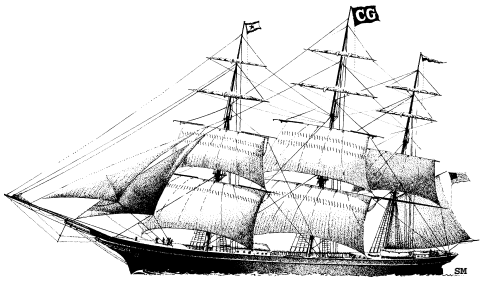
THE CLIPPER GROUP
Navigator™



Navigating Information Technology Horizons℠

Published Since 1993 Report #TCG2008004 January 17, 2008

Sharing Business Data Properly Requires Master Data Management Discipline

Analyst: Anne MacFarland

Management Summary

Business has expanded beyond the days when information was used only internally and was communicated up and down the hierarchy through carefully groomed reports. In a world where communications has shrunk distances and accelerated the pace of business, effective *ad hoc* sharing of trustable, complete and accurate information – not just reports – becomes essential to good business practices. It also can be a really big challenge.

Irreconcilable data exploded as a huge and unmanageable issue in the late 20th Century, when institutional mergers exposed massive dysfunctionalities. Many data quality strategies have emerged since then, but most focus on *aspects of data quality*, not on the particular problems of *data to be shared*. Sharing operational data – even within an organization – is intimate, and can be messy. While other people's data enhances our understanding of markets and opportunities, inconsistent data pollutes databases and make analysis impossible. Better organizational collaboration requires rationalized operational information, be it financial numbers or design specs. It mandates consistent definitions of structured data, disciplined use of both those definitions and common concepts such as customer, product, etc., and a willingness to apply – and extend – these disciplines on an ongoing basis. With these basics, business can be done accurately and consistently. Without them, organizational change is difficult, and collaboration becomes reduced to the least common denominator of understanding – a congenial best effort that is often insufficient.

It is not only the nature of day-to-day operations that demands a better discipline. Innovation these days comes with a caveat of *reusability*. This reusability mandates a well-defined, sharable *operational vocabulary* of entities and categories that is accurate, current, and consistent with the practices of all parties the innovation will touch. This is a language of business – quite different from the traditional social languages in which we do business, but just as important. It is a broader concept than the also-important industry data standards that make data understandable by different programs, processes, and databases, which are focused on format and formulation of key industry artifacts**.** ***Master Data Management* (*MDM*) is all about this language of business *information*, and the support of its coherence and truthfulness that support both business operations and business intelligence.** What needs to be accurately sharable depends on the nature of the business, initiative, or project. It will change over time.

The success of an organization rests on the quality of the truths it can bring to bear on a situation, how well it can analyze them, and how quickly it can come up with a solution. As collaboration gets more outward facing, it is easy for organizational coherence to slip – but very important that it should not. This is a data management problem that must be addressed, as IBM has done, with both software and human effort. For more details on IBM's Master Data Management approach to this problem, read on.

IN THIS ISSUE

What is Master Data?

Data that is useful beyond the application from which it was spawned qualifies to become *Master Data*[1]. It may be useful to other parts of an organization and to external partners. It may be sold or traded as an information asset. It may be aggregated with other data sources to provide a more complete view of a product, or a customer. With this more complete view, decisions on how to move forward become easier, and the tactics to use more apparent. However, if the aggregate data from various applications about a particular customer, product, or situation is inconsistent or incomplete, decisions often must be postponed, and strategy brought to a standstill, until the inconsistency is resolved.

Varieties of Master Data

Master Data can come in many forms. It is the specifications that follow a product from design through production and persist for warranty support throughout the life of the product. It is the information an organization aggregates about its members or customers. In collaborations, it is the key concepts whose definition should not become subject to tactical compromise when the going gets tough. It is contract templates and terms.

