

Managing Meta Data for the Business

By David Marco

The balance sheet of most any *Forbes* Global 2000 company would show entries for assets such as property, cash, equipment, and accounts receivable. Unfortunately, one item that is not seen in the asset section of the balance sheet is data. Data is every bit as valuable as property, equipment, and accounts receivable. Low-quality data, poorly understood data, mismanaged data, redundant applications, and poorly built applications prevent companies from effectively managing their other assets. How can a company convert accounts receivables into cash when the accounts receivable system has transaction records with data quality issues that prevent them from ever becoming billable? Moreover, all of the needlessly redundant applications create a substantial cost drain on the enterprise's cash assets.

This is an especially important concept because all companies are desperately trying to increase shareholder value. Corporate executives spend numerous hours looking for ways to increase their company's value, because 95% of these executives' compensation is directly linked to shareholder value. They realize that shareholder value is tied not just to the assets on the balance sheet, but also to nonphysical factors (e.g. intellectual capital, customer loyalty, brand recognition, etc.). Moreover, CEOs and CIOs are using successful technology implementations as trophies to improve shareholder value by enhancing the company's reputation as an innovator and by attracting better employee talent.

Understanding and leveraging technology is critical for any enterprise. The average company spends 5.3% of its gross revenues on its IT applications. This means that a company with \$1 billion in revenues spends, on average, \$53 million annually on their IT systems. Several *Fortune* 50 companies and large government organizations have IT budgets that approach or exceed \$1 billion annually. These same organizations have implemented systems to manage almost every aspect of their business, including payroll systems, accounts receivable applications, order entry systems, marketing campaign management, human resources systems, logistics, invoicing applications, and even systems to track the placement of office furniture and employee holidays. In fact, a great number of the same organizations have systems (though unauthorized) that manage the weekly football pool. Despite this massive investment, most companies do not have an application to systematically manage their IT systems. This reveals a fundamental truth about data management:

We build systems to manage every aspect of our business, except one to manage the systems themselves.

Despite spending these exorbitant amounts on IT, most companies still do not value data as an asset, whether on the balance sheet or in the board room. Most companies' IT development process can best be described as piecemeal. Typically, systems are built through "heroic effort" by group of developers and business users get together to implement a new application (e.g. data warehouse, customer relationship management, enterprise resource planning). These heroes embark on a perilous journey to understand the existing (and often convoluted) system's architecture, work many long hours, make business assumptions with little or no common business understanding, and then hope that their efforts will be successful. This situation explains why 60% to 75% of large IT initiatives fail. Even successful initiatives may not be repeatable, because an application development process is not in place and the standards for building the application have not been formulated or documented. Moreover, even those companies whose IT processes are repeatable find that they can be repeated only by a single group of developers. Their work and effort cannot be transferred to other groups. Even with a failure rate so high and an investment in IT so great, most companies still do not manage their applications systematically.

A managed meta data environment (MME) can manage a company's systems by cataloging the applications, data, processes, hardware, software (technical meta data), and business knowledge (business meta data) possessed by an organization. This information (meta data) can then be utilized to identify redundancies before they occur and to eliminate duplication that already exists. A world-class meta data management solution dramatically improves data quality by providing a full understanding of the data and is absolutely essential for having repeatable and transferable IT

processes because it centrally and completely documents data and applications. Business executives are beginning to realize the importance of managing their data as an asset and thus are starting to look to the MME as the key technical solution for data asset management.

The key to your company's prosperity is how well you gather, retain, and disseminate knowledge. Managed meta data environments are the key to gathering, retaining, and disseminating knowledge.

One of the chief challenges in building a successful MME is to decide on the specific business objectives for your MME. Over the next several months I will describe the most pressing challenges that organizations are currently facing, including:

- Reduce IT redundancy
- Provide IT portfolio management
- Prevent IT applications failure
- Reduce IT expenditures
- Enable knowledge management
- Adhere to regulatory requirements
- Enable enterprise applications

Reduce IT Redundancy

“CIO” is commonly defined as chief information officer; however, there is another possible meaning for this acronym—career is over. One of the chief reasons for this is that most IT departments are “handcuffed” by needless IT redundancy that too few CIOs are willing and able to fix. There are several CIO surveys that are conducted annually. These surveys generally ask, “What are your top concerns for the upcoming year? Data integration is usually high on the list. Data integration focuses on two key areas:

- Integration of data across disparate systems for enterprise applications
- Removal of IT redundancies

Some IT redundancy is a good thing. For example, during a power outage when one of your data centers is not operational, you need a backup. The discussion here focuses on *needless* IT redundancy, or IT redundancy that only exists because of insufficient management of the IT systems. I was working with a Midwestern insurance company that, over a 4-year span had initiated various decision support efforts. After this 4-year period, they took the time to map out the flow of data from their operational systems to their data staging areas and finally to their data mart structures. What they discovered is shown in Figure 1.

