The Importance of Meta Data and Data Governance in Process and Data Management

By David Marco
President
EWSolutions
**EW Solutions’ Background**

**EW Solutions** 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, meta data management, data governance and 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.

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@EW Solutions.com
## EWSolutions’ Partial Client List

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

For more information on our Strategic Consulting Services, Implementation Services, or World-Class Training email us at Info@EWSolutions.com
Best known as the world’s foremost authority on metadata management, he is an internationally recognized expert in the fields of data warehousing, data governance and enterprise information management. In 2004 David Marco was named the “Melvil Dewey of Metadata” by Crain’s Chicago Business as he was selected to their very prestigious “Top 40 Under 40” list. David 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).

- Selected to the prestigious 2004 Crain’s Chicago Business “Top 40 Under 40”
- 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...


Marco Masters Series

- Teaching the full 3 day version of our Meta Data Management course from June 6 – 8, 2011 in Chicago, IL
- Teaching the full 3 day version of our Data Governance course from October 3 – 5, 2011 in Chicago, IL
- Visit www.MarcoMastersSeries.com for details
Fundamentals and Defining Terms
Enterprise Information Management (EIM): The systematic processes and governance procedures for applications, processes, data, and technology at a holistic enterprise perspective

The purpose of enterprise information management is to bring enterprise order, purpose, structure, efficiency, and performance to applications, processes, data, meta data and technology

EIM is not a single technology or component, but a coordinated framework of disciplines for managing data, meta data and information assets throughout the organization

Data Does Not Manage Itself!!
Data Management is the foundation for all of the other EIM focus areas. Regardless of which focus area you target first, you will need to do Data Management.
Process Management

- **Process Management**: is the application of knowledge, skills, tools, techniques and systems to define, visualize, measure, control, report and improve processes with the goal to meet customer requirements.

- Why do we need to manage processes?

- What does one process look like for a large corporation?
Process management: looks to holistically manage the Information Technology (IT) and business processes that exist within the organization in order to streamline, optimize, historically track, ensure quality and prevent redundancy of the IT processes at an enterprise level.

Process (people & technical) Management Target Areas
- What processes exist?
- What processes are used?
- What function does the process perform?
- Who uses the process?
- What data is utilized by the process?
- What inputs does the process get from which people?
- Which processes are used in which applications?
- Workflow (process dependencies and scheduling)
- Process performance

Vital for Service Oriented Architectures (SOA)
Process Management

- Why do we need to manage processes?
- What does one process look like for a large corporation?
Data Management: Data Resource Management is the development and execution of architectures, policies, practices and procedures that properly manage the full data lifecycle needs of an enterprise (DAMA International)
Data Management:
is the function of managing the data within an organization
This is the keystone focus area
Without this area it is almost impossible to address the other focus areas
Business requirements will drive this initiative
Data Management Objectives

- Data Management looks to answer questions on the data in a company:
  - What does it mean (data lineage)?
  - What is its source (data heritage)?
  - What are the valid values?
  - What formulas were used to calculate it?
  - What are its business rules?
  - What are its technical rules?

- Subject area definition
- Define business entities
- Enterprise conceptual model
- Enterprise logical model
Maybe data management looks better than process management?
Clearly we can see that there is a problem

Question: What are the key disciplines to resolve this problem?

Answer: Data Governance (people) and Meta Data Management (technology)
Data Governance
Background & Fundamentals
Data Governance Defined

- **Data Governance**: defines the people, processes, framework and organization necessary to ensure that an organization’s information assets (data and meta data) are formally, properly, proactively and efficiently managed throughout the enterprise to secure its trust, accountability, meaning and accuracy.
Understanding Data Governance

Data Governance

Policies, Procedures, Consensus, Knowledge, Information, Data, Meta Data

Data Stewards

Data Stewards

Data

Misunderstood
Inaccurate
Misleading

I need to make profitable decisions
I don’t know what I’m looking at

Transform Data into Information

Understood
Accurate
Consistent

Actionable Information

www.EWSolutions.com
© 2011 Enterprise Warehousing Solutions, Inc. (EWSolutions) – 22
How Do You Manage Information Assets?

