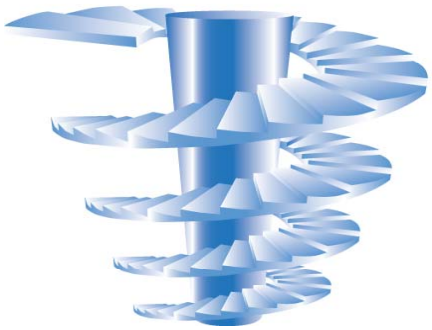
Meta Data Management: The Key Enabler of Data Integration



By David Marco

President

EWSolutions

EWSolutions' Background

EWSolutions is a Chicago-headquartered strategic partner and full life-cycle systems integrator providing both **award winning** strategic consulting and **full-service implementation services**. This combination affords our clients a full range of services for any size enterprise information management, managed meta data environment, and/or data warehouse/business intelligence initiative. Our notable client projects have been featured in the Chicago Tribune, Federal Computer Weekly, Crain's Chicago Business, and won the 2004 Intelligent Enterprise's RealWare award, 2007 Excellence in Information Integrity Award nomination and DM Review's 2005 World Class Solutions award.

Information Integrity Coalition

2007 Excellence in Information Integrity Award Nomination

Chicago Tribune

DM Review 2005

World Class Solutions Award Data Management

intelligent enterprise REAL wave Awards TRANSFORM

2004 WINNER

Best Business Intelligence Application
Information Integration
Client: Department of Defense

For more information on our Strategic Consulting Services, Implementation Services, or World-Class Training, call toll free at 866.EWS.1100, 866.397.1100, main number 630.920.0005 or email us at Info@EWSolutions.com

EWSolutions' Partial Client List

Arizona Supreme Court
Bank of Montreal
BankUnited
Basic American Foods
Becton, Dickinson and Company
Blue Cross Blue Shield companies
Branch Banking & Trust (BB&T)
British Petroleum (BP)
California DMV
California State Fund
College Board
Comcast
Corning Cable Systems
Countrywide Financial
Defense Logistics Agency (DLA)
Delta Dental
Department of Defense (DoD)
Driehaus Capital Management
Eli Lilly and Company
Environment Protection Agency
Federal Aviation Administration
Federal Bureau of Investigation (FBI)
Fidelity Information Services
Ford Motor Company
GlaxoSmithKline
Harris Bank
The Hartford
Harvard Pilgrim HealthCare
Health Care Services Corporation
Hewitt Associates
HP (Hewlett-Packard)
Information Resources Inc.
International Paper
Janus Mutual Funds
Johnson Controls
Key Bank
LiquidNet
Loyola Medical Center
Manulife Financial
Mayo Clinic
Microsoft
National City Bank
Nationwide
Neighborhood Health Plan
NORC
Physicians Mutual Insurance
Pillsbury
Quintiles
Sallie Mae
Schneider National
Secretary of Defense/Logistics
South Orange County Community College
SunTrust Bank
Target Corporation
The Regence Group
Thomson Multimedia (RCA)
United Health Group
United States Air Force
United States Army
United States Navy
United States Transportation Command
USAA
Wells Fargo
Wisconsin Department of Transportation
Zurich Cantonal Bank

GSA Schedule
Contract GS-35F-0453M

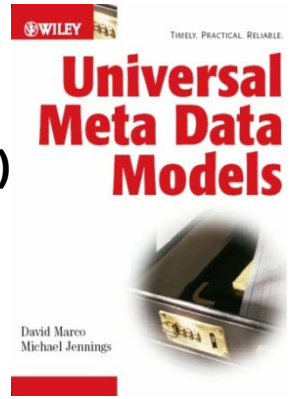
SeaPort-e

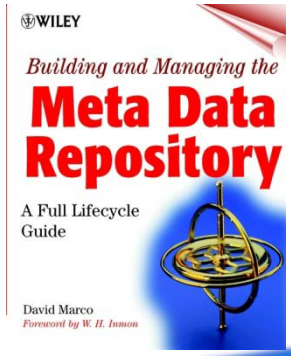
For more information on our Strategic Consulting Services, Implementation Services, or World-Class Training, call toll free at 866.EWS.1100, 866.397.1100, main number 630.920.0005 or email us at Info@EWSolutions.com