As companies grow by acquisitions, each new group brings obvious legacies, in the form of cultural and organizational expectations, and ways of doing business. The less obvious legacies are how costs are managed and cumulated (i.e., what elements comprise administrative costs) and how variable costs (like compensation packages) are derived. Companies based in certain geographies must meet certain local business standards, while still collaborating fully with the new organization of which they are a part. All these factors are challenges that can make Master Data Management an opportunity for defense of fiefdoms, rather than the operational boon that it should be.

Analyzed business data must be consistent with operational data, if the analysis is to have any meaning. This is obvious – but insuring that a full range of operational data is available for analysis is not easy. In an organization of any size and complexity, locating, gathering, processing and preserving master data can be a challenge. Organizations change, customers change, and what we want to know about them

[1] Typical Entities that are a part of Master Data Initiatives are discussed in Exhibit 1, at right.

Exhibit 1 – MDM Entities

The following categories of entities form a framework common to most organizations. Many precursor approaches focused on one or a piece of one (e.g., customer or product data) but all four categories are needed to get information congruent in most organizations. Other categories may be useful in a particular industry. These domains are interrelated and cannot be treated as silos. Segregating information in labeled buckets (repositories, or folders) masks the relationships that transcend categories.

- **Party** is broad term that covers lines of business, customers, partners, suppliers, and other business entities. Any party may have a variety of relationships with other parties. An individual may act either as a party and or as a representative of any of a number of parties. The attributes of parties re enriched by careful harvesting of attributes and duplication is resolved by matching. The need to prevent fraud underlies many efforts that have parties as a focus. Management of parties is made more complex and important due to government regulations.
- **Product** information comes from many sources over the course of its lifecycle. It can include information about parts, bundles, SKUs, and warranties. Pricing and other sales information across geographies may add intentionally conflicting information. The rate of change (due to things such as limited time offers) of this kind of information is very high. Maintaining coherence and validity must be an ongoing effort.
- **Account** information includes contracts, a transaction history, and rewards programs and other special considerations. The completeness of this information is highly important. What constitutes account information, and how it is derived, may set off ownership issues and political squabbles within an organization.
- **Location** information, such as addresses, contact methods and other details must be kept current and used consistently. This is not just for day-to-day use, but also as operational metrics are rolled up into aggregates (regions and geographies).

changes. It is easy to add a new category without finding or defining its relationship to other similar categories already in use, particularly if you are only thinking in terms of an immediate project. Over the long term, these ad-hoc decisions become a problem.

Over the years, many approaches have been used to address this problem. A *system of record* is an authoritative version of business documents, best represented in the days of paper records by the locked file cabinets in the office of the president's secretary. Master Data Management, at its most disciplined incarnation, produces a system of record. However, in a modern, highly-distributed enterprise, such discipline does not come easily. Putting everything "of importance" in a data warehouse or repository becomes a very complex undertaking, for data is interrelated – all business information, by its nature, contributes some value to the whole. The more information that is included, the more time it takes to load, which – these days – degrades its value. Much of the information that is shared within an organization does not fit the categories within a data warehouse, and, were new categories to be developed, the information would be too sparse and irregular to be useful. Finally, putting information in a warehouse does not get it used.

A *hub that transforms data into a usable state* is another classical approach to reconciliation issues – but, like a data warehouse or repository, it comes with limitations. If it is the only path to the truth, unless it is scalable and robust, it can become a bottleneck or single point of failure. Multiple hubs are more robust, but they become another system to be managed.

In a Service Oriented Architecture (SOA), a *service registry* can list and describe information sources, allowing them to become services invoked by applications. For this to be effective, of course, the data must be in a state to be used by all of them. This can be problem. High-quality data, i.e., accurate and useful data, from one application may be unusable by other applications until the inconsistencies are addressed – often by lopping off inconvenient anomalies. This results in data loss, which may affect the quality of the decisions that can be made using it.

