

Infostructure Associates

Master Data Integration: A Key Enabler of Master Data Management

White Paper

April 2006

Infostructure Associates
an affiliate of Valley View Ventures
24 Muzzey St.
Lexington, MA 02421
Telephone: 508-265-1558
www.valleyviewventures.com

Master Data Management: A Definition

This whitepaper discusses the role that Master Data Management (MDM) is playing in the business strategies of many organizations, and in these organizations' evolution to on-demand enterprises. In order to enable users to make these strategies a reality, the paper provides the reader with an important perspective on the value of IBM's Master Data Integration discipline and the importance of integration to successful MDM deployments.

To begin with a working definition of MDM: Master Data Management (MDM) is the business process, the applications, and the technical integration architecture used to create and maintain accurate and consistent views of core business entities across disparate applications in the enterprise. In other words, MDM helps a financial institution ensure that no matter how many ways you interact with it — via ATM, internet, branch office, credit card, marketing or custom service department — the institution knows you.

But isn't that what companies have been doing all along? Not really. The reality for many corporations is that each line of business and its systems are autonomous, and the corporation lacks the ability to establish data consistency across systems. For years now, companies have known of the need for MDM, and tried to approximate it using a number of solutions. Unfortunately, they have met with only partial success, which has made overcoming business process challenges, rationalizing outdated infrastructure, and dealing with mergers and acquisitions extremely complex and time consuming. In short, a lack of successful MDM has reduced organizational agility and led to missed business opportunities.

However, if MDM is implemented successfully it can enable:

- **Faster time to market.** MDM provides a single system for creating and maintaining product information, promotions, and rich, accurate consumer communications via online and traditional channels. One major retailer used MDM to reduce the time to introduce a new product from four weeks to one day.
- **Supply chain improvements.** A single, accurate and agreed-upon definition of products and suppliers, made possible by MDM, eliminates duplication, increases buying power, and provides insight into supplier relationships. A major European retailer is combining MDM with IBM WebSphere® Portal to allow more than 30,000 suppliers to manage their own master information.
- **Closer customer relationships.** By creating a single 360 degree view of customers, MDM helps sales, marketing, and service teams better anticipate customer needs, provide targeted offers, and improve customer service.
- **Better integration.** IT departments use MDM to eliminate information silos that have developed across their companies, reducing integration costs, enabling collaboration, and improving business productivity. A major European manufacturer uses MDM to cleanse and synchronize rich, accurate master information across more than 200 instances of SAP and other ERP systems.
- **Compliance** with industry mandates and government legislation. IBM offers MDM solutions and deep expertise that help customers comply with information-centric mandates like Sarbanes-Oxley, the Patriot Act, Basel II, ACORD, and HIPAA (Health Insurance Portability and Accountability Act), and enable global data synchronization among retailers, distributors and manufacturers.

Essentially, there are two important facets of implementing a Master Data Management strategy successfully:

1. An appropriate level of *business* participation, including the development of data governance best practices.
2. *Technical* best practices, a technical architecture, and tools for responding to the MDM challenge.

Therefore, a complete MDM solution has two major technology components: the MDM platform embodying business best practices, and the integration foundation — Master Data Integration — responsible for its deployment and ongoing data integrity.

Master Data Integration

IBM Master Data Integration (MDI) is the use of the WebSphere Information Integration platform, with related technical best practices, to support Master Data Management deployments within application and business process infrastructures.

Master Data Integration is essential because of the state of information in the typical enterprise. Information about key enterprise data is hard to find or non-existent; enterprise data is typically poor in quality and widely differing in format. As a result, information from across the enterprise needs to be identified, de-duplicated and matched prior to being loaded into an MDM system. And each of these steps is typically complex, risky and time consuming.

IBM Master Data Integration:

- Assesses the scope of data inconsistency in the source silos.
- Enables faster deployment of MDM by creating the MDM repository.
- Ensures ongoing data standardization, rationalization and access to key data and content.
- Delivers ongoing data monitoring to characterize the level of data compliance.
- Helps to embed data “best practices” in the organization’s culture through data governance.

The Past is Prologue

Clearly, Master Data Management is not the first attempt to rationalize the key data within the enterprise. There have been many past attempts that have not succeeded in delivering a “golden master” of data across the enterprise. What has gone wrong with these past attempts, and why is MDM better?