Professional Profile/Contact Information

Mr. Marco is an internationally recognized expert in the field of enterprise information management, data warehousing, Capability Maturity Model (CMM), business intelligence, and is the **world's foremost authority on meta data management**. Mr. Marco has authored several books including the widely acclaimed **"Universal Meta Data Models"** (Wiley, 2004) and the classic **"Building and Managing the Meta Data Repository: A Full Life-Cycle Guide"** (Wiley, 2000). These groundbreaking books have been broadly endorsed by many of the largest software companies in the industry and by several major magazines.

- Selected to the prestigious **2004 Crain's Chicago Business "Top 40 Under 40"**
- Crain's Chicago Business anointed him the "Melvil Dewey of Metadata"
- **2008 DAMA Data Management Hall of Fame (**Professional Achievement Award**)**
- Chairman of the Enterprise Information Management Institute (EIMInstitute.ORG)
- **2007 DePaul University named** him one of their **"Top 14 Alumni Under 40"**
- Presented hundreds of keynotes/seminars across four continents
- Published hundreds of articles on information technology
- Author of several best selling information technology books
- Taught at the **University of Chicago** and **DePaul University**



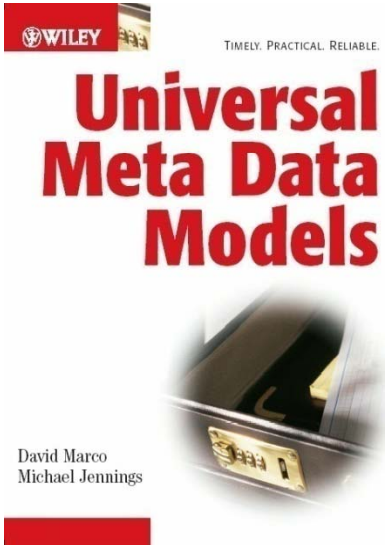


Email: DMarco@EWSolutions.com

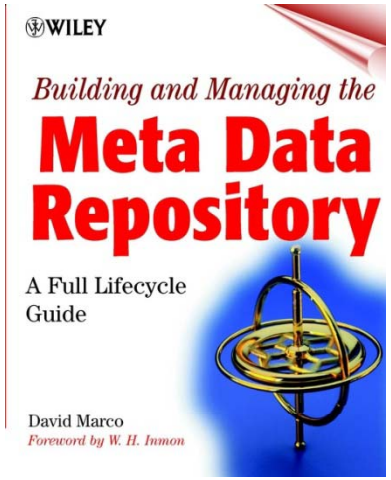
Acknowledgements

Session materials adapted from the books...

Universal Meta Data Models (Wiley, 2004)

