Data Governance & Stewardship

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, 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, Journal of the American Medical Informatics Association (JAMIA), 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, email us at info@EWSolutions.com or call at 630.920.0005
### EWSolutions’ Partial Client List

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

- 2014 EWSolutions was inducted into the Hinsdale business Hall-of-Fame after 6 consecutive years of receiving “Best of” awards in Enterprise Information Management
- Selected to the prestigious 2004 Crain’s Chicago Business “Top 40 Under 40”
- 2008 DAMA Data Management Hall of Fame (Professional Achievement Award)
- 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
- Holds both a CDMP and a CBIP certification

Email: DMarco@EWSolutions.com
Agenda

- Enterprise Information Management’s (EIM) Role in Data Governance
- Data Governance Return on Investment (ROI)
- Introduction to Data Governance
- Introduction to Data Stewardship
- Data Governance Framework
- Top 10 Reasons for Failure
Your Data Governance Program

- No two governance and stewardship programs are the same – each is unique
- Adopt as many of the best practices presented here as possible
- Tailor them to your organization, recognize the uniqueness in the nuances of your program while keeping the direction of your governance approach true
- The best approaches to governance maintain the enterprise focus while implementing iteratively
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.
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)
Data Governance Fundamentals

- **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!!*
What Does All This Have To Do With Data Governance?
Let’s look at the definition for EIM again:

- **Enterprise Information Management (EIM):** The systematic processes and governance procedures for applications, processes, data, and technology at a holistic enterprise perspective

- 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

*EIM is impossible to accomplish without data governance*
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.

### EIM Knowledge Areas

<table>
<thead>
<tr>
<th>Enterprise Information Management</th>
</tr>
</thead>
<tbody>
<tr>
<td>Process Management</td>
</tr>
<tr>
<td>Data Architecture</td>
</tr>
<tr>
<td>Information Quality</td>
</tr>
<tr>
<td>IT Portfolio Management</td>
</tr>
<tr>
<td>Master Data Management</td>
</tr>
<tr>
<td>Information Delivery</td>
</tr>
<tr>
<td>Information Security</td>
</tr>
</tbody>
</table>

### Data Management

<table>
<thead>
<tr>
<th>Data Governance</th>
<th>Meta Data Management</th>
</tr>
</thead>
</table>
Common Data Governance Activities

Common Data Governance Activities/ Goals
- Data identification and definition
- Effective use of meta data
- Data quality metrics definitions and rules
- Brings order, purpose, form, efficiency, performance
- The capture, store, protect and ensure the integrity of the information assets
- Many others!!
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
Current State of Data Governance

- EIM goes by many names:
  - Data Governance
  - Data Stewardship
  - Data Management
  - Enterprise Data Management
  - Centers of Excellence
  - Information Management
  - Information Asset Management

- I don’t care which name you use as long as you do it successfully.

- Many companies are now pursuing this activity at an enterprise level.

- Almost all companies are doing some level of data governance; however, the majority do it piecemeal.
What Does One Process Look Like for a Large Company?
What Do Our Current IT Systems Look Like?
Islands of Data

Operational Applications

Islands of Data

End User Reporting Systems

Legend: Multiple colors indicate the system is used by multiple divisions.
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 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 \times 12 \text{ months} \times 1,000,000 \text{ (1.6 petabytes)} = $153,600,000
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/addxed/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
Data Governance
Data Governance

- Data is one of the most important assets in a corporation
- Data has value when it is delivered timely, properly formatted, concise, accurate and understood
- Business ownership of the data and active participation are critical
- The role of the data steward has grown considerably over the years
 **Data Stewardship:** The process of having data stewards work with the data and meta data of an organization to ensure its quality, accuracy, formats, domain values, and that it is properly defined and understood across the enterprise

 Data Stewardship’s role is to ensure organizational data and meta data meet quality, accuracy, format and value criteria; ensuring that data is properly defined and understood (standardized) across the enterprise
Data Stewardship

- **Data Steward**: A person(s) responsible for working with the data and meta data. There are different types of data stewards:
  - Involved in defining subject area boundaries
  - Collects feedback and enhancements for specific subject area
  - Resolves data integration issues
  - Acts as the conduit between business and IT
  - Serves as quality control point for the subject area

- **Note**: one data steward can be responsible for multiple subject areas

- The data steward acts as the conduit between IT and the business. The data steward (often not just one person, but a collection of people) aligns the IT systems (both decision support and operational) with the business’ requirements. The data steward has the challenge of guaranteeing that one of the corporation's most critical assets--its data--is used to its fullest capacity.
Understanding Data Governance

DATA GOVERNANCE

I need to make profitable decisions

I don’t know what I’m looking at

Data

Misunderstood Inaccurate Misleading

Data Governance

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

Transform Data into Information

Understood Accurate Consistent

Actionable Information

Data Stewards

Data Stewards
Data Governance Fundamentals

- Data steward acts as the conduit between IT and the business
- Align the IT systems (both decision support and operational) with the business’ requirements
- Every company has data stewards...they may not have the official designation but they exist
- No two data stewardship groups are exactly the same
- The task of being a data steward is typically not a full-time assignment
- There are different types of data stewards
Introduction to Data Governance
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

The true tangible value/measure of data governance – when the artifacts are used successfully

Artifacts

Work Results

Kept in Tools
Data governance is the method for connecting information management and the corporate business strategy.
Data Governance Organization
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 8 – 20 Subject Areas in an organization
- The “nouns of an entity. Examples:
  - Legal Entity
  - Location
  - Account
  - Product
  - Customer/Party
  - Sale/Transaction
### 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

**Managed Meta Data Environment**

**Information Technology**
- Data Custodian Team
- Technical Stewards

**Recommendations**

**Requirements**
Subject Area Groups

- Subject Area Groups utilize business stewards for business related subject areas (e.g. attribute definitions) and technical stewards for technical related subject areas (e.g. data transformations)

- A Subject Area Group will likely (definitely) interact with multiple EIM Focus Areas
Each Subject Area Group can be comprised of:

- Business Steward(s)
- Technical Steward(s)
- Interested Parties

Interested parties includes those individuals or groups that have a vested interest and ability to contribute positively to the Subject Area User Group.

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
The Data Stewardship Coordination Group is the centralized organization that houses the EIM Focus Area Steward Teams.

It ensures that the meta data and data being brought in by the Subject Area User Groups is being allocated to the correct EIM Focus Area/Project Steward Team.
The Data Governance Council oversees the data governance organization

- This group handles any serious issues or questions that may arise
- Approve/create policies, procedures and standards that the Data Stewardship Group & Subject Area Groups

Data Governance Council

Members:
- Executive Sponsor(s)
- Program Manager
- Chief Stewards
- CIO
- Key business staff
- Key IT staff

Policies, Procedures, Standards, etc.

Enterprise Oversight

Data Stewardship Coordination Group

EIM Focus Area #1 Steward Team

Subject Area User Group #1
- Chief Steward
- Business Stewards
- Technical Stewards
- Interested Parties

EIM Focus Area #2 Steward Team

Subject Area User Group #2
- Chief Steward
- Business Stewards
- Technical Stewards
- Interested Parties

EIM Focus Area #3 Steward Team

Subject Area User Group #3
- Chief Steward
- Business Stewards
- Technical Stewards
- Interested Parties

EIM Focus Area #4 Steward Team

Subject Area User Group #4
- Chief Steward
- Business Stewards
- Technical Stewards
- Interested Parties

EIM Focus Area #5 Steward Team

Subject Area User Group #5
- Chief Steward
- Business Stewards
- Technical Stewards
- Interested Parties

EIM Focus Area #6 Steward Team

Subject Area User Group #6
- Chief Steward
- Business Stewards
- Technical Stewards
- Interested Parties
Enterprise Oversight

- Enterprise Oversight represents any oversight committee that the organization may have that looks over the Data Governance Executive Committee.

- These groups go by many names including steering committees, project management office, etc.

- Larger companies may have such an organization while medium-sized organizations may not.
Top 10 Reasons for Data Governance Failures
These are some of the basic stumbling blocks that many data governance programs have experienced.

Feel free to ignore these keys...if you want to have a failure on your hands.
Key #1: Overcoming Political Challenges

- Number #1 reason for failure in data governance programs
- People want to do it “their way” instead of the “organization’s way”
- Don’t want to make time for the data stewardship tasks
Key #2: Lack/Loss of Executive Support

- Executive support is vital as these programs require coordination and cooperation from across the enterprise.
- Critical for leaping political hurdles.
- Executive support keeps the program funded.
Key #3: Balance Short-term & Long-Term Goals

- Strategic and tactical
- Very few organizations will wait more than a year for results
- Must identify short-term tactical projects that upon completion would benefit the company but that ALSO move the strategic (long-term) vision forward
- Keep the enterprise in mind
Key #4: Build Iteratively

- Avoid the “Boiling the Ocean” approach to data governance
  - It takes years to fully build a world-class data governance organization
  - Your environment is always changing
- Use 3 – 6 month project cycles
- Government organizations can go to 6 – 12 month project cycles
Key #5: 80/20 Rule

- Data exists all over the organization
  - Some business users will be managing data on their desktops
  - It may be in very old and antiquated systems
  - It may be housed in a remote location

- For large global entities it is almost impossible to govern every scrap of data that may exist

- Govern the “key” data in your organization as this will give you the vast majority of the value for your efforts
Many companies focus exclusively on the initial construction of their Data Governance program but neglect to address ongoing adherence.

Tie your data stewards MBOs (management by objectives) into their data stewardship activities.

Make sure to keep your MME (automated)
Key #7: Business Is Actively Involved

- Data governance is not a IT project. It is a business program.
- Business people must be actively involved in data stewardship roles and oversight.
- Without business’ help this task is almost impossible.
Key #8: Inability to Measure Success

- Many data governance programs neglect to define measurable metrics
- Most programs fail to take a baseline measurement
- This prevents them from being able to measure success
Key #9: It’s a Program Not a Project

- Projects have end dates, programs do not
- This is program
- If properly built your data governance structures will last for many years
Key #10: It’s Not A Science Project

- These efforts are embarked on to provide REAL value to an enterprise
- They are NOT “science projects”
- Make sure to define specifically the value that data governance will provide
Data governance programs are not easy and take many years to truly develop.

There are NO MAGIC BULLETS.

You must ALWAYS understand your data.

Don’t cut corners and do it on the “cheap”.

Be smart, diligent and patient.

Never stop selling your program.

You will achieve GREAT results.