Customer Information Files and Operational Data Stores

Many solutions that have provided applications and operational processes with golden masters of data were deployed as Customer Information Files (CIFs) and Operational Data Stores (ODSs). While these were effective in supporting specific lines of business, with simple inter-party relationships, they typically lacked the real-time updates and sophistication of modern solutions, instead relying upon batch updates and basic functions. For example, they were not designed to handle entitlement and privacy rule checking, to support composite data transactions, or to perform duplicate data processing; in short, they were

very inflexible. By contrast, today's MDM systems can be used to feed multiple CIFs in order to bring them into alignment, or to provide a replacement strategy in support of new business processes as the organization re-engineers its IT infrastructure.

Customer Relationship Management (CRM) and ERP

These applications provide solutions for managing customer-facing or manufacturing processes such as sales, customer service, inventory management and general ledger. They are designed to manage their own data services that support their applications' user interfaces and processes. CRM systems, for example, were designed to coordinate sales, marketing and customer service data, but lack the necessary tools to manage master customer data from within the application itself. There are three points to consider:

1. CRM and ERP still offer only one application's perspective on enterprise master data, and they do not easily lend themselves to being extended to support the broader role of providing application-neutral master data services.
2. While Enterprise Application Integration (EAI) has provided connectivity and orchestration of transactions among major application suites, EAI has done very little to ensure data consistency across all of the applications.
3. As new key data and new functional requirements such as web-enabled e-commerce, radio frequency identification (RFID), and Global Data Synchronization have surfaced, these typically fall outside the purview of ERP and CRM — and they are likely to remain so, because ERP and CRM lack an enterprise-wide master data perspective.

Data Warehousing

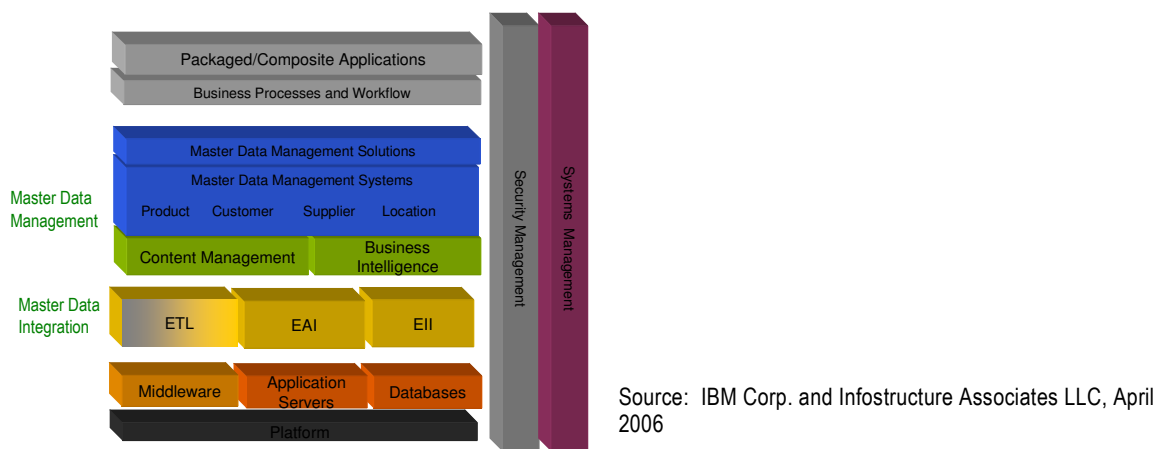
The first attempt at creating one consolidated set of "master data" for customers and products across the enterprise was the data warehouse. Data warehousing best practice consolidates multiple enterprise data sources to a common data store. The warehouse usually includes master data about customers and product items along with associated historical event data such as purchase orders, to create a usable perspective for analysis. While the data warehouse provides invaluable insight into the business, it is not typically designed for synchronized updates of master data, and it lacks the necessary transactional robustness and infrastructure to support centralized Master Data Management.

Data in the warehouse is consistent, reliable and accurate, due to the use of data integration best practices and tools to establish and enforce data integrity. In fact, there is a natural synergy between the disciplines, concepts and tools used in Master Data Management and those used in data warehousing. But these data integration tools by themselves do not represent an MDM solution. Note that once a full MDM system is in place, the MDM system of record becomes the primary source of data warehouse dimensional data, to ensure consistency in operational and management reporting.

Master Data Management Architecture

Companies planning to implement Master Data Management today can take advantage of a wide ranging of offerings, including master data repositories with management facilities, solution accelerators with industry data models, associated business data services, and turn-key templates for such entities as product catalogs, RFID, global data synchronization, and account management, as well as Master Data Integration solutions. Figure 1 shows an example of an MDM Architecture.