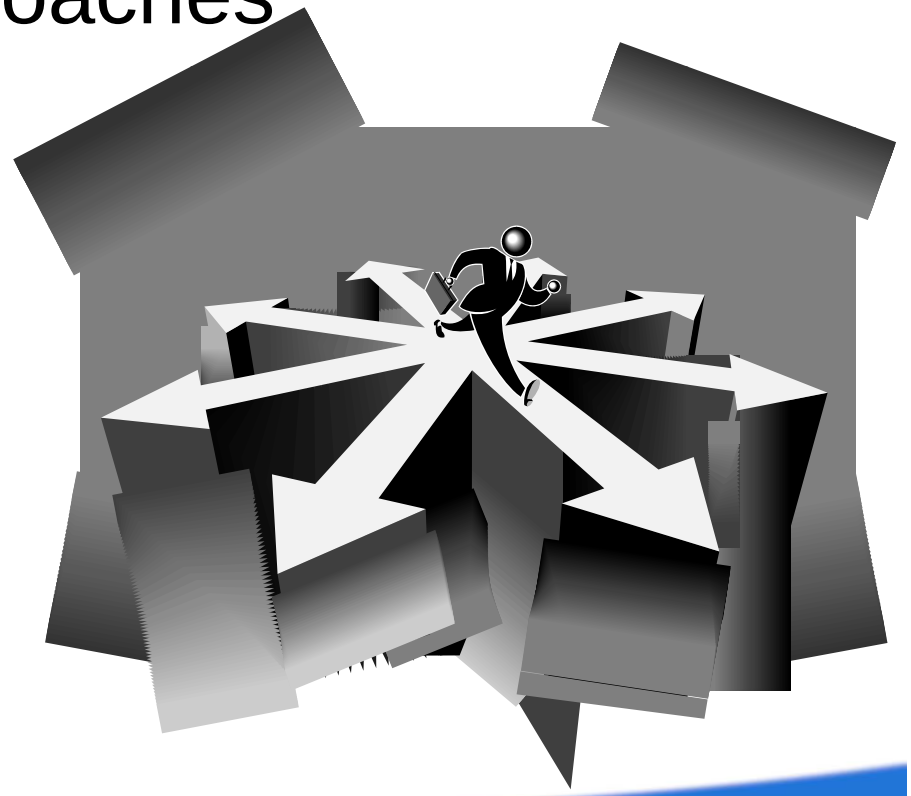


Building and Managing the Meta Data Repository: *A Full Life-Cycle Guide* (Wiley, 2000)



Agenda

- Data Integration Overview
- Data Integration Need
- Data Integration Approaches
- Business Meta Data
- Technical Meta Data
- Real-World Example