Typical IT Architecture

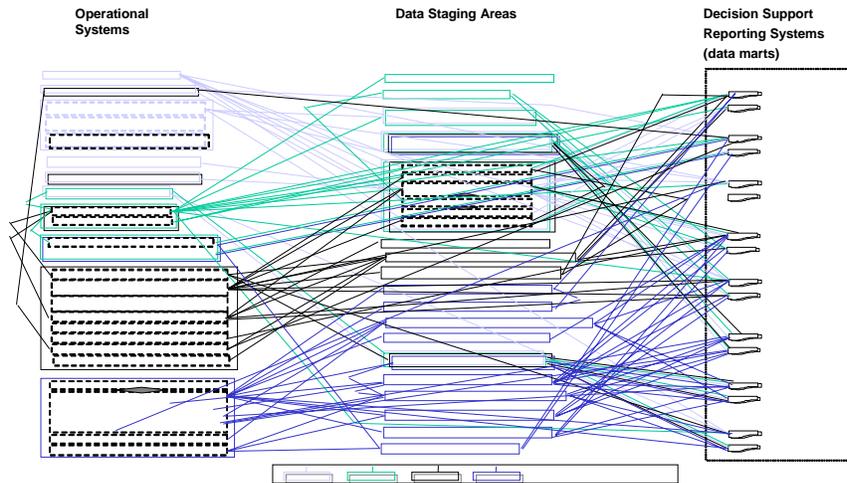


Figure 1: Typical IT architecture.

The typical response to Figure 1 is, “Where did you get a copy of our IT architecture?” If you work at a Global 2000 company or any large government entity, Figure 1 represents a simplified version of your IT architecture, which is actually no architecture at all. Poor architecture creates a litany of problems, including:

- Redundant applications, processes, and data
- Needless IT rework
- Redundant hardware and software

The following sections will discuss each of these problems in greater depth.

Redundant Applications, Processes, and Data

It has been our experience in working with large government agencies and *Forbes* Global 2000 companies that needlessly duplicated data is running rampant throughout organizations. One large banking client asked us to analyze its IT environment. During this analysis, we discovered a tremendous amount of application and data redundancy. The company had over 700 unique applications. During this analysis, we compared this client to a bank that is more than twice its size; however, this larger bank has a world-class MME and uses it to properly manage their systems. As a result, they have less than 250 unique applications. Clearly, the bank with more than 700 applications has a great deal of needless redundancy. The bank with the world-class MME was also 14 times more profitable than the one maintaining over 700 applications. Obviously, the less-profitable bank would become much more profitable if it removed its needless redundancy.

In our experience, a typical large organization’s data has 60% to 75% needless data redundancy. Some organizations have hundreds of “independent” data mart applications spread all over the company. Each one of these data marts is duplicating the extraction, transformation, and load (ETL) processes typically done centrally in a data warehouse. This greatly increases the amount of support staff required to maintain the data-warehousing system, and these tasks are the largest and most costly data-warehousing activities. Each data mart also copies the data, requiring even more IT

resources. It is easy to see why IT budgets are straining under the weight of all of this needless redundancy.

Fortunately, these large organizations are beginning to realize that they can't continue to operate in this manner. Therefore, they are targeting MME technology to assist them in identifying and removing existing application and data redundancy. This can be accomplished because the MME can identify redundant applications through analysis of the data and the application's processes. These same companies are starting IT application integration projects to merge these overlapping systems, in conjunction with their MME initiatives to ensure that future IT applications do not proliferate needless redundancy.