Figure 1: Example of an MDM Architecture



In this architecture, the **Master Data Management System** is a transactional software infrastructure that manages a repository of reference data through a package of business services interfaces. These services include administrative and user screens, repository management functionality, workflows, and events that define and maintain the data hierarchies, relationships and attributes associated with specific data elements and industries. For example:

- For retailers and consumer packaged goods manufacturers, an MDM system such as IBM WebSphere® Product Center manages part numbers, descriptions, pricing, images and packaging details, suppliers and vendors. This form of MDM system is commonly referred to as a Product Information Management (PIM) system.
- For financial services, retail banking, telecommunications, and insurance companies, an MDM system such as IBM WebSphere Customer Center manages financial profiles, location, demographic data, billing and account information. This type of MDM system implements a concept of “party” that covers all roles associated with individuals and legal entities and is commonly referred to as a Customer Data Integration (CDI) system.
- Organizations often build custom applications for the Master Data Management System to provide product- or customer-oriented information management.

The **Master Data Integration** portion of the architecture provides a single integration infrastructure necessary to solve implementation and ongoing operational data challenges across any industry, business function and scope of business data. The functionality includes data matching, linking, correction, and standardization across the enterprise. In addition, federation middleware enables MDM systems to dynamically access external data sources for content such as images and documents related to the entities managed by the MDM system.

The **Master Data Solutions** comprise industry-specific data models, workflows and business processes that ensure the speedy customization of MDM systems targeted at core business domains. These solutions include reusable integration templates that help to rapidly move and cleanse data from common application sources such as ERP and CRM during the initial load of the MDM repository.

An Application-Neutral Infrastructure for MDM

As noted earlier, past attempts to deliver consistent master data have resulted in an inflexible and narrow support for master data across the enterprise. The weaknesses of the previous approaches highlight the importance of “taking a stand” to provide an application-independent infrastructure that focuses exclusively on the accuracy, integrity, and richness of the data managed.

An “application-neutral” approach means using a “common data format” (also known as a data model) that has been designed exclusively for the business, and that ensures that existing and future applications can be easily connected to the MDM system through integration. Additionally, while there is just one source of authoritative master data, the application-neutral approach does not require all of the data of the enterprise to be consolidated into the single MDM system. Instead, each enterprise application will continue to maintain its own data store, while the MDM system is responsible for enforcing business rules that govern the reliability of the company’s key master data, e.g., customer or product data.

Within the MDM repository, the common data format represents the key attributes of a given entity that overlap with enterprise application data and content, plus external data such as SKUs that may require synchronization with partners. This synchronization can be accomplished in conjunction with an EAI infrastructure. For example, a Master Data Management system for customer information may manage account and customer data across multiple banking systems, including a consumer portal, ATM and back office systems. Once established, this master data is then used to synchronize customer data across other enterprise applications as part of the organization’s usual business processes, often by using EAI. This approach helps to eliminate redundancy, improve accuracy, and facilitate broad usage of the data.

In summary, Master Data Management:

1. Focuses on synchronizing specific portions of the entire enterprise data inventory that are deemed common and strategic across the company’s internal systems and external interactions.

2. Leaves existing applications in place, and superimposes an infrastructure that enforces corporate oversight on data elements related to key business entities such as products and customers, while providing maximum flexibility for evolution.
3. Supports synchronization of master data using real-time updates across all participating applications, effectively linking every instance of master data as if there were one gigantic, logical database, thereby creating data consistency.
4. Supports the deployment of solution templates to address discrete operational, compliance and rationalization challenges such as new product introductions, new account openings, ATM preference management, and RFID.

MDM and MDI Best Practices

Master Data Integration provides a critical foundation for any MDM initiative. First, it enables MDM solutions to be deployed faster. Second, the resulting solutions deliver higher data quality. For example, IBM's data-quality offering, WebSphere QualityStage, can dramatically improve the quality and the effectiveness of key organization information. Third, it provides additional flexibility that enables an MDM solution to integrate with future tools and applications.

Master Data Integration is one of the core disciplines of the IBM information integration strategy. The IBM MDI solution includes a best practices roadmap, a reference architecture, and the use of the WebSphere Information Integration platform, which can be applied to any Master Data Management project.