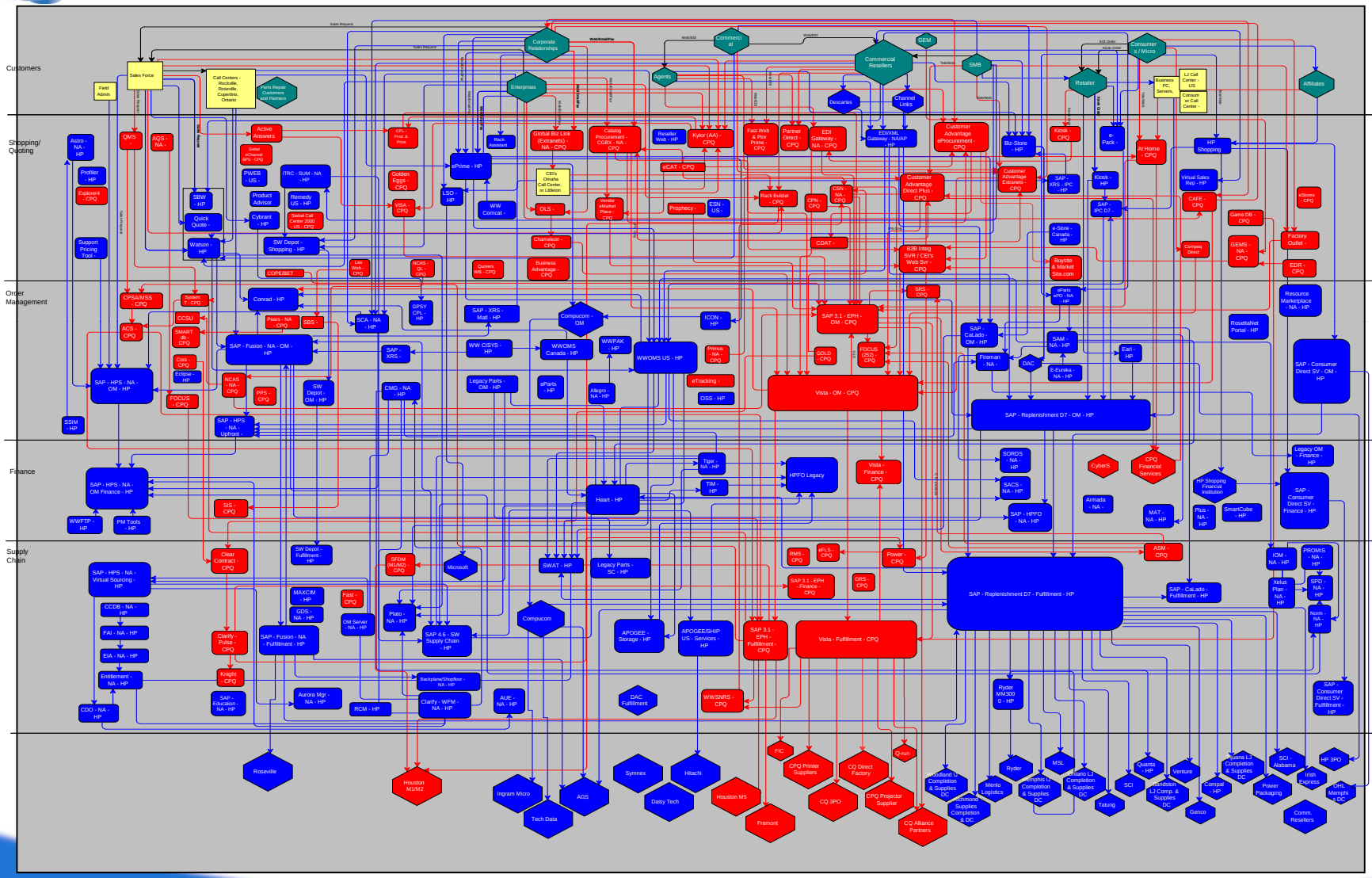


Data Integration

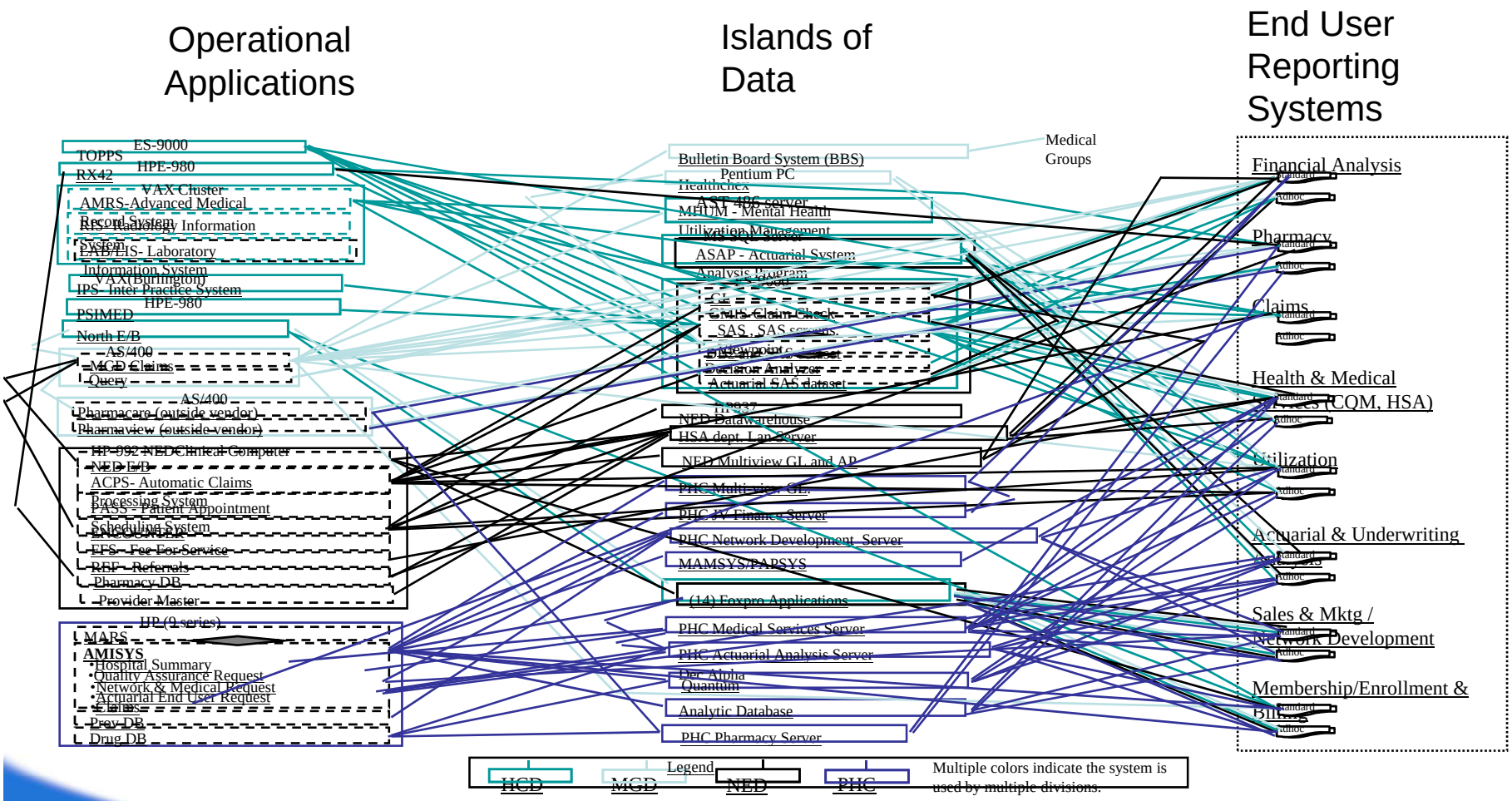
Data Integration

- Data integration is at the top, or near the top, of most every CIO's list of objectives
- Many CEOs and CFOs will mention the importance of data integration and the need for simplification of their IT portfolio
- Why has this issue become so important?

Fortune 100 – One Process



Typical System's Architecture



Need for Data Integration

- Exceedingly difficult and costly to manage system changes in this environment
- Inaccurate data
- Poor data quality
- Difficult to create enterprise applications (data warehouse, CRM, supply chain, etc.)
- Tremendous levels of needless redundancy:
  - Data
  - Process/Application
  - Hardware
  - Software
- Most Global 2000 companies and large government organizations have 3 – 4 fold needless data redundancy
- This needless redundancy costs money!

The Cost of Redundancy

- Large healthcare insurance company
- Has a $1.6 billion IT budget
- They estimate it costs them $2 per month to store each gigabyte of data
- $8 per month if you add in services and maintenance
- They estimate that they have 1.6 petabytes of redundant data
- What does this cost them yearly? Simple math
- $8 x 12 months x 1,000,000 (1.6 petabytes) = **$153,600,000**

How Organizations Approach Data Integration

Achieving Data Integration

Approach #1: Talk about the problem, but do nothing

- Lots of talks about the importance of data integration
- No financial backing to fix the problem
- Will not alter business processes
- Business as usual

Effectiveness

- Not effective at all
- Problem continues to proliferate

Achieving Data Integration

Approach #2: Fund a large effort to find systems to integrate

- The corporation will fund a large project, or series of projects to integrate systems that appear to be redundant
- No creation of a meta data repository to enable the data integration process
- No strategy/approach on how to ensure that duplication does happen again
- Will not alter business processes
- Business as usual

Effectiveness

- Is somewhat effective in the short-term as targeted duplication is removed
- Data integration problem continue to occur time and time again

Achieving Data Integration

Approach #3: Believing EAI (enterprise application integration) technology, on its own will solve the problem