How IBM Approaches MDM

IBM has offered Master Data Management products and services for many years and is actively helping over a thousand organizations to develop Master Data Management strategies. In offering *Multiform Master Data Management* as an *Information On Demand Solution*, IBM seeks to offer a full variety of capabilities that are useful to building out a full hybrid MDM environment. It also can work with existing legacies of competing products.[2]

From its experience, IBM has seen the wisdom of giving people various tactical approaches to Master Data Management in addition to the implementation that produces a System of Record that, in highly regulated industries, is part of doing business. IBM notes that, while many well-targeted pilots are very successful, they are often followed by another pilot, rather than a broader initiative. This is unfortunate, for better-managed silos of information is not the philosophical objective of Master Data Management.

Most organizations will opt for the simplest deployment that will meet their immediate needs, knowing that the extensibility of the platform will let them add more functionality when and where it is needed. As they realize the benefits of Master Data Management, they extend it to new areas of the business – but in a style that will meet the needs of that area.

IBM gives them the underlying commonality to pull more into a common whole when there is a need or desire to do so. By meeting the customer on his or her ground, IBM gives more choice with less risk than some of its more deterministic competitors. By supporting and end-state vision as well as deliverables with short-term benefit, IBM offers customers a way to get started with less commitment. The Master Data Management "styles" that IBM offers have enough commonality to play well with each other in a hybrid deployment. This strategy lets their customers get started comfortably, picking areas where data reconciliation is crucial, and picking the style that will suit the situation. Then, the initiative can spread using the paradigm that works best in various organizational situations.

IBM Styles of Master Data Management Adoption

At its simplest, Master Data Management involves information transformation to a more sharable norm. At its most sociologically complex, it involves renegotiations of the definition of naming conventions and accounting categories – institutions that are surprisingly hard to change. Certain kinds of information –

[2] More on multiform MDM in the next section.

customer information subject to privacy constraints and research information subject to industrial espionage – will require particular care, special handling, and is not broadly useful for operations or analysis. It is the run-of-the-mill operational data that will tell you how, exactly, the business is doing – if you can aggregate it accurately and analyze it deftly.

The four approaches discussed below, which correspond to what IBM calls its MDM *styles*, give you ways to approach making your huge store of operational data more coherent. The advantage of the first three is that they are less intrusive to implement and give benefits usually in a few months. The fourth offers full benefits, but will be attractive as a place to start only to those organizations with a lot of experience in data management precursors like integrated customer information systems and with substantial critical needs.

(1) Reconciliation Approach (Consolidation Style)

Consolidation would seem to be the most obvious starting point for Master Data Management. It is the post-merger and acquisition mode for many organizations. Of course, this approach is most satisfactory to the party whose vocabulary and definitions persist (and who may be defined, less kindly, by others as the tyrant). So, this approach is easiest to deploy in situations where some party has the upper hand and can impose standards. Otherwise, propagating the solution on a wider basis can be culturally difficult.

In a consolidation approach, the MDM database becomes the point of reconciliation. Consolidation of existing systems increases business effectiveness while giving hardware and software licensing savings in the near term. If your critical problem set of information is limited, this may be a good approach to take.

(2) Brokering Approach (Registry Style)

In many enterprises, lines of business have become viewed as components – capable of external and internal relationships, instead of dangling boxes on an organizational, hierarchical tree. In such situations, defining *information sources* by what they have to offer and *information requests* by what they wish to consume as a service may be a satisfactory approach.

These declarations become the basis of a brokering or registry approach. The MDM database and indices becomes a hub offering cross-reference links. This style is particularly attractive to organizations that have not had experience with integrated customer information systems. It offers a quick way to offer access to data sources, together with descriptions of their requirements and limitations, to facilitate broader information use. This may meet needs for internal collaboration well.

(3) Diplomatic Approach (Coexistence Style)

There are organizations, particularly those that arise between mergers of equals, where large existing information systems each are internally consistent, but which cannot be easily reconciled. In this case, some form of diplomatically-arbitrated coexistence may be the best way to move forward. It may be a great paradigm for clearly-defined associations or partnerships, in organizations where collaboration has a narrow focus, or where there is a low rate of information change. Where territories have long been staked out, coexistence can be politically attractive. Over time, the efficiency of more common usage will become evident.