Note that the versatility of the WebSphere Information Integration platform increases its value. It can accelerate multiple data-centric projects such as MDM, ERP infrastructure rationalization, legacy system migration, and business intelligence/data warehousing projects. The economies of scale and improvements in data cleanliness engendered by this approach become extremely advantageous to the business. Additionally, WebSphere Information Integration can be a keystone of an organization's coherent vision for information and process services, as the company transforms itself into an on demand enterprise. Hence, the WebSphere Information Integration platform is useful whether or not the organization implements MDM.

Recommended Best Practices

To aid customers considering MDM, IBM recommends the following best practices based upon its MDM engagement experience. Customers should:

1. Focus on using an application-neutral definition of "master data," one which defines core business entities such as customer, employee, citizen, account, or supplier with their related data and content without any dependencies on a particular application. Such an application-neutral definition can evolve with the needs of the business.

2. Implement a “blueprint” process that maps out the enterprise information at the inception of the MDM project. This will create better understanding of the organization’s data inventory and enable the establishment of standards that can be continuously monitored.
3. Provide management oversight through a data governance program, e.g., via a center of excellence for MDM and MDI as well as data-stewardship best practices. Data governance can be an effective tool to ensure that data-centric business initiatives, such as the creation of an eCommerce presence using a unified product catalog, get implemented correctly and on time.
4. Examine the flow of current and future data and master data, in and out of corporate systems and partner systems. This examination will help to determine where master data synchronization needs to occur and where dependencies exist across applications and processes.

IBM Master Data Integration Methodology and Process

There should be a *phased* approach to the use of Master Data Integration in MDM deployment and on-going maintenance. To set up the MDM data store, organizations must identify all variants of key data, and the relationships between them, and then perform an initial cleansing and translation of the data (See Table 1). This can be an exacting process — but the results are well worth it. Simply identifying and cleansing all variants of customer data can lead to dramatic increases in customer satisfaction, as customers avoid constant re-entry of their vital statistics and irritation at misinformation.

Table 1: Example of the IBM MDI Methodology and Process

Phase	Tasks	MDI Functions Used
Phase 1 (Requirements): Capability Assessment	<ul style="list-style-type: none"> • Confirm Requirements & Current Footprint • Assess Capabilities • Perform Gap Analysis • Develop Alternatives and Benefit Statement 	Profile & Understand (e.g., WebSphere ProfileStage™)
Phase 2 (Requirements): Blueprint	<ul style="list-style-type: none"> • Form Project Team • Create Initial Architecture • Extract & Load Data Samples • Develop Project Timeline 	Profile & Understand, Connect & Populate (e.g., WebSphere DataStage®)
Phase 3 (Deployment): Data Assessment	<ul style="list-style-type: none"> • Discover and Validate Sources • Establish Technical Assessment • Validate Business Rule Requirements 	Federate (e.g., WebSphere Information Integrator Standard Edition), Profile & Understand, Connect & Populate
Phase 4 (Deployment): Data Alignment	<ul style="list-style-type: none"> • Conduct Initial Data Alignment • Cleanse Data Sources • Confirm Technical Specifications • Agree upon Test Specifications • Establish Data Validation Routines • Create Lookup Tables 	Transform & Enrich (e.g., WebSphere DataStage and WebSphere QualityStage™), Match & Standardize (e.g., WebSphere QualityStage), Profile & Understand, Connect & Populate, Manage Metadata (e.g., WebSphere MetaStage™)
Phase 5 (Deployment): Data Harmonization	<ul style="list-style-type: none"> • Create Batch Updates: Consolidate, Integrate, Cleanse, Normalize, & Harmonize Data 	Federate, Transform & Enrich, Match & Standardize, Profile & Understand, Connect & Populate, Manage Metadata

Phase 6 (Operation): Data Integrity	<ul style="list-style-type: none"> • Determine Hierarchy & Object Management • Synchronize Data and Access External Content • Enforce Data Integrity Rules • Manage Events & Transactions • Manage Repository and Entitlements • Determine Data Governance Reporting needs 	Federate, Synchronize (e.g., WebSphere Business Integrator, WebSphere DataStage TX, WebSphere Information Integrator Replication Edition), Monitor Data Quality (e.g., WebSphere ProfileStage), Transform & Enrich, Match & Standardize, Profile & Understand, Connect & Populate, Manage Metadata
Phase 7 (Operation): Evaluation and Assessment	<ul style="list-style-type: none"> • Monitor Data Quality Within MDM system • Prioritize Forthcoming Initiatives and Anticipated Data Content 	Monitor Data Quality (e.g., WebSphere ProfileStage)

Source: IBM Corp. and Infostructure Associates LLC, March 2006

Business Benefits of MDM

MDM provides a platform to overcome rampant master data fragmentation across applications and external systems. This defragmentation delivers business benefits to companies in a wide range of industries. Business benefits of MDM include:

- Competitive advantage.
- More effective reengineering.
- Better compliance with government legislation and industry mandates.