- The corporation will fund a large EAI project
- They will not construct a meta data repository to manage the business rules and IT portfolio
- Will neglect to do business process re-engineering (BPR)
- No strategy/approach on how to ensure that duplication does happen again
- Will not alter business processes
- Business as usual

Effectiveness

- Not effective at all
- These companies are using message broker technology to create point-to-point interfaces
- Redundancy problem continues to proliferate

Achieving Data Integration

Approach #4: Implementing a managed meta data environment (MME) solution, data governance organization, along with migration efforts

- Implement a MME, its associated processes/procedures and conduct system consolidation efforts
- MME incorporates the business process changes required of the enterprise
- IT development will need to go through the MME for their development work
- Data governance will help define the processes and business meta data

Effectiveness

- Highly effective
- Requires a change in mentality
- Redundancy continues to shrink
- Redundancy is no longer occurring

Meta Data Management

What is Meta Data?

Meta Data Definition

All physical data (contained in software and other media) and knowledge (contained in employees and various media) from within and outside an organization, containing information about your company's physical data, industry, technical processes, and business processes.

Meta Data Is Knowledge

Managed Meta Data Environment

Managed Meta Data Environment (MME): The managed meta data environment represents the architectural components that are required to properly and systematically gather, retain and disseminate meta data throughout the enterprise.

Managed Meta Data Environment ROI

- Meta Data for the Business (business meta data)

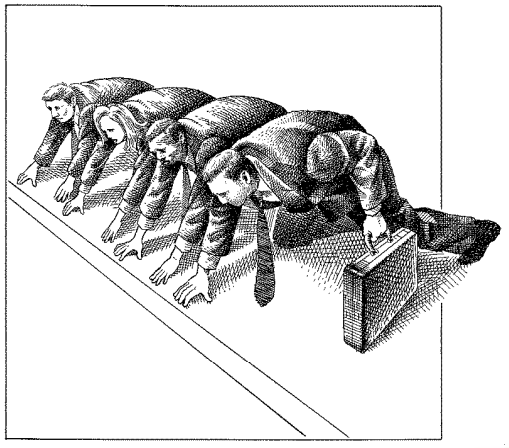
- Meta Data for the IT Department (technical meta data)

Business Meta Data

Managed Meta Data Environment ROI

Meta Data for the Business (business meta data)

- Provides the semantic layer between a company's systems (operational and business intelligence) and their business users



Meta Data for the Business

- Reduces training costs
- Makes strategic information (e.g. data warehousing, CRM, SCM, etc.) much more valuable as it aids analysts in making more profitable decisions
- Create **actionable** information
- Limits incorrect decisions
- Assists business analysts in finding the information they need, in a timely manner
- Bridges the gap between business users and IT professionals
- Increases confidence in the IT system data

Business Meta Data In Action

Meta Data for the Business

2007 Monthly Global Sales Report				February 7, 2008
Month	**Product Category**	**Sales $ (in thousands) U.S**	**Sales $ (in thousands International**	**Sales $ (in thousands Total**
December	TV	22,101	10,200	32,301
	VCR	11,190	4,300	15,490
	Cellular Phone	12,190	7,193	19,383
	Digital	4,002	1,301	5,303
	Miscellaneous	1,209	870	2,079
November	TV	42,000	22,200	64,200
	VCR	21,190	9,878	31,068
	Cellular Phone	28,193	12,193	40,386
	Digital	8,901	901	11,802
	Miscellaneous	2,730	0	4,260
October	TV	70,100	32,	103,050
	VCR	31,900	14,87	46,778
	Cellular Phone	41,700	17,550	59,250
	Digital	20,000	4,100	24,100
	Miscellaneous	4,850	2,850	7,700

"Sales $ U.S." is comprised of aggregated sales revenues from the United States, Canada, and Mexico, but does not subtract sales dollars from returned orders