Needless IT Rework

During the requirements-gathering portion of one MME initiative, an IT project manager brought up the challenges that he faced in analyzing one of the mission-critical legacy applications that fed the data-warehousing application his team had been asked to build. During our interview he stated, "This has to be the 20th time that our organization is analyzing this system to understand the business rules around the data." This story is all too common, because almost all organizations reinvent the IT wheel with every new project. This situation occurs when separate teams build each of the IT systems without an MME, so the teams can't leverage each other's standards, processes, knowledge, and lessons learned. This results in a great deal of rework and reanalysis. A good MME captures this invaluable IT information, including business rules, business processes, attribute definitions, entity definitions, and data lineage. This information is invaluable to an organization that is looking to maximize its IT investment.¹

Redundant Hardware and Software

All of this redundant application and IT work generates a great deal of needless hardware and software redundancy, forcing the enterprise to retain skilled employees (usually at great cost) to support each technology. In addition, more money is spent because standardization doesn't occur. Often software, hardware, and tools can be licensed at a considerable discount to an enterprise. These economies of scale can provide tremendous cost savings to the organization.

In addition, the hardware and software is not used in an optimal fashion. For example, one client purchases hardware for each individual IT project. As a result, they have a bunch of servers running at 25 percent capacity.

Software presents even more problems. One IT project leader responded to the question of what software vendors his company was standardized on with, "All of them!" This leads to the old joke "What is the most popular form of software on the market? Answer: shelfware!" Shelfware is software that a company purchases and winds up never using, so it just sits on a shelf collecting dust.

IT Portfolio Management

Over the years, we have performed dozens data-warehousing assessments. During these assessments, clients are routinely asked how much they spend annually on data warehousing. Most cannot accurately estimate on what they actually spend. In order to manage these and any other costly IT initiatives, it is critical to measure each one, but it is impossible to measure them when most

¹ For more information on meta data sources and the types of meta data that they include, see Chapter 2 of "Building and Managing the Meta Data Repository", David Marco, Wiley 2000

companies do not understand them (see Figure 1). This is where IT portfolio management enters the picture.

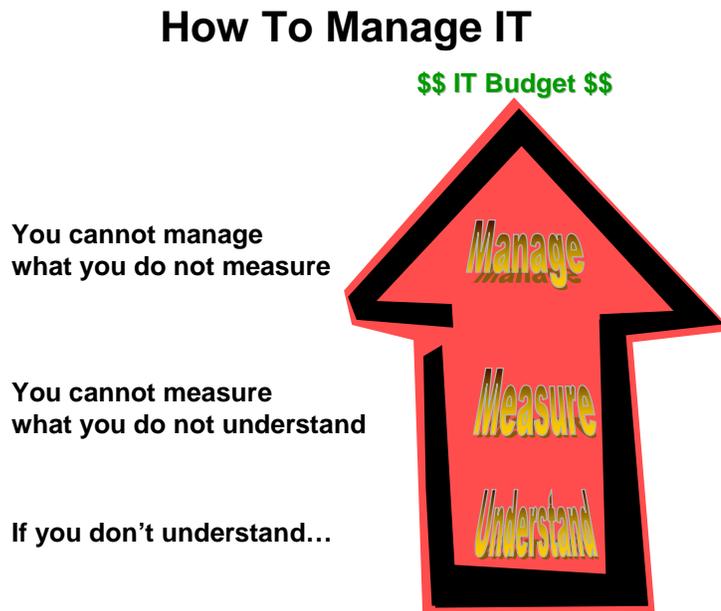


Figure 1: How to manage IT

IT portfolio management is the formal process for managing IT assets. An IT asset may be software, hardware, middleware, an IT project, internal staff, an application, or external consulting. As with other new disciplines, many companies are not setting up IT portfolio management efforts correctly. Following are some of the keys to building successful IT portfolio management applications.

The information that portfolio management targets is meta data (both business and technical), and, as such, needs to be stored in the MME. The MME will allow the corporation to aggregate the portfolio-related meta data into an executive view so that it is clear which projects are proceeding well and which are lagging behind. It can also publish information to a Web site so that front-line IT staff can see the status of other projects. This will greatly aid their project timelines and planning. It should be directly linked to the more granular business and technical meta data to allow developers to understand what IT projects are underway, what technology are these projects implementing, and what data sources are they accessing. Most large companies have a great deal of duplicate IT efforts (see the next section for more on this topic). This happens because the meta data is not accessible. At my company, we have a couple of large clients whose only goal is to remove these tremendous redundancies, which translates into tremendous initial and ongoing IT costs.