Competitive Advantage

MDM helps companies to develop a competitive-advantage strategy that is information-driven and not application-focused.

Today, we are seeing a fundamental shift in effective strategies for achieving competitive advantage via IT. The new strategy is to use proprietary information about customers and suppliers to gain insight and create value that competitors cannot duplicate.

The key barrier to achieving competitive advantage by leveraging proprietary information is the complexity and disparity of data and its distribution across a wide range of data sources. Therefore, IT needs to implement a broad strategy of (a) placing key proprietary information where it is easily accessible, and (b) ensuring that the rules for data and relationships between the data elements are clearly identified and consistently enforced — in effect, a Master Data Management strategy.

Business Process Re-engineering

MDM is a natural enabler for an organization’s push to re-engineer and integrate business processes for greater efficiency and agility. Additionally, new composite applications that enhance existing business processes can be created based on consistent handling of key data across all sub-applications using MDM — speeding composite-application creation by as much as 25-30 % (source: Infostructure Associates development TCO studies). An ex-

ample of such a composite application using an MDM core is a CRM-oriented portal or e-commerce application.

Data Governance

While MDM enables corporations to comply with industry mandates and governmental legislation, many companies are taking the next step and investing in “data governance” as a key component of an effective business compliance strategy. Usually, groups are set up with a data governance mandate to help organizations navigate the business implications of a far-reaching strategy. For example, rolling out a single global instance of SAP R/3 requires that multiple lines of business agree on data standards and data audit controls.

With executive management support and both business and IT-level participation, data governance functions ensure that the data and processes needed to support a project are created and managed, and that they are auditable. Standard processes for identifying and maintaining master data effectively increase the likelihood that critical data will be easy to find, and will be immediately available when needed for compliance with legislation or industry best practices. The value of data governance increases as MDM is used to bring more lines of business, subsidiaries, channels, customers, and retailers into alignment.

Global Food Retailer Improves Quality of Customer Information

A global dairy-food retailer has recently undergone a merger and rapid growth. The retailer seeks to “be one company” by integration, simplification of the architecture, and optimization of processes, especially supply chain processes. The retailer



is creating a company architecture in which local systems exchange master data with a corporate SAP R/3 implementation by way of a Global Master Data Repository (GMDR). The master data project provides global definitions for customer, material, supplier, and employee in accordance with relevant standards (e.g., EAN-UCC); identifies global data that can be fed into the repository; sets up translation mechanisms between this global data and the repository’s “common format”; populates the repository; and creates mechanisms for access to the repository from the SAP R/3 implementation and other applications. The food

retailer has chosen to make its master data administration a workflow solution in order to make it “as simple and consistent as possible.” The repository is housed in a separate system for performance. Global data is fed into a “work database” for cleansing and translation, then into the repository. The repository, in turn, integrates with applications via EAI.

As part of the master data project, the retailer has set up a “virtual master data organization,” with existing employees tasked with being “object owners” (overall responsibility), “MD responsible” (guiding maintenance and handling MD budget within a business unit), “MD owners” (responsibility within a business unit), and “maintainers” (maintenance im-

plementers in a business unit). Other organizational tasks include training, networking across business units, and creating forums for information sharing.

Creating a common format has involved integrating more than 30 global data stores with about 224,000 records. A key difficulty has been handling duplicates, as the same material may have different records in different global data stores.

A vital factor in the project's success has been the use of the data-cleansing capabilities of IBM WebSphere QualityStage. This approach has helped the retailer determine how much of the data is already consistent and therefore how it can be harmonized. It has also provided tools for assessing the quality of the existing data, for standardizing "key" fields such as addresses for better duplicate determination, and for handling "near-matches." This standard process, in turn, has reduced the number of records needed by about 25%, with corresponding improvements in performance and data quality.

Getting Started

IBM recommends that companies considering MDM take the following initial steps:

1. Assess the scope and criticality of your organization's current master data problem, and the areas where MDM can yield business process improvements.
2. Identify the set of master data yielding the quickest value-add from MDM. For example, you might focus on customer data for CRM applications or supplier and supply-chain data for SCM applications.
3. Perform due diligence on MDM vendors, assessing them according to the criteria listed below — being sure to assess the extent of their Master Data Integration capabilities.
4. Consider "kicking the tires" by prototyping your MDM implementation or doing a run-through on your MDM business process. Prototypes can often identify areas of difficulty or of high importance before actual implementation, and assess any potential scalability limitations.
5. Assess the strength of the preferred vendor's services offerings to support your MDM initiative, including best practices, blueprints, and project roadmaps.
6. Ask the vendor for references for customers with profiles similar to yours, including similar industries, comparable IT environments, similar data standards, and comparable transactional data volumes.

Be sure that your MDM solution provides broad capabilities, including:

- *Adaptability.* Are there multiple interfaces into the MDM system (e.g., Web Services, Java, COBOL)? What pre-built integrations with EAI/Business Process tools exist?
- *Solution Readiness.* Is there e-commerce integration? How about support for standards such as Global Data Synchronization, product catalogs, and RFID? Is there access to rich data types and external content?

- *Out-of-the-box Templates.* Is there an adaptable data model? Is there a library of pre-built and customizable business services, with fine and coarse-grained transactional update capability?
- *Data Integrity.* What MDM services are provided for key data integrity functions such as validation, standardization and matching?
- *Support Infrastructure.* What Master Data Integration capabilities such as data profiling, data quality and data transformation are available to speed deployment, and increase operational data integrity?
- *Scalability and Manageability.* Can the MDM system handle real-time, high-volume update transaction streams across very large data sets? Does the solution offer extensive data administration and systems management capabilities?
- *Vendor Services and Roadmap.* Are the vendor's services well regarded and sophisticated? Are there specific services for MDM and MDI? Does the vendor incorporate MDM best practices and training for creating an MDM business process? Does the vendor offer a roadmap that is credible and that meets your ongoing needs?

Where Should I Go For More Information?

For an introduction to IBM's Master Data Management solutions and Master Data Integration capabilities, please visit:

<http://www-306.ibm.com/software/data/integration/solutions/mdm.html>

Infostructure Associates Conclusions

MDM offers users not only the hope of beginning to link and rationalize their archipelagoes of under-leveraged data, but also the ability to achieve major business benefits:

- Competitive advantage by leveraging proprietary information more easily.
- Cost-cutting by eliminating costly administration of duplicate and unstandardized data.
- Support for business-process reengineering, e.g., in mergers and acquisitions.
- More effective business compliance, where rapid access to information is key to cost-effective compliance.

In all of these cases, MDI allows users to carry out MDM more effectively and cost-effectively. Moreover, MDI makes MDM forward-looking — its standards-based approach and flexibility provide insurance against future technologies, standards, and requirements.

Corporate strategists assessing MDM should note another key point: The reason that we recommend that large organizations move rapidly towards MDM/MDI implementation is not just that MDM/MDI offers the potential for competitive advantage and more effective business compliance and risk management, but also that *organizations are already seeing that potential in the real world*. As we have seen in our case studies, users are seeing dramatic, quantifiable improvements in data quality and hence in decision quality and customer satisfaction.

Finally, MDM and MDI are also effective building blocks for further value-add. To cite one example, the key to achieving the “real-time enterprise” is an “information on demand solution” such as MDM that provides transactional access to all actionable organization data, and instant visibility of updates. MDI can be a key enabler of the on demand enterprise, helping companies to standardize critical data elements for broad use across the enterprise.

We therefore strongly recommend that large organizations investigate MDM and MDI — *now*.

Infostructure Associates
24 Muzzey St.
Lexington, MA 02421
USA
Telephone: 508 265 1558
Fax: 781 862 6236
www.valleyviewventures.com

© 2006 Infostructure Associates
All rights reserved
April 2006

Infostructure Associates is an affiliate of Valley View Ventures that aims to provide thought leadership and sound advice to both vendors and users of information technology.

This document is the result of sponsored research performed by Infostructure Associates. Infostructure Associates believes that its findings are objective and represent the best analysis available at the time of publication.

About the analyst: Wayne Kernochan is a long-established expert and thought leader in the database, development solutions, and Enterprise Information Integration (EII) areas. He has been a strong advocate of strategic information management and agile programming. In 2002, 2003, and 2004-5 he has issued yearly database, development, and EII reports as well as in-depth qualitative research in these areas, including TCO/ROI and case studies.

During an eleven-year tenure at Aberdeen Group, Kernochan was the author of their unique approach to TCO studies and held the positions of senior director, managing vice president, and senior vice president. Kernochan is frequently quoted in publications such as Information Week and CIO Magazine.

Email the analyst:

wayne.kernochan@valleyviewventures.com