from reporting to operations. Improvement to operational processes requires a bi-directional flow of data between the master data (i.e., policy) hub and the participating local systems. The local systems continue to update master data records or attributes, causing periods when master data may not be consistent across all systems. Master data consistency is reestablished upon completion of the synchronization process.

- **Single point of origination.** This is the most demanding scenario, where real-time consistency of master data is required for continuous coordination of operations. All changes to master data must originate at the hub, rather than at local systems. This principle requires the redefinition of key processes (e.g., new product introduction). Specific roles and responsibilities must be assigned for the maintenance of master data down to the attribute level.

It is possible to combine elements of multiple scenarios in a single project. Moreover, organizations may begin with one scenario and move to another. The scenarios are listed in order of implementation complexity, but degree of complexity is not a measure of value. The key is to determine the most appropriate scenario in the light of business requirements. Some organizations will only be ready for the first scenario. Over time, it is likely that many organizations will move from the simpler to the more complex scenarios. This represents an incremental or "one step at a time" approach to master data management.

The Management Reporting Scenario

It's not uncommon to hear master data management identified with enterprise dimension management or hierarchy management. This is because reference or master data typically is organized in multiple levels. For example, parts are composed of other parts in the specification of a bill of materials. Employees are organized within business units or profit/loss centers. Customers can be organized by region or by what they buy. It may be necessary to define alternate hierarchies in order to support different types.

Management reporting requires care in the maintenance of these dimensions, whether the reporting is motivated by regulatory compliance or by business performance management. Business changes and reorganizations (of organizational units or products) may require dimension changes, but older versions must be saved for historical reporting.

Data quality practices, so important in data warehousing, are equally important here. But using deduplication and matching techniques to ensure data quality is not always sufficient. It also may be necessary to reconcile and rationalize the master data, as in deciding which products should be retained or whether sales territories should be reassigned.

A key function of the master data or policy hub is to maintain a cross-reference table. Since the flow of master data is from the participating systems to the hub, identifiers

or keys for the same entity may differ across systems. For example, cust-id #1152 in the billing system may refer to the same individual as cust-num #5468 in the call center system. The cross-reference table maintains mapping records that relate the application-specific keys or identifiers for each master data instance, such as a specific customer or product. This cross-reference capability enables the hub to serve as a master data registry, providing a mechanism for linking together all information about a specific person, product, account, or location.

Vertical Industry Perspectives

The management-reporting scenario can be found in virtually any industry. Regulatory compliance is a major driver for this work, whether the regulation applies across all industries or only to one industry. A regulation such as Sarbanes-Oxley applies to public institutions across all industries. But international banks must also comply with Basel II, requiring reporting of the overall operational risk exposure faced by the institution. Hence, Basel II requires bringing together information about customers and products across the various divisions or separate business units that make up a bank. In other words, this is a mandate for a single view of master data to enable this rollup. Given the velocity of mergers and acquisitions in this industry, the requirement to maintain a single view of master data is an ongoing challenge.

Healthcare is another industry where the management reporting scenario can be effectively deployed. Provider institutions such as hospitals have long-standing issues caused by a myriad of separate systems implemented at each department within the institution. There are clinical and administrative systems, applications for radiology, physician's billing, and so forth. It's not feasible to envision a single application or suite emerging to cover all of these functions. Hence, master data management meets a real need. Without the capability to bring together all relevant information about a patient, an institution may make errors that are not only injurious to patients but can lead to financial jeopardy. In the United States, the government is starting to mandate an electronic patient record to enable a single view of all relevant patient information. One step down that road is to create a patient registry, ensuring that patient master data in one system can be associated with patient master data as maintained in another system. Establishing the identity of a patient in order to retrieve all relevant medical history is not only valuable, but can be a matter of life and death.

The Data Synchronization Scenario

Master data management in the first scenario pays attention to the management of dimensions so that data can be rolled up properly for reporting. But what if the primary requirement is not analysis of business data, but coordinating business operations? To impact operations, changes to master data that are approved at the central hub must be sent to the participating operational and analytical systems. This bidirectional flow of master data is the hallmark of the data synchronization scenario.

