A Comprehensive Guide to Metadata Management using SAP Information Steward

Session Code: 0207

Michael Briles, Enterprise Information Management
August, 2012
Data Quality in Context

Overview – SAP Information Steward

Metadata Management with SAP Information Steward

Demo
Data Quality in Context

INTRODUCTION
Dealing with Known Data Quality Issues ...

Source: Dilbert, May 2008
http://dilbert.com/strips/
... But How to Identify and Manage the Hidden Data Quality Problems?

Where is data stored within the organization?

What content is really stored within the organization?

Who is using data within the organization?

How is data transferred within the organization?

What are User's requirements on the data?
Information Governance to Optimize Every Business Process and Ensure Compliance

- **Decrease COGS**
  - Consolidate supplier data to optimize purchasing

- **Decrease days sales outstanding**
  - Warehouses use bar code data to speed shipping

- **Increase promotion effectiveness**
  - Social media data for segmentation, sentiment, behavior information

---

**Suppliers**

**Plant**

**Warehouse**

**Distributor**

**Retailer**

**End Consumer**

---

- **Decrease compliance issues**
  - Plant operators use carbon input data to identify compliance

- **Decrease time-to-delivery**
  - Geo spatial data identifies best routes/number of trucks/timing
Overview

SAP Information Steward

EMPOWERMENT

Empower your Data Governance Project Team
Business Challenges with Data Quality and Data Integration Projects Today

**No** environment for business users to collaborate with IT regarding data issues

**No** consistent repeatable way to measure and score data quality

**No** clear strategy and discipline for improving data quality

**Minimal** reuse of data assets; lots of data duplication

**Not** sure what the business definition is

**No** ability to analyze data dependencies across systems

**Lack** of visibility to where the numbers or data is coming from and many more…
SAP Information Steward
Collaborative environment for your IT and business users

Discover
Understand and catalog enterprise data

Define
Rules and ownership

Assess
Overall data quality

Monitor
Quality continuously

Improve
Data quality and governance

Empower business and IT users with a single environment to manage the quality of their enterprise data assets
How Can Business and IT Users Leverage SAP Information Steward?

- **Data Quality Scorecard**
  - Apply “balanced scorecard” towards data quality performance
  - Measure and track data quality performance against metrics
  - Anticipate and spot data quality weakness

- **Data Profiling**
  - Analyze if data matches business definitions and expectations
  - Validate data completeness, sparseness, redundancy, pattern distribution
  - Analyze cross system data dependencies using business views

- **Validation Rules**
  - Work with IT to define data validation rules
  - Apply validation rules against sources to continuous monitor data quality
  - Reuse validation rules in data migration and integration

- **Metapedia**
  - Work with business users to define and agree on business terms
  - Link business terms to data elements to assist business users understand data definitions

- **Metadata management**
  - Trace data lineage from BI reports to data sources
  - Assess impact of data quality or changes
  - Reuse data assets that is available

- **Cleansing Package Builder**
  - Define party or non-party cleansing packages
  - Apply business knowledge directly into defining the package
  - Automatically generate parsing rules used by Data Cleanse transform in Data Services
Key Innovations

Data Profiling

- Run data validation policies on data in the context of business processes
- Visualization of ‘fit for purpose’ data quality
- Gather statistical information, analyze and identify weak spots
- Drill down to the data for frequency distribution, pattern analysis
- Identify redundancy and dependency of columns and tables

Metadata Management

- Automated collection of technical and business metadata from applications repositories
- Correlate metadata across repositories for lineage and impact understanding
- Follow the data from report to source
- Lower TCO by tracking usage and promote reuse of data and reports
- Improve decision making and regulatory compliance
**Key Innovations**

**Business Term Glossary**
- Promote proactive data governance with common understanding and agreement on business taxonomies
- Central repository that links and organizes terms to data assets
- Assign roles and responsibilities for definitions, approvals and accountability

**Cleansing Package Builder**
- Empowers domain experts to develop custom data cleansing solutions for any data domain (Party and Non-Party)
- Intuitive drag and drop UI for non-technical users
- Native Integration with Data Service allows publishing of Cleansing Packages to developers in Data Services Repository
- Data Services can use Cleansing Package in any Real-time or Batch Job for Data Cleansing Action
Key Innovations

Monitor Data Quality
- Provide Data Stewards with graphical representation of overall data quality
- Measuring and tracking data quality against defined governance standards
- Historical tracking and trending allows for measuring effectiveness of governance program.

Identify Root Cause or Impact
- Drill down to root causes of poor data quality as it exists in the source system
- Quickly isolate data anomalies through preview of failed data.
- Establish impact of bad data on critical business analytics, processes and who it affects.
Metadata Management with SAP Information Steward
Metadata Management
Consolidate, Integrate, Audit, and Trust in your Metadata

- Consolidates metadata from various BI related sources
- Integrates metadata in a central metadata repository
- Relates metadata to enable auditing, usage, change impact, and data lineage analysis
- Supplements metadata with custom attributes, annotations
- Creates universe from relational objects
- Metapedia, a business term encyclopedia
- Advanced search
Metadata Management
Metadata Integrators (native and through additional MITI Bridge)

Metadata Management Integration to:

BI Systems

Metadata Integrators bundled with Information Steward

SAP Software: BI Platform, Data Services, Data Federator and SAP NetWeaver BW
NON SAP Software: CWMXML (CWM), RDBMS: MSSQL Server, DB2, Teradata and JDBC Sources

Other Metadata Integrators available *

Altova  MicroStrategy
Borland  NoMagic
CA  OMG
COBOL  Oracle
Embarcadero  SELECT
EMC  SPARX Systems
Gentleware  SUN
Silverun  Sybase
Knightbridge (HP)  Teradata
IBM  Tigris
Informatica  Visible Systems
Micsosoft  Corp W3C

• Other metadata integrators available via Meta Integration Technology, Inc. (MITI);
  • http://www.metaintegration.net/Products/MIMB/SupportedTools.html
Metadata Management
Information You Can Trust with End-to-End Metadata Management

- Understand where the data used in your applications comes from and how it is transformed
- Assess change impact analysis for source to reports
- Manage metadata from various data sources, data integration technologies, and BI systems
- Lower TCO by tracking usage and promote reuse of data and reports
- Improve decision making and regulatory compliance

“Where did this number come from?”

Impact Analysis

“How will this change in the source impact my BI reports?”
Information Steward Metadata Management Administration

Configuring Sources

- SAP BusinessObjects Enterprise Metadata Integrator
- SAP NetWeaver Business Warehouse Metadata Integrator
- Common Warehouse Metamodell (CWM) Metadata Integrator
- SAP BusinessObjects Data Federator Metadata Integrator
- SAP BusinessObjects Data Services Metadata Integrator
- Relational Database Metadata Integrator
- Meta Integration Metadata Bridge (MIMB) Metadata Integrator - also known as MITI Integrator
Running a Metadata Integrator - Properties
Metadata Management Metadata Explorer

Information Steward

Metadata Management

Directory

Business Intelligence (1)

- BCE BI - Oracle Sales RapidMart
  Description: BOE BI - Oracle Sales RapidMart
  User name: Administrator
  Authentication mode: secEnterprise

- Universes (5) Reports (119) Folders (11) More

Data Modeling (1)

- CWM XML Data Model - ERwin eMovies
  Description: CA ERwin eMovies CWM XML model
  CWM File name: freswin sample modal cwm 1.0.xml
  Catalog (1) Schema (1) Tables (12) More

Data Integration (1)

- Data Services - Oracle Sales RapidMart
  Description: Data Services - Oracle Sales RapidMart
  Database Name: ORASALES_RapidMart
  Database Type: Microsoft_SQL_Server

- Projects (1) Jobs (3) Work Flows (58) More

Data Insight Connections (2)

- Sales Application
  Description: Sales Application
  Connection Type: Microsoft SQL Server
  Server Name: localhost
  Schemas (3) Tables (4)

- Tamarin System
  Description: Tamarin System
  Connection Type: Microsoft SQL Server
  Server Name: localhost
  Schemas (2) Tables (3)
Metadata Management Metadata Explorer
Searching for objects
Metadata Management Metadata Explorer
Integrating Data Quality Content
Metadata Management within BI Launchpad

Trusted Information: Where does it come from AND how good is it?
Metadata Management within BI Launchpad

Trusted Information: Where does it come from AND how good is it?
Metapedia
Business Metadata Encyclopedia

Group Business Terms in Category Hierarchy

Export & Import Business Terms

Define Business Terms and their description

Assign Data Steward to each Business Term
Metapedia
Linking technical objects to business terms

Define Category and Sub-category; associate Terms

Associate Objects to the terms
Demo
Key Takeaways
The best choice for EIM

✅ One place for data stewards and business analysts to collaborate and govern their data asset

✅ First and only integrated solution - metadata management, business glossary, data quality assessment, data quality monitoring, and cleansing package builder solution

✅ Various perspectives to understand and analyze trustworthiness of data

✅ Easy and secure access to a rich set of metadata data sources
Thank You!

Contact information:

Michael Briles
Director - Enterprise Information Management
SAP Labs LLC
Seattle, WA

T: 206-355-6059
E: michael.briles@sap.com
Alliander Achieves Data Quality Controls and Reduces IT Costs

Why we were successful?

- Control of data quality and data quality measurements
- Data quality reports that can be created from a single screen with no need for IT support
- 60% reduction in IT costs by eliminating the programming and data design cycles in the SAP NetWeaver Business Warehouse component
- Extend Data Governance capabilities
- Improved business rules definition that everyone understands
- Ease of use solution and fast user enablement with only 4 hours of training

Customer quote

“SAP Information Steward solution helps our data owners to bring data quality to the next level. The biggest benefit is measuring the quality itself so allowing data owners to actually measure the quality and not go by, let’s say gut feeling but actually making sure that they know how many data objects are complied, are in compliance and how many data objects are not.”

John Oud, Alliander

Products purchased

- SAP Data Services
- SAP Information Steward
Data Drives Agility at Colgate-Palmolive

Products purchased

SAP Master Data Governance for Finance (Cost Center)
SAP Information Steward

Why we were successful?

• Support the need to identify defect data and pinpoint data quality using metrics offered in the solution.
• The business impact of defect data impact is huge for Colgate because it affects production planning.
• With Information Steward, they are able to produce the right products at the right time and reduce the wrong items being manufactured.
• Information governance support by enabling business and IT collaboration, measure and fix the quality of information so that it is fit for use across the enterprise.

• More accurate and predictive analytics, standardized global processes, and enhanced enterprise-wide reporting

Customer quote

“When we combine the integrated data with improved processes, that is when we see the real benefits,”

Ruben Panizza
Global IT Director, Colgate-Palmolive