Meta Data for the Business

2007 Monthly Global Sales Report				February 7, 2008
Month	**Product Category**	**Sales $ (in thousands) U.S**	**Sales $ (in thousands International**	**Sales $ (in thousands Total**
December	TV	22,101	10,200	32,301
	VCR	11,190	4,300	15,490
	Cellular Phone	12,190	7,193	19,383
	Digital	4,002	1,301	5,303
	Miscellaneous	1,209	870	2,079
November	TV	42,000	22,200	64,200
	VCR	21,190	9,878	31,068
	Cellular Phone	28,193	12,193	40,386
	Digital	8,901	2,901	11,802
	Miscellaneous	2,730	1,530	4,260
October	TV	70,100	32,950	103,050
	VCR	31,900	878	46,778
	Cellular Phone	41,700		59,250
	Digital	20,000	4	24,100
	Miscellaneous	4,850	2,8	7,700

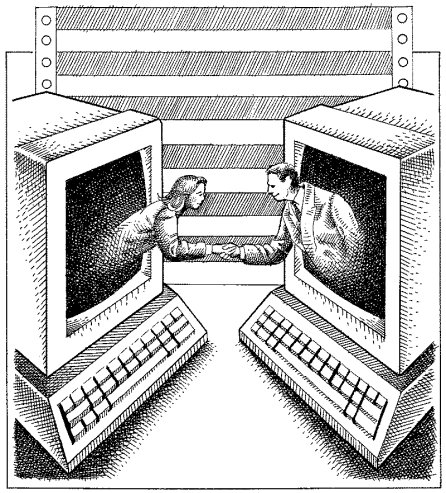
Information Quality Tracking Statistics
8.4% of the dollar values were not loaded
1.7% of the records were not loaded

Technical Meta Data

Managed Meta Data Environment ROI

Meta Data for the IT Department (technical meta data)

- Help IT departments better manage, maintain and grow their IT systems and assets



How Does a Lack of Meta Data Management Impact IT Development?

Case Study – NASA

Problem

- NASA has a history of financial mismanagement. “The agency’s contract-management function has earned a spot on the GAO’s “high risk” watch list every year since 1990
- In early 2004 NASA’s auditor (PricewaterhouseCoopers) proclaimed several issues with NASA’s 2003 financial statements
- “NASA couldn’t adequately document more than $565 billion – *billion* – in year end adjustments”
- Because of “the lack of a sufficient audit trail…it was not possible to complete further audit procedures”
- NASA has a $204 million line item called “Other” that “could not be explained or supported, indicating that NASA had not correctly reconciled its budgetary resources to its net cost of operations”
- NASA’s stated fund balance was $2 billion more than the balance in the treasury account
- NASA’s proposed 2005 budget is $16.244 billion (source: NASA)

Source: CFO magazine, “NASA, We have a problem”, May, 2004

Case Study – NASA

- Why does this problem exist?
- NASA says this problem is caused by enterprise software implementation called Integrated Financial Management Program (IFMP)
- NASA's CFO Gwendolyn Brown said the conversion to the new system caused the problem with the audit, specifically the agency had great difficultly converting the historical financial data from 10 legacy systems to the new system
- NASA has a "stovepipe" structure, in which each center behaves as an independent entity with a unique history and culture that is loath to brook "outside" interference from other parts of NASA
  - For example, NASA has 10 centers, each with a different financial reporting system
- "It's like a dozen dueling fiefdoms," says Keith Cowing, editor of NASA Watch

Meta Data for the IT Department

- **Dramatically reduces the probability of project failure**
- Speeds system's time-to-market
- Reduce system development life-cycle time
- Limit redundant data
- Limit redundant processes
- Managing IT portfolios
- Leverage work done by other teams
- Reduced rework
- Reduce research time
- Reduce unproductive work
- Lowers the impact of staff turnover