The goal of this scenario is to synchronize the canonical master data, managed in the hub, with the participating local systems. Though more complex than the first scenario, the data synchronization scenario retains the autonomy of local systems to initiate master data changes.

Vertical Industry Perspectives

Consider the case of a retailer with a long-standing mail order catalog business, a newer online business, and possibly a chain of stores as well. A customer could be added at any of these points. Assume that the retailer wishes to run a monthly cross-selling campaign, requiring the association of customer information across multiple channels. A prerequisite is the ability to identify customers established via one channel with the same customer established via another channel. Is it essential that there be real-time consistency of customer master data across all channels? From the perspective of the monthly cross-selling report, real-time coordination is not required. Daily or weekly synchronization of customer master data would be sufficient.

The Single Point of Origination Scenario

The most rigorous master data management scenario requires that all changes to master data originate from the central hub. This single point of origination strategy may be an effort to reclaim a vision that went along with big ERP implementations. Over the last decade, many global companies replaced separate departmental applications with a single enterprise application suite where master data was maintained by the suite.

Whether or not this vision of an ERP suite enforcing master data consistency ever existed, it became invalidated when multiple instances of ERP suites appeared in the same company. With multiple instances separately maintaining reference data, an integration layer apart from each instance is required to achieve a single view of master data that is consistent at all times.

Moving to a single point of origination of master data is the most challenging MDM strategy, since it requires the most fundamental changes to business process changes of any of the scenarios. Rather than enabling master data changes across an organization, this scenario requires that all master data changes originate at the hub. Existing systems must be adapted so that the workflows leading to an update to master data must be redirected from the local system to the hub. This might be straightforward in a world when all applications are built out of Web services and comply with a services oriented architecture. In such a world, the process steps (implemented via services) could be rearranged easily via a process design tool and coordinated at run time via a process orchestration layer. In the world of today's systems, significant application code changes would be needed to implement the changes to process flow.

Vertical Industry Perspectives

A global CPG company has embarked on a plan to provide a single point of origination to changes in item, vendor, and customer data, beginning with their European operations. This is required not only for operations within the enterprise but for delivering syndicated data to their partners. It is an ambitious effort to reestablish control on master data that was lost in the move from the mainframe to a distributed environment.

Customer and product hierarchies were normalized, leveraging the Master Data Repository (MDR) as the system of record (based on Websphere Product Center). This required establishing a consensus on the definition of customer, vendor, and item (product) across systems. The company uses external standards when available, such as UCCnet — 150 in all for different types of data. Data custodian responsibilities in the organization are assigned down to the attribute level for all attributes that are of more than local interest — these are termed *global attributes.* Master data changes are propagated from the hub to participating systems using MQ Series middleware.

But implementing a repository and setting up a data synchronization environment are not all that were needed. The CPG company's global application architect recognized that implementing a single point of origination scenario requires business process adaptation. In particular, adding or updating master data records or changing any of the global attributes must be driven from the hub. This requires disabling functionality to make these changes in participating systems (even their ERP suite), and redirecting the workflow to do the updates via the hub. The combination of data and process integration achieves a strong degree of data consistency for the shared master data critical to this global organization.

Examining Business Processes and Their Impact on Master Data

The third scenario calls attention to the impact of business processes on master data. Organizations who study the root causes of master data integrity issues recognize that they are caused by a lack of consistency in business processes such as new product introduction or customer update.

For example, a retailer may have redundant, separate processes for introducing new customers via an application in support of their online business versus their catalog business. Or a manufacturer may have uncoordinated product introduction processes at regional sites versus headquarters. Separate processes are usually accompanied by separate applications, different data models, and multiple points of responsibility.