The Coexistence style makes existing data sources work like read-only operational data stores. A daily or weekly snapshot of changes stores. A publish-subscribe type of process keeps everybody "pretty darn close to the truth."

(4) Integrated Approach (Transactional Style)

The fully-integrated, transactional style of MDM is a superset of all the other approaches, with the addition of full audit capabilities and authorization at the data-attribute level. It supports atomic transactional system integrity, including all SOA transactions, and produces a system of record.

It is also the most intrusive approach to deploy, involving more rip and replacement than other modes. However, for many enterprises, including those in financial services and other regulated industries, all information related to regulated processes must be addressed, and the Transactional approach offers the cleanest path to the desired end state.

Hybrid Approaches

Most customers start with a style, but find that, organizationally, they don't fit into just one style. IBM has found that many of its MDM customers evolve to a multi-modal hybrid, for the information they care about differs in its rate of change and the risks associated with a particular

kind of information. If they have instituted MDM to improve operational problems, they may discover that it also addresses collaborative needs. If they start with MDM to foster collaboration, they may find it benefits operations that seemed pretty cut and dried. The value of MDM is cumulative, for the more you know about a customer, a situation or a market the better your service can be. Therefore, IBM has composed a *Multiform Master Data Management* offering that will support their customers in both the short and long term.

What is Needed to Support a Viable Solution Over Time

To support any and all of the modes described above, a viable solution consists of the following elements:

- ***A server***[3] that serves as the basis of operations.
- ***A data model*** is the basis for master data management. It holds data definitions, attribute definitions, relationships and service definitions. Services can be fine-grained (create, read, update, and delete for data attributes) or more coarse grained (privacy preferences). Because the services are linked by designated business rules, adding a new customer adds a party to the MDM Server and cascades other definitions.
- ***A workflow tool and integrated development environment*** *(IDE)* to add data discipline to operations on an ongoing basis. As long as the business is reacting to changing situations, new concepts and initiatives will arise. New data relationships and attributes are inevitable – and must be propagated across the relevant domains.
- ***Entity Analytics*** to sort through the ongoing complexities of analyzing near-duplicate parties. As a component of a platform, these analytics can be used where they are needed, and turned off where they are not. Businesses that are vulnerable to internal fraud will opt to use entity analytics pervasively, while others will limit it to particular target areas.
- ***Data Tolerance*** is needed, for new kinds of data, like that from RFID tags, sensors, and analysis of surveillance tapes, may become part of a master data management initiative. A platform can deal with them in a way that a single-purpose appliance cannot.
- ***Data Quality and Staging elements*** to inspect additional data sources, as they become available, to determine their best use. In the case of IBM, its Information Server provides a rich assortment of needed functionality.
- ***Auditing*** MDM operations is prudent, and often required by regulations.
- ***Industry models*** can have great value as a template of usual relationships between operational parts. They can accelerate consolidation efforts and guide the orchestration of a coexistence mode. IBM has developed detailed models for several industries.
- ***Metadata Management*** is a core functionality in Master Data Management, for it is in metadata that definitions are expressed.
- ***Delivery Mode***, well thought out, will make master data a part of ongoing operations. Master Data will be used not just by business analysts and marketers, but by operations, partners and strategists – and all of the applications they employ. The ability to feed customized views to relevant stakeholders is also part of the delivery strategy. The ability to mask particular data is a key capability for satisfying the needs of external partners safely. Without a compellingly stable delivery mode, Master Data Management may not be effective.

The Positive Effects of MDM

The business value of Master Data Management comes – most obviously – in *cost avoidance* realized by using fewer, better-managed information stores. By offloading master data management functions as a centralized function, production systems run more efficiently, often using fewer CPUs (or, if on a mainframe, fewer MIPS).

Risk avoidance is harder to quantify, but just as welcome. With Master Data, it becomes less expensive and less risky to extend existing business operations. Existing Master Data becomes a point of data congruence for new processes to design to, rather than making their own designations.

The most important benefit is one of *time*. People are more productive, decisions can be reached more expeditiously, and sales people can respond to opportunity with greater assurance.