This is all Data Governance

You cannot manage what you do not measure

You cannot measure what you do not understand

You do not understand…….
Data Governance ROI
Data Governance Metrics

- Defined “hard” and “soft” dollar savings and earnings
- Testimonials from participating areas (lines-of-business, divisions, etc.)
- Improvements in process and data performance (limiting redundancy, increasing reuse, improving performance, etc.)
- Documented improvements in information quality
- Number of times meta data is read/updated/added/deleted from the MME
- Number of participating Data Stewards
- Number of defined Subject Areas
- Number of entities, attributes and relationships actively managed
- Number of entities, attributes and relationships with corresponding meta data
Large healthcare insurance company
Has a $1.6 billion IT budget
They estimate it costs them $2 per month to store each gigabytes 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 Does a Lack of Data Governance 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 difficulty 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
Information As A Corporate Asset
“Information and knowledge are the primary resources of the knowledge society of the 21st century.” P. Drucker, 1992

“Organizations that do not understand the overwhelming importance of managing data and information as tangible assets in the new economy will not survive.” T. Peters, 2001
Data Governance Components

- **Thought Ware**: Mission/Core Values, Goals/Objectives, Charters/Principles, Critical Success Factors, Plans, Documents/Policies, Communication Plan (messages and vehicles), Roles/Functions and Responsibilities Definitions, Accountability Matrix, Organizational Interdependencies, Workflow

- **People Ware**: Structures, Organizations, Committees, Teams/Groups, People

- **Work Ware**: Managed Meta Data Environment, Software, Training and Education, References, Templates, Standards

- **Artifacts**: Meta Data, Data Rules and Definitions, Decision Rights, Accountabilities, Controls
Illustration of the Four Components

- Thought Ware
  - Guides
- People Ware
  - Assists
- Work Ware
- Artifacts

The true tangible value/measure of DG – when the artifacts are used successfully.

Kept in Tools

Work Results
Data governance is the method for connecting information management and the corporate business strategy.
Data Governance Organization
Every organization forms their data governance organization a little differently.

Some have a more or less complex organization.

What is critical is that the organization:

- is actively using the MME
- has clear lines of communication
- has a defined and well understood decision making process
- well defined feedback loop
Subject Area

- A logical grouping of items of interest to the enterprise, or areas of interest within the company.
- About 10 – 20 Subject Areas in an organization.
- The “nouns of an entity. Examples:
  - Legal Entity
  - Cost Center
  - Account
  - Product
  - Customer
  - Sale
Data Governance Organization

**Subject Area Groups**
- **Subject Area User Group #1**
  - Chief Steward
  - Business Steward(s)
  - Technical Steward(s)
  - Interested Parties
- **Subject Area User Group #2**
  - Chief Steward
  - Business Steward(s)
  - Technical Steward(s)
  - Interested Parties
- **Subject Area User Group #3**
  - Chief Steward
  - Business Steward(s)
  - Technical Steward(s)
  - Interested Parties

**Data Stewardship Coordination Group**
- Program Manager
- Chief Stewards
- EIM Focus Area/Project #1 Steward Team
- EIM Focus Area/Project #2 Steward Team
- EIM Focus Area/Project #3 Steward Team

**Enterprise Oversight**
- Data Governance Council
  - Members:
    - Executive Sponsor(s)
    - Program Manager
    - Chief Stewards
    - CIO
    - Key business staff
    - Key IT staff
  - Policies, Procedures, Standards, etc.