With separate, uncoordinated processes, inconsistent master data is introduced. Transforming and integrating the data in a data warehouse does not remove the inconsistency. If the goal is to coordinate operations across divisions of a company or even across enterprises (e.g., across a supply chain), attention must be given to uncover the business process issues that lead to inconsistent data. Fixing the problem requires attention to reconciling and rationalizing the relevant business processes and establishing clear lines of personal responsibility for master data entries and attributes across an organization.

Which Master Data Management Scenario Is Appropriate?

Here are some points to consider in selecting a master data management scenario:

- Is the goal improved reporting or increased efficiency of operations? If the objective is to provide better reporting, either for compliance or performance management reasons, the management reporting scenario makes sense. If the

goal is also to coordinate operations, such as across a supply chain, the data synchronization or single point of origination scenarios should be considered.

- How often must updates to master data be propagated from the hub to the participating systems? The answer depends on how quickly decisions must be made based on the data. Monthly marketing campaigns don't require near real-time updates. But adjusting workflows on the shop floor is another matter. The single point of origination scenario, though the most complex, provides the capability to get to near real-time synchronization if the business decision cycle warrants that level of consistency.

- Can business processes be altered? The single point of origination scenario requires change of processes, such as new product introduction or customer registration. Is the organization ready to reevaluate how these processes are done and centralize responsibilities for the maintenance of master data attributes that must be shared across the organization.

- What is the degree of heterogeneity of the relevant systems? If there are a large number of separate applications or application instances, it may be worth considering combining or consolidating systems. It may not be possible to move to a single application or instance, but the reduced complexity will ease the task of implementing a master data hub and spoke architecture.

Whichever path you take, anticipate that MDM activity will grow over time, adding more types of master data to the mix, and possibly moving from the relatively less sophisticated reporting scenario to the more difficult data synchronization scenario. Organizations should take great care in considering the third scenario, single point of origin, in that such an approach would *require* making profound changes to *all* related applications with respect to the update of master data, including packaged applications. Yet this scenario, if implemented successfully, provides the highest degree of control and greatest level of consistency of master data.

In addition to the choice of project scenario, you need to consider the type of software to be used. In particular, you'll need to decide whether to build of buy a master data management solution.

What Type of MDM Software Is Appropriate?

Buy

As with most application software decisions, the tendency is to buy and apply already existing packaged applied MDM software for those situations where it makes sense. This option should be examined whenever such master data as product, customer, location, or financial master data is involved. It is not necessarily the case, however, that such MDM software will be able to handle the master data in a way that is consistent with corporate strategy or practice. This must be examined, and the question of how easily the software is modified must be taken into account. It may be that in order to make the package work in a manner consistent with corporate policy or practice, such modifications must be made either by staff or by well-qualified service providers. Are such service providers available? How expensive are they?

Another important consideration is the MDM scenario the enterprise wants to adopt, both now and in the future (management reporting, data synchronization, or single point of origin). Does the proposed package support the plans of the enterprise, whether they are to adopt one of these scenarios for all time, or to adopt one for now, then evolve to the others? If not, the investment in time, money, and effort to install the software and adjust the other business software to work with the package may prove wasted in the longer term.

Another concern is this: it may be desirable, either from an information governance or from an operational or technical management perspective to ensure that all master data is managed the same way. If the packaged applied MDM software for, say customer master data management, is remarkably different from that for financial master data management, it may be necessary to purchase one or the other, and build a solution for the unsupported master data that is consistent with the approach of the package. This dissonance may also be sufficient to drive consideration of the "build" option.

Build

The other option is to build custom MDM solutions on top of MDM infrastructure software. This approach provides maximum flexibility and ensures conformance with corporate policies and strategy, but may prove too expensive if the enterprise lacks personnel qualified to execute the data analysis, modeling, database design, data transformation rules definition, data movement scheduling, and other tasks necessary to carry it out.

Hybrid

A third option is a blend of the first two. This involves using packaged applied MDM software for that master data that is sufficiently standardized that such packages make sense, but integrate them into a MDM framework built for flexibility and in a manner consistent with future plans using MDM infrastructure software. This approach will become more attractive as standards emerge in this area that make the integration of packaged applied MDM modules into a standards-compliant MDM infrastructure practical.