Finally, the MME should contain both a business and a technical component. In most cases, the business component is more complicated than the technical. For example, a project manager communicates a project's funding, status, and technology to the group responsible for the portfolio management application. The meta data in the portfolio application is only as good as the information that these managers provide, so it is vital to integrate the IT portfolio application into the company's IT production turnover procedures. This ensures that the portfolio application stays current².

² A detailed IT portfolio management meta model walkthrough is provided in Chapter 6 of "Universal Meta Data Models", David Marco & Michael Jennings, Wiley 2004

Prevent IT Applications Failure

When a corporation looks to undertake a major IT initiative such as a CRM, ERP, a data warehouse, or an e-commerce solution, the likelihood of project failure is between 65 percent and 80 percent, depending on the study referenced. This is especially alarming considering that these same initiatives traditionally have executive management support and cost many millions of dollars. For example, one large client was looking to roll out a CRM system (for example, Siebel, Oracle) and an ERP system (for example, SAP, PeopleSoft) globally within 4 years, with an initial project budget of over \$125 million. In my opinion, they have a 0 percent probability of delivering all of these systems on time and on budget. When was that last time you saw an ERP or CRM initiative delivered on time or on budget?

When we examine the causes of project failure, the following themes become apparent. First, the company did not address a definable and measurable business need. This is the number one reason for project failure data warehouse, CRM, MME, or otherwise. Second, they did not incorporate the existing IT environment and business rules. This includes custom applications, vendor applications, data elements, entities, data flows, data heritage, and data lineage.

A MME dramatically reduces the likelihood of IT project failure, because it allows a corporation or government agency to decipher its IT environment and to greatly speed up its IT development life cycles. The MME captures the key meta data around a company's IT applications and uses this meta data to generate an *impact analysis*. An impact analysis is a technical meta-data-driven report that shows the impact of a potential change to an organization's IT systems. Chapter 3 presents several MME use-case examples utilizing impact analysis. The MME case studies featured in that chapter also discuss impact analysis.

Reduce IT Expenditures

If your MME reduces the applications, processes, data, software, and hardware, and it lowers the likelihood for IT project failure and speeds up the IT development life cycle, then clearly it will greatly reduce your company's IT expenditures. The majority of MMEs are focused on reducing and managing IT expenditures.

Enable Knowledge Management

One of the most vital functions of any MME is to provide the technical architecture and processes to manage corporate knowledge. All corporations strive to become more intelligent. To attain a competitive advantage, a business needs its IT systems to manage more than just their data; they must manage knowledge (that is, meta data). As a corporation's IT systems mature, they progress from collecting and managing data to collecting and managing knowledge. Knowledge is a company's most valuable asset, and an MME is the key "technological backbone" for managing a company's corporate knowledge³.

George Bernard Shaw once said, "The greatest problem in communication is the illusion that it has been accomplished." Executives often do not realize how not properly managing a company's knowledge (meta data) can negatively affect their bottom-line profits, or they are under the illusion that proper communication and understanding are occurring. This couldn't be further from the truth.

NASA Mars Orbiter: A Case Study in a Lack of Meta Data Management

On December 11, 1998, the National Aeronautics & Space Administration (NASA) launched the Mars Climate Orbiter to study the climatic conditions on the planet Mars. The engineers who worked on this mission calculated rocket firing using feet per second. However, the orbiter was programmed in Newtons per second (metric) of thrust. The difference between these calculations was 4.4 feet per second. "Each

³ For more information on knowledge management, refer to Chapter 11 of "Building and Managing the Meta Data Repository", David Marco, Wiley 2000

time there was a burn (rocket firing) the error built up,” said Art Stephenson, Director of the Marshall Spaceflight Center and head of the NASA Investigation Team. “We entered the Mars atmosphere at a much lower altitude (than planned),” said Ed Weiler, NASA's chief scientist. “It (the spacecraft) either burned up in the Martian atmosphere or sped out (into space). We're not sure which happened.” This situation clearly illustrates a lack of meta data management. If NASA had a MME, they would have had meta data on the business rules programmed into the orbiter. The engineers would have known that it was programmed for metric measurements, rather than making incorrect assumptions. What was the cost of this lack of proper meta data management? The cost of this mission was over \$250 million. Who paid for this mistake? The U.S. taxpayer⁴.