**Managed Meta Data Environment**
- Information Technology
  - Data Custodian Team
  - Technical Stewards

**Recommendations**
- Requirements

www.EWSolutions.com
© 2011 Enterprise Warehousing Solutions, Inc. (EWSolutions) – 38

Strategic Partner & Systems Integrator
Intelligent Business Intelligence™
Data Governance Organizations

- **Chief Stewards**
- **Business Data Stewards**
- **Technical Data Stewards**

- **Data Governance Program Mgr.**
- Various Data Stewards

- **Data Governance Executive Committee**
  - Data Governance Program Mgr.
  - Various Data Stewards

- **Data Stewardship Groups**
  - Data Governance Program Mgr.
  - Chief Stewards
  - Business Stewards
  - Technical Data Stewards

- **Subject Area Groups**
  - Chief Stewards
  - Business Data Stewards
  - Technical Data Stewards

Operational – by Subject Area, not by LOB

80-85% Conflicts Resolved at this level

Strategic

Tactical

< 20%

< 5%

Escalation Path
Meta Data Management
Meta Data vs. Data

- **Meta Data:** Meta data contains the knowledge that a 1) field is called “Customer_Name”, is 40 characters in length, and exists in systems A, B, and C; 2) that our company has 3 systems which contain customer master data. These systems are…

- **Data:** Data would be a specific instance of “Customer_Name” equaling “John Doe”

- **Information:** Data that is meaningful to a business user. They understand it and they know what to do with it
Information = Data + Meta Data

(content)     (context)
**Meta Data**

By definition meta data is

1. “Data about data”.
2. “Everything that data is not”.

*So What!!*
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
“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”
Managed Meta Data Environment ROI

- Meta Data for the Business (business meta data)

- Meta Data for the IT Department (technical meta data)
Intel finds huge ROI in managing meta data

- Estimates $6 in savings for every $1 spent
  - Emphasis on reducing developer’s average research time of 30%
  - Uses “Centralized” architecture
  - Key to success is based on regular, frequent “scan updates”

**Source: Computer World magazine – July 2005**

A Canadian government agency achieved:

- More than 90 percent reuse of existing data definitions
- 85 percent improvement in application integration
- 25 percent reduction in DW analysis and design
- Impact analysis study in 2 hours vs. 36 person-days of consulting

“Knowledge workers spend up to 2.5 hours each day looking for information… but find what they are looking for only 40% of the time.”

Kit Sims Taylor – Economist and Researcher in “Economics of information and networks”
Bellevue Community College, WA
Business Meta Data
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
Application for Employment

[Blank lines]

1. ____________________________  2. ____________________________

[Blank lines]
Example with Meta Data

Application for Employment

PERSONAL INFORMATION

Last Name | First Name | Initials

Present Address: Street, apt. | City | Province | Postal Code

Permanent Address: (leave blank if same as present address) | City | Province | Postal Code

Telephone: Residence: ______________ | Business: ______________ | Other: ______________

TYPE OF WORK OR POSITION APPLIED FOR

1. __________________________  2. __________________________

Salary expectations: __________________________ | Date Available: __________________________

Type of employment sought:

- [ ] Full Time
- [ ] Regular Part-Time
- [ ] Casual/Contract
- [ ] Summer
- [ ] Co-op
- [ ] Internship
Meta Data Providing the Semantic Layer
1. “Global Sales by Month”
This report shows a years worth of U.S., international, and Totals, of summarized sales figures by product category, on a monthly basis.

2. “Global Sales by Region, by Month”
This report shows a years worth of U.S., international, and Totals, of summarized sales figures by product category, on a monthly basis by region.

3. “Global Product Sales by Region, by Month”
This report shows a years worth of U.S., international, and Totals, of detailed product sales figures, on a monthly basis by region.
### 2007 Monthly Global Sales Report

<table>
<thead>
<tr>
<th>Month</th>
<th>Product Category</th>
<th>Sales $ (in thousands)</th>
<th>Sales $ (in thousands)</th>
<th>Sales $ (in thousands)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>U.S</td>
<td>International</td>
<td>Total</td>
</tr>
<tr>
<td>December</td>
<td>TV</td>
<td>22,101</td>
<td>10,200</td>
<td>32,301</td>
</tr>
<tr>
<td></td>
<td>VCR</td>
<td>11,190</td>
<td>4,300</td>
<td>15,490</td>
</tr>
<tr>
<td></td>
<td>Cellular Phone</td>
<td>12,190</td>
<td>7,193</td>
<td>19,383</td>
</tr>
<tr>
<td></td>
<td>Digital</td>
<td>4,002</td>
<td>1,301</td>
<td>5,303</td>
</tr>
<tr>
<td></td>
<td>Miscellaneous</td>
<td>1,209</td>
<td>870</td>
<td>2,079</td>
</tr>
<tr>
<td>November</td>
<td>TV</td>
<td>42,000</td>
<td>22,200</td>
<td>64,200</td>
</tr>
<tr>
<td></td>
<td>VCR</td>
<td>21,190</td>
<td>9,878</td>
<td>31,068</td>
</tr>
<tr>
<td></td>
<td>Cellular Phone</td>
<td>28,193</td>
<td>12,193</td>
<td>40,386</td>
</tr>
<tr>
<td></td>
<td>Digital</td>
<td>8,901</td>
<td>3,901</td>
<td>11,802</td>
</tr>
<tr>
<td></td>
<td>Miscellaneous</td>
<td>2,730</td>
<td>0</td>
<td>4,260</td>
</tr>
<tr>
<td>October</td>
<td>TV</td>
<td>70,100</td>
<td>32,100</td>
<td>103,050</td>
</tr>
<tr>
<td></td>
<td>VCR</td>
<td>31,900</td>
<td>14,800</td>
<td>46,778</td>
</tr>
<tr>
<td></td>
<td>Cellular Phone</td>
<td>41,700</td>
<td>17,550</td>
<td>59,250</td>
</tr>
<tr>
<td></td>
<td>Digital</td>
<td>20,000</td>
<td>4,100</td>
<td>24,100</td>
</tr>
<tr>
<td></td>
<td>Miscellaneous</td>
<td>4,850</td>
<td>2,850</td>
<td>7,700</td>
</tr>
</tbody>
</table>

“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.
# Carrier/Usage Summary Report

<table>
<thead>
<tr>
<th>Month</th>
<th>Carrier Name</th>
<th>Usage Type</th>
<th>Discounted Usage (M seconds)</th>
<th>Regular Usage (M seconds)</th>
<th>Total Usage (M seconds)</th>
</tr>
</thead>
<tbody>
<tr>
<td>October</td>
<td>BigTeleCo</td>
<td>Long Distance</td>
<td>7,201</td>
<td>4,288</td>
<td>11,489</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Regular Usage</td>
<td>72,033</td>
<td>42,000</td>
<td>114,033</td>
</tr>
<tr>
<td></td>
<td>TeleBell</td>
<td>Long Distance</td>
<td>630</td>
<td>777</td>
<td>1,407</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Regular Usage</td>
<td>23,000</td>
<td>17,255</td>
<td>40,255</td>
</tr>
<tr>
<td></td>
<td>NewBell</td>
<td>Long Distance</td>
<td>220</td>
<td>530</td>
<td>310</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Regular Usage</td>
<td>1,100</td>
<td>757</td>
<td>1,857</td>
</tr>
<tr>
<td>September</td>
<td>BigTeleCo</td>
<td>Long Distance</td>
<td>6,400</td>
<td>4,000</td>
<td>10,400</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Regular Usage</td>
<td>73,450</td>
<td>42,702</td>
<td>116,152</td>
</tr>
<tr>
<td></td>
<td>TeleBell</td>
<td>Long Distance</td>
<td>645</td>
<td>750</td>
<td>1,395</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Regular Usage</td>
<td>23,500</td>
<td>17,923</td>
<td>41,423</td>
</tr>
<tr>
<td></td>
<td>NewBell</td>
<td>Long Distance</td>
<td>124</td>
<td>175</td>
<td>299</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Regular Usage</td>
<td>1,175</td>
<td>703</td>
<td>1,878</td>
</tr>
<tr>
<td>August</td>
<td>BigTeleCo</td>
<td>Long Distance</td>
<td>6,220</td>
<td>4,010</td>
<td>10,230</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Local</td>
<td>71,207</td>
<td>41,918</td>
<td>113,125</td>
</tr>
<tr>
<td></td>
<td>TeleBell</td>
<td>Long Distance</td>
<td>652</td>
<td>754</td>
<td>1,406</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Local</td>
<td>20,010</td>
<td>15,500</td>
<td>35,510</td>
</tr>
<tr>
<td></td>
<td>NewBell</td>
<td>Long Distance</td>
<td>92</td>
<td>110</td>
<td>202</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Local</td>
<td>1,177</td>
<td>708</td>
<td>1,885</td>
</tr>
</tbody>
</table>

**Discounted Usage:** Any local or long distance phone usage that has a discount applied to it. Discounts include **non-prime**, **holiday** and **rate specials**.

**Non-Prime Usage:** Any local or long distance phone usage that occurs between the time of 8:00pm – 7:00am.

Last Updated: 3/31/2004, Bob Jones

Last Updated: 1/08/2005, Tony Ragone
Meta Data Makes for Better Decisions
## Meta Data for the Business

### 2007 Monthly Global Sales Report

<table>
<thead>
<tr>
<th>Month</th>
<th>Product Category</th>
<th>Sales $ (in thousands)</th>
<th>Sales $ (in thousands)</th>
<th>Sales $ (in thousands)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>U.S</td>
<td>International</td>
<td>Total</td>
</tr>
<tr>
<td>December</td>
<td>TV</td>
<td>22,101</td>
<td>10,200</td>
<td>32,301</td>
</tr>
<tr>
<td></td>
<td>VCR</td>
<td>11,190</td>
<td>4,300</td>
<td>15,490</td>
</tr>
<tr>
<td></td>
<td>Cellular Phone</td>
<td>12,190</td>
<td>7,193</td>
<td>19,383</td>
</tr>
<tr>
<td></td>
<td>Digital</td>
<td>4,002</td>
<td>1,301</td>
<td>5,303</td>
</tr>
<tr>
<td></td>
<td>Miscellaneous</td>
<td>1,209</td>
<td>870</td>
<td>2,079</td>
</tr>
<tr>
<td>November</td>
<td>TV</td>
<td>42,000</td>
<td>22,200</td>
<td>64,200</td>
</tr>
<tr>
<td></td>
<td>VCR</td>
<td>21,190</td>
<td>9,878</td>
<td>31,068</td>
</tr>
<tr>
<td></td>
<td>Cellular Phone</td>
<td>28,193</td>
<td>12,193</td>
<td>40,386</td>
</tr>
<tr>
<td></td>
<td>Digital</td>
<td>8,901</td>
<td>2,901</td>
<td>11,802</td>
</tr>
<tr>
<td></td>
<td>Miscellaneous</td>
<td>2,730</td>
<td>1,530</td>
<td>4,260</td>
</tr>
<tr>
<td>October</td>
<td>TV</td>
<td>70,100</td>
<td>32,950</td>
<td>103,050</td>
</tr>
<tr>
<td></td>
<td>VCR</td>
<td>31,900</td>
<td>3,778</td>
<td>46,778</td>
</tr>
<tr>
<td></td>
<td>Cellular Phone</td>
<td>41,700</td>
<td>6,300</td>
<td>59,250</td>
</tr>
<tr>
<td></td>
<td>Digital</td>
<td>20,000</td>
<td>2,870</td>
<td>24,100</td>
</tr>
<tr>
<td></td>
<td>Miscellaneous</td>
<td>4,850</td>
<td>2,870</td>
<td>7,700</td>
</tr>
</tbody>
</table>

**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
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
Technical Impact Analysis
### Impact Analysis Report

#### Question:
Show all decision support tables/files, programs, and fields impacted by a change to the “CUST” table in the “Order Entry” system

<table>
<thead>
<tr>
<th>Source System</th>
<th>Source Table</th>
<th>Impact Field</th>
<th>Program Impacted</th>
<th>Tables/Files Impacted</th>
<th>Table Type</th>
<th>Fields Impacted</th>
</tr>
</thead>
<tbody>
<tr>
<td>Order Entry</td>
<td>CUST</td>
<td>Customer_Name</td>
<td>CUSTOMER_PR02</td>
<td>DW_CUSTOMER</td>
<td>T</td>
<td>Cust_Name_First</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CUSTOMER_PR01</td>
<td>I02_CUSTOMER</td>
<td>I</td>
<td>Cust_Name_Middle</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Cust_Name_Last</td>
</tr>
<tr>
<td>Customer_Addr</td>
<td>CUSTOMER_PR02</td>
<td>Cust_Name_Address</td>
<td></td>
<td></td>
<td></td>
<td>Cust_Name_Middle</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cust_Name_City</td>
<td></td>
<td></td>
<td></td>
<td>Cust_Name_State</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cust_Name_State</td>
<td></td>
<td></td>
<td></td>
<td>Cust_Name_Zip</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cust_Name_Zip</td>
<td></td>
<td></td>
<td></td>
<td>Cust_Name_Address</td>
</tr>
</tbody>
</table>

*Legend*

“T” = Target  
“I” = Intermediate  
“S” = Source
Question: Show all systems, tables/files, fields, and their domains impacted by a change to the length of all occurrences of the Customer_Name field

<table>
<thead>
<tr>
<th>Field</th>
<th>System</th>
<th>Tables/Files</th>
<th>Fields</th>
<th>Domain</th>
</tr>
</thead>
<tbody>
<tr>
<td>Customer Name</td>
<td>Order Entry</td>
<td>CUSTOMER_BILL_TO</td>
<td>CUST_NAME</td>
<td>Alphanumeric 20</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CUSTOMER_SELL_TO</td>
<td>CUST_NAME</td>
<td>Alphanumeric 20</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CUSTOMER_SHIP_TO</td>
<td>CUST_NAME</td>
<td>Alphanumeric 20</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ORDER_HEADER</td>
<td>CUST_NAME</td>
<td>Alphanumeric 20</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ORDER_DETAIL</td>
<td>CUST_NAME</td>
<td>Alphanumeric 20</td>
</tr>
<tr>
<td>General Ledger</td>
<td>CUSTOMER</td>
<td></td>
<td>Cust_Name</td>
<td>Alphanumeric 35</td>
</tr>
<tr>
<td></td>
<td>EXPENSES</td>
<td></td>
<td>Cust_Name</td>
<td>Alphanumeric 35</td>
</tr>
<tr>
<td></td>
<td>DW_CUSTOMER</td>
<td></td>
<td>Cust_Name</td>
<td>Alphanumeric 20</td>
</tr>
<tr>
<td></td>
<td>I01_CUSTOMER</td>
<td></td>
<td>Cust_Name</td>
<td>Alphanumeric 20</td>
</tr>
<tr>
<td></td>
<td>I02_CUSTOMER</td>
<td></td>
<td>Cust_Name</td>
<td>Alphanumeric 20</td>
</tr>
<tr>
<td></td>
<td>I03_CUSTOMER</td>
<td></td>
<td>Cust_Name</td>
<td>Alphanumeric 20</td>
</tr>
<tr>
<td>Data Mart - Marketing</td>
<td>DM_CUSTOMER</td>
<td></td>
<td>DM_Cust_Name</td>
<td>Alphanumeric 20</td>
</tr>
<tr>
<td></td>
<td>I01_DM_CUSTOMER</td>
<td></td>
<td>DM_Cust_Name</td>
<td>Alphanumeric 20</td>
</tr>
<tr>
<td></td>
<td>I02_DM_CUSTOMER</td>
<td></td>
<td>DM_Cust_Name</td>
<td>Alphanumeric 20</td>
</tr>
</tbody>
</table>
MME For Systems Consolidation
<table>
<thead>
<tr>
<th>Attribute Name</th>
<th>Attribute Definition</th>
<th>Entity Name</th>
<th>System Name</th>
<th>Attribute Name</th>
<th>Attribute Definition</th>
<th>Entity Name</th>
<th>System Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cust_Nbr</td>
<td>Cust_Nbr is the attribute of record for BigCity Bank customer numbers</td>
<td>Cust_Tbl</td>
<td>Central Customer System</td>
<td>CUSTNUM</td>
<td>Customer numbers from the deposit system.</td>
<td>Cust TABLE</td>
<td>CUSTAPPL</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Purchase_No</td>
<td>Customer numbers from the purchase in the legacy deposit system</td>
<td>Purch_Tbl</td>
<td>CUSTSYS</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Borwr_No</td>
<td>Customer numbers from the loan system.</td>
<td>Borrower_File</td>
<td>LoanSys</td>
</tr>
<tr>
<td>Cust_Type</td>
<td>Cust_Type is the attribute of record for BigCity Bank customer types (affluent, upward, standard, high risk).</td>
<td>Cust_Tbl</td>
<td>Central Customer System</td>
<td>CUSTCDE</td>
<td>Customer types from the general ledger system.</td>
<td>GL_CUST</td>
<td>GLAPPL</td>
</tr>
<tr>
<td>Cust_Card_Ind</td>
<td>Cust_Card_Ind is the attribute of record for BigCity Bank customer ’s that have a BCB credit card.</td>
<td>Cust_Tbl</td>
<td>Central Customer System</td>
<td>None applicable</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cust_Crdt_Ratg</td>
<td>Cust_Crdt_Ratg is the attribute of record for BigCity Bank customer credit ratings (Superior Risk, Low Risk, Standard Risk, High Risk, Extreme Risk).</td>
<td>Cust_Tbl</td>
<td>Central Customer System</td>
<td>Credit_Rate</td>
<td>Customer rate is from the general ledger system and refers to the credit rating/worthiness of a customer.</td>
<td>GL_CUST</td>
<td>GLAPPL</td>
</tr>
<tr>
<td>Entity Name</td>
<td>Attribute Definition</td>
<td>Domain Value</td>
<td>Transformation Rules</td>
<td>Attribute Name</td>
<td>Domain Value</td>
<td>Entity Name</td>
<td></td>
</tr>
<tr>
<td>-------------</td>
<td>--------------------------------------------------------------------------------------</td>
<td>--------------</td>
<td>--------------------------------------------------------------------------------------</td>
<td>----------------</td>
<td>--------------</td>
<td>-------------</td>
<td></td>
</tr>
<tr>
<td>Cust_Tbl</td>
<td>Cust_Type is the attribute of record for BigCity Bank customer types: 1 = affluent</td>
<td>Cust_Type</td>
<td>1 Cust_Type = 1 WHEN CUSTCDE = 3 AND CUSTBAL &gt; 500,000</td>
<td>CUSTCDE</td>
<td>3</td>
<td>GL_CUST</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2 = upward</td>
<td></td>
<td>2 Cust_Type = 2 WHEN CUSTCDE = 4 AND CUSTBAL &lt;= 500,000 AND CUSTBAL &gt; 200,000</td>
<td>CUSTCDE</td>
<td>3</td>
<td>GL_CUST</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3 = standard</td>
<td></td>
<td>3 Cust_Type = 3 WHEN CUSTCDE = 1 or 2 AND CUSTBAL &lt;= 200,000 AND CUSTBAL &gt; 75,000</td>
<td>CUSTCDE</td>
<td>3</td>
<td>GL_CUST</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4 = high risk</td>
<td></td>
<td>4 Cust_Type = 4 WHEN CUSTCDE = 0 AND CUSTBAL &lt; 75,000 AND Credit_Rate &lt; 22</td>
<td>CUSTCDE</td>
<td>3</td>
<td>GL_CUST</td>
<td></td>
</tr>
<tr>
<td>Cust_Card_Ind</td>
<td>Cust_Card_Ind is the attribute of record for BigCity Bank customer’s that have a BCB credit card.</td>
<td>Cust_Tbl</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Managed Meta Data Environment
When implementing a meta data management system there is a lot more to it than just a meta data repository

**Important:** The MME is an operational system, not a data warehouse
Managed Meta Data Environment (MME): The managed meta data environment represents the architectural components, people and processes that are required to properly and systematically gather, retain and disseminate meta data throughout the enterprise.
Managed Meta Data Environment

Meta Data Sourcing Layer

- Software Tools
- Documents/Spreadsheets
- Messaging/Transactions (EAI, web services, XML, etc.)
- Applications (CRM, ERP, data warehouses, etc.)
- Websites/E-Commerce
- Third Parties (business partners, vendors, customers, government agencies)

Meta Data Management Layer

- Meta Data Repository

Meta Data Delivery Layer

- Meta Data Marts
- End Users (business and technical)
- Applications (CRM, ERP, etc.)
- Websites/E-Commerce
- Third Parties (vendors, customers, government agencies)
- Messaging/Transactions (EAI, web services, XML, etc.)
- Business Users
- End Users (business and technical)
- Software Tools

www.EWSolutions.com © 2011 Enterprise Warehousing Solutions, Inc. (EWSolutions) – 72
Meta Data Management in Governance

- Stewards manage data (instances of data values) and meta data (information concerning the data)
- Meta data management is the key technical enabler for the performance of successful governance
- Difficult to do governance successfully without managed meta data environment
“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.”
Understanding the Big Picture
The DAMA-DMBOK Framework

Version 3

Data Quality Management
Data Architecture Management
Data Development
Database Operations Management
Data Security Management
Data Warehousing & Business Intelligence Management
Reference & Master Data Management
Document & Content Management
Meta Data Management
Data Governance

EIM Maturity Model

**Level 1: Informal Processes**
- Redundant, undocumented data.
- Disparate databases without architecture.
- Little or no business meta data.
- Diverging semantics.
- Minimal data integration.
- Minimal data cleansing.
- Dependent on a few skilled individuals.
- Responsibilities assigned across separate IT groups.
- Few defined IT roles.
- Some commonly used approaches but with no commitment to their use.
- Some management awareness, but no enterprise-wide buy-in.
- Little or no business involvement, no defined business roles.
- General purpose tools used as point solutions.
- Reactive monitoring and problem solving.
- Data regarded as a minor by-product of business activity, with no estimated business impact.

**Level 2: Emerging Processes**
- Growing intuitive executive awareness of the value of data assets in some business areas.
- Initial forays in data stewardship and governance but roles are unclear and not ongoing.
- Initial efforts to implement enterprise-wide management, but with contention across groups with differing perspectives.
- New skills requirements are recognized and addressed with training.
- Enterprise architecture and MME projects underway.
- Data Distribution Services are deployed as strategic solutions.
- Some processes are repeatable.

**Level 3: Engineered Processes**
- Active executive involvement across the enterprise.
- Ongoing, clearly defined business data stewardship.
- Central EDM organization.
- Standard processes, metrics, and tools used enterprise wide.
- Enterprise data architecture guides implementations.
- Centralized meta data management.
- Quality SLA’s are defined and monitored regularly.
- Commitment to continual skills development.
- Periodic audits and proactive monitoring.

**Level 4: Controlled Processes**
- Measurable process goals are established for each defined process.
- Measurements are collected and analyzed.
- Quantitative (measurement) analysis of each process occurs.
- Beginning to predict future performance.
- Defects are proactively identified and corrected.
- Quantitative and qualitative understanding used to continually improve each process.
- Value is monitored continuously.
- Understanding of how each process contributes to the business strategies and goals of the enterprise.

**Level 5: Optimizing Processes**
- Quantitative and qualitative understanding used to continually improve each process.
- Value is monitored continuously.
- Understanding of how each process contributes to the business strategies and goals of the enterprise.

---

www.EWSolutions.com
© 2011 Enterprise Warehousing Solutions, Inc. (EWSolutions) – 79

2010 © EWSolutions
All Rights Reserved
Conclusion
Conclusion

- Data & Process Management are enterprise activities that address the structured management of our informational systems.
- Data governance & Meta Data Management are foundational activities:
  - Master Data Management
  - 360 degree view of customer
  - Cross sell
- These disciplines are not easy…but what enterprise activity is?
- Focus on the best practices
- Stay focused
- You WILL have a big success!!