Opportunities and Challenges

Master data management presents an opportunity to establish or reestablish control of the maintenance of data that must be shared across many applications within your enterprise and between your enterprise and your business partners. There are many benefits to this process.

Here are some ways in which MDM can reduce costs and also lead to increased revenue:

- There is a high cost when knowledge workers are unable to find the information necessary to perform their jobs. Master data management enables the rollup of business information according to product hierarchies, organizational groups, and the like. Much time is wasted trying to find information and trying to reconcile

inconsistent information produced by multiple systems. This is frustrating, saps employee productivity, and impedes the monitoring of business activity and performance.

- Complying with regulations such as Sarbanes-Oxley, Basel II, and HIPAA requires taking an enterprise view of business entities such as products and customers.

- Many organizations have service level agreements to provide data to their partners and stakeholders, such as syndicated data about product sales, warranty claims, and the like. Without the proper maintenance of master data, a company is at risk in meeting these agreements.

- Besides cost reduction, master data management can enable new revenue opportunities. The ability to sell products across lines of business requires care for the maintenance of customer and product master data across these units.

But these benefits do not come without significant challenges. The technical challenges may be the first that come to mind. The three scenarios for master data management reveal increasing levels of technical complexity in data and process integration. However, most organizations prepare for master data management by tackling the organizational challenges first.

The root cause for the proliferation of systems for maintaining master data may lie in political divisions within an organization, impeding progress in sharing data. That is why master data management goes hand in hand with information governance. Companies that are achieving success in master data management take the establishment of policies and procedures for information governance very seriously. They are forming master data management committees to understand the problem. And they are assigning responsibility down to the data-attribute level for information that must be shared across applications, across business units, and even across enterprises.

Future Outlook: MDM and SOA

There is a seminal movement underway in IT shops around the world; that movement is the development of a component-oriented approach to application development and deployment known as service oriented architecture (SOA). This will emerge as the dominant application architecture, leading to the development and deployment of composite applications.

SOA involves the deployment of applications not as monolithic programs but as collections of components that perform application services and invoke one another's services through standard message formats and protocols. At present, the code that invokes these services, and the services themselves, are under the control of the same project teams that tend to develop them together. Over time, this will be less true. Over time also, existing applications will become increasingly componentized, exposing their operations as services and enabling application functions that span across existing applications; these "spanning" application functions represent a new class of applications known as "composite applications."

As composite applications develop, a great danger exists that the application functions they join together will be based on separate databases that do not agree, resulting in errant operations and producing puzzling and, in many cases, subtly incorrect outcomes that could cause real damage (operational costs, legal liabilities, etc.) to an enterprise. In order to avoid such an outcome, data integration technologies will need to be brought to bear to synchronize and coordinate the data that underlies the underlying components or services that make up the composite applications. Such a task is greatly eased if key reference data is already uniformly managed under a master data management rubric. Thus, MDM stands to play a critical role in enabling the growth in SOA deployments, generally, and composite applications, in particular.

CONCLUSION

Master data management is gaining greater recognition as a priority for organizations seeking to reestablish control over the data that needs to be shared across the organization. This paper presented several models of projects for master data management. With any of the project scenarios and whether you build or buy the needed software, keep the following in mind as you plan a master data management initiative:

- Recognize that organizational issues of data ownership are a prime cause of master data management problems. Establish clear lines of responsibility down to the attribute level for data that needs to be shared. Ensure that there is a clear mandate from senior management in support of these efforts.
- Pay attention to those business processes where master data is introduced and maintained. Business process reconciliation and organizational alignment are at the heart of information governance.
- Select a business area to focus on where the returns can be measured because enterprise master data management cannot be achieved in a single step. Focus on those processes that need to be standardized in order to ensure consistent master data. Measure your improvement, refine, and expand.
- Partner with a vendor who understands the organizational, technical, and vertical industry dimensions of master data management.