Adhere to Regulatory Requirements

Every industry has its own dizzying array of regulatory requirements (see the next section of this chapter for more details on these requirements). Many companies look to MME architecture to assist them in meeting these requirements. Since most information stored in a company is stored electronically, a MME can greatly assist an organization in tracking, monitoring, and recording changes to its data and documents. Often this type of information is needed to meet regulatory requirements. Corporations must be able to meet these requirements, because the penalties can be severe. They range from large monetary fees to felony charges levied against the senior executives within the organization.

Enable Enterprise Applications

Without an enterprise-wide MME, your organization does not have a common understanding of the data across its systems. How do you build an enterprise-wide application (that is, a customer relationship management (CRM), a data warehousing, or an enterprise resource planning (ERP) system) without an enterprise-view understanding of the data? The answer for most companies is that they can't. This is why many research firms state that CRM initiatives fail over 90 percent of the time. In fact, Tom Siebel proclaimed that "There's no market for CRM . . . It's not there. There are a lot of companies trying to get into the space . . . But they are going after a market that doesn't exist."⁵

Realistically, this is not the case, because there is a clear business need to understand and manage all of a customer's *touch points*. So why does the man who marketed CRM the most and made the most money from it, claim that it is dead? Maybe it's because these CRM initiatives are highly complex and as a result have created a large number of expensive failures. These failures, along with a compressed IT economy, may have caused companies to abandon CRM. Personally, I don't expect Siebel Systems to stop making CRM software. On the other hand, I interpret Tom Siebel's comments to mean that we have to get ready for a new term for the same old thing.

In January 2000 article for *DM Review* ("Data Warehousing Trends for 2000"), I gave predictions for the future of the data-warehousing industry. In writing this article, I realized that the most important thing that I could do was to make sure that my predictions were accurate. Therefore, the first prediction was that most large government agencies and Global 2000 companies would spend a great deal of time and money rebuilding their data-warehousing investments. This was a simple and easy prediction to make. During the 1990s, corporations raced to build their data warehouses as quickly as they could. In their zeal to do this, many organizations neglected to build the architecture necessary to grow their systems over time. In many cases, these companies built "independent" data marts, which are directly sourced from operational systems without a data warehouse. These companies also neglected to implement a MME. Without a MME, it is exceedingly difficult, if not impossible, to maintain and enhance a data warehouse.

⁴ Story and corresponding quotes are from Associated Press Science Writer Paul Recer <http://www.anomalous-images.com/news/news537.html>

⁵ "The Siebel Observer," <http://www.siebelobserver.com/siebel/trends/vision.htm>

This is especially true if a company has implemented multiple, disparate data-warehousing systems. Having multiple data warehouses can lead to the following problems:

- Duplicate data
- Duplicate reports and queries
- Increased strain on operational batch windows
- Multiple versions of the truth

A MME addresses the excessive IT redundancies in this type of environment and acts as the “glue” that binds the data warehouses together. This makes MMEs even more vital in a cluttered environment.

Global 2000 companies spend *billions* of dollars analyzing existing data-warehousing systems for requirements and business rules, which have long since been forgotten. All of these requirements and business rules should have been initially managed in the MME. These same companies will continue to invest equally large expenditures into CRM and other enterprise-spanning applications. If you do not properly manage your meta data, you will have a very difficult time implementing enterprise-wide systems.

MMEs may not be as glamorous as CRM applications; however, these projects deliver real value as opposed to the 90 percent plus failure rate. Running a successful business is about concentrating single-mindedly on the fundamentals and then executing those fundamentals.

*Mr. Marco is an internationally recognized expert in the fields of enterprise architecture, data warehousing and business intelligence, and is the world's foremost authority on meta data. Mr. Marco is the author of several widely acclaimed books including “Universal Meta Data Models” and “Building and Managing the Meta Data Repository: A Full Life-Cycle Guide”. Mr. Marco has taught at the University of Chicago, DePaul University, and in 2004 he was selected to the prestigious **Crain's Chicago Business “Top 40 Under 40”**. He is the founder and President of EWSolutions, a GSA schedule and Chicago-headquartered strategic partner and systems integrator dedicated to providing companies and large government agencies with best-in-class business intelligence solutions using data warehousing, enterprise architecture and managed meta data environment technologies (www.EWSolutions.com). He may be reached directly via email at DMarco@EWSolutions.com*