Master Data Management can unify blended organizations It con improve corporate planning by surfacing aspects of a relationship that were formerly hidden in disparate systems. It can increase cross-sell and up-sell opportunities by

[3] The IBM MDM Server can run on any IBM server and on third-party Linux, UNIX, and Windows servers as well.

providing a more complete view of the customer. It can increase a product's time to market by increasing data congruence across the product lifecycle. In general, it just makes an organization work better.

Master Data Management is an operating strategy, not a quick fix. It addresses something that always will be a challenge of any successful company. It fosters the organizational coherence that was difficult even before the complications of multiple modes of communications and the Web. With increasing use of collaboration to harvest expertise and shrink the costs of running a business, the disciplines of Master Data Management become more and more important. The more you collaborate as a business, the more you must be a diplomat – finding common grounds and definitions on which you can agree or disagree. You must consolidate disparate approaches. You must offer information as a service. Moreover, you must work with all the available information, not just that which is easy.

To some extent, all business – like all politics – is local, for customer satisfaction, in the fullest sense of the term[4], is they key around which successful businesses are built. Having trustable and comprehensive information at your fingertips is the only way to satisfy them all.

Conclusion

Growing or evolving a business demands deft use of trustable data at all levels of the organization. Without trustable information, business planning and promises are just pipe dreams. It's just that simple – and that complex.

The business climate is such that businesses will be seeking to leverage their information more, not less. They will be partnering more, not less. The cost of sloppy data practices will be high, and getting higher. It is time to start implementing Master Data Management, and IBM offers a number of routes by which to accomplish the journey.

[4] Within an organization or value chain, there are usually many internal and external "customers" looking for satisfaction.

About The Clipper Group, Inc.

The Clipper Group, Inc., is an independent consulting firm specializing in acquisition decisions and strategic advice regarding complex, enterprise-class information technologies. Our team of industry professionals averages more than 25 years of real-world experience. A team of staff consultants augments our capabilities, with significant experience across a broad spectrum of applications and environments.

- ***The Clipper Group can be reached at 781-235-0085 and found on the web at www.clipper.com.***

About the Author

Anne MacFarland **is Director of Enterprise Architectures and Infrastructure Solutions for The Clipper Group.** Ms. MacFarland specializes in strategic business solutions offered by enterprise systems, software, and storage vendors, in trends in enterprise systems and networks, and in explaining these trends and the underlying technologies in simple business terms. She joined The Clipper Group after a long career in library systems, business archives, consulting, research, and freelance writing. Ms. MacFarland earned a Bachelor of Arts degree from Cornell University, where she was a College Scholar, and a Masters of Library Science from Southern Connecticut State University.

- ***Reach Anne MacFarland via e-mail at Anne.MacFarland@clipper.com or at 781-235-0085 Ext. 28. (Please dial "1-28" when you hear the automated attendant.)***

Regarding Trademarks and Service Marks

The Clipper Group Navigator, **The Clipper Group Explorer**, **The Clipper Group Observer**, **The Clipper Group** ***Captain's Log***, **Clipper Notes,** and *"clipper.com"* are trademarks of The Clipper Group, Inc., and the clipper ship drawings, *"Navigating Information Technology Horizons"*, and *"teraproductivity"* are service marks of The Clipper Group, Inc. The Clipper Group, Inc., reserves all rights regarding its trademarks and service marks. All other trademarks, etc., belong to their respective owners.

Disclosure

Officers and/or employees of The Clipper Group may own as individuals, directly or indirectly, shares in one or more companies discussed in this bulletin. Company policy prohibits any officer or employee from holding more than one percent of the outstanding shares of any company covered by The Clipper Group. The Clipper Group, Inc., has no such equity holdings.

Regarding the Information in this Issue

The Clipper Group believes the information included in this report to be accurate. Data has been received from a variety of sources, which we believe to be reliable, including manufacturers, distributors, or users of the products discussed herein. The Clipper Group, Inc., cannot be held responsible for any consequential damages resulting from the application of information or opinions contained in this report.