MME For Systems Consolidation

Meta Data for the IT Department

Systems Consolidation Report		BigCity Bank			August 15, 2008		
BigCity Bank				**Small Town Bank**			
Attribute Name	Attribute Definition	Entity Name	System Name	Attribute Name	Attribute Definition	Entity Name	System Name
Cust_Nbr	Cust_Nbr is the attribute of record for BigCity Bank customer numbers	Cust_Tbl	Central Customer System	CUSTNUM	Customer numbers from the deposit system.	CUSTTABLE	CUSTAPPL
				Purchase_No	Customer numbers from the purchase in the legacy deposit system	Purch_Tbl	CUSTSYS
				Borwr_No	Customer numbers from the loan system.	Borrower_File	LoanSys
Cust_Type	Cust_Type is the attribute of record for BigCity Bank customer types (affluent, upward, standard, high risk).	Cust_Tbl	Central Customer System	CUSTCDE	Customer types from the general ledger system.	GL_CUST	GLAPPL
Cust_Card_Ind	Cust_Card_Ind is the attribute of record for BigCity Bank customer ‘s that have a BCB credit card.	Cust_Tbl	Central Customer System		None applicable		
Cust_Crdt_Ratg	Cust_Crdt_Ratg is the attribute of record for BigCity Bank customer credit ratings (Superior Risk, Low Risk, Standard Risk, High Risk, Extreme Risk).	Cust_Tbl	Central Customer System	Credit_Rate	Customer rate is from the general ledger system and refers to the credit rating/worthiness of a customer.	GL_CUST	GLAPPL

Meta Data for the IT Department

Systems Consolidation Report		BigCity Bank			Small Town Bank		August 15, 2008
BigCity Bank							
Entity Name	Attribute Definition	Attribute Name	Domain Value	Transformation Rules	Attribute Name	Domain Value	Entity Name
Cust_Tbl	Cust_Type is the attribute of record for BigCity Bank customer types: 1 = affluent 2 = upward 3 = standard 4 = high risk	Cust_Type					
			1	Cust_Type = 1 WHEN CUSTCDE = 3 AND CUSTBAL > 500,000	CUSTCDE	3	GL_CUST
					CUSTBAL	High cardinality field	GL_CUST
			2	Cust_Type = 2 WHEN CUSTCDE = 4 AND CUSTBAL <= 500,000 AND CUSTBAL > 200,000	CUSTCDE	3	GL_CUST
					CUSTBAL	High cardinality field	GL_CUST
			3	Cust_Type = 3 WHEN CUSTCDE = 1 or 2 AND CUSTBAL <= 200,000 AND CUSTBAL > 75,000	CUSTCDE	3	GL_CUST
					CUSTBAL	High cardinality field	GL_CUST
			4	Cust_Type = 4 WHEN CUSTCDE = 0 AND CUSTBAL < 75,000 AND Credit_Rate < 22	CUSTCDE	3	GL_CUST
					CUSTBAL	High cardinality field	GL_CUST
					Credit_Rate	High cardinality field	GL_CUST
Cust_Card_Ind	Cust_Card_Ind is the attribute of record for BigCity Bank customer 's that have a BCB credit card.	Cust_Tbl					

Managed Meta Data Environment ROI

“We Build Systems To Manage Every Aspect Of Our Business, Except One To Manage The Systems Themselves.”

“A Managed Meta Data Environment Is A System That Manages Our Systems.”

Does This Really Work?

IT Portfolio Simplification Example

- RBC Financial Group is world-class in meta data management and has an award winning MME
- Let's compare RBC Financial Group to a large banking client of EWSolutions
- There is a tremendous amount of redundancy in our client's IT environment
- RBC has around 200 applications while our client has approximately 700 applications, although RBC's revenue is approximately double that of our client's ($16.1 billion as compared to $7.3 billion)
- **RBC's 2001 net income was $1.5 billion while our client's 2001 net income was $132 million**
- Obviously RBC's MME is a key contributor to their success

Don’t Limit Yourself

Questions