

Aligning Business and IT to achieve a successful Enterprise Data Warehouse

Presented to DAMA NJ by:

Nick Bonamassa
President, Info Architects, Inc.

16 Esther Drive, Suite 101
Freehold, NJ 07728

Phone/Fax: (732) 866-4119
eMail: nickb@infoarchitectsinc.com

www.infoarchitectsinc.com

Introduction

One would assume that successful organizations have a handle on their data, but my experience shows there are still many IT issues that stem from poor data and even more so meta data management. This becomes especially evident when developing a major system like an Enterprise Data Warehouse.

This presentation focuses on the critical issues surrounding an Enterprise Data Warehouse and the points that require Business and IT alignment in order to be successful.

Agenda

- Company Background and Speaker Bio
- Overview of an Enterprise Data Warehouse Architecture
- Overview of Meta Data for Business Intelligence (BI)
- Alignment between the Business and IT on Meta Data
- Implementation of Meta Data for a Successful Enterprise Data Warehouse BI Solution

Company Background and Speaker Bio

Nick has 20 years of IT expertise and has been focusing on Data Warehouse, Business Intelligence and Meta Data in the strategic planning and design of Information Environments for over 14 years. He also possesses Project Management and development experience in full system lifecycle support. Nick has developed strategies, architectures and models that have directed major Financial and Consumer Goods companies to achieve immediate value for their business communities while fulfilling long term IT vision.

Prior to Nick forming his own company, he held several positions in IT, including at the management and executive level, for Fortune 500 Financial Institutions.

Nick was on the Board of Directors of the Data Administration Management Association New York Chapter (DAMA NY) from 1996 through 2005. Nick has presented at industry conferences around the country.

Nick can be contacted at Info Architects, Inc.

office/fax (732) 866-4119 or nickb@infoarchitectsinc.com

Enterprise Data Warehouse Architecture

Enterprise Data Warehouse Architecture

The first aspect for a successful Enterprise Data Warehouse is its architecture. The architecture needs to be flexible so that it can be quickly adaptable to changes in the business direction.

Key Points:

- Multiple layers
- Land source data unchanged, then transform so you have an audit trail
- Leverage ETL tools
- Source to Target Mappings (STTMs)
- Provide user views upon a detailed data model

Enterprise Data Warehouse Architecture

Multiple layers

- The simplest division is 3 layers:
 - Gather - the extraction layer
 - Store - the main data warehouse database
 - Deliver - the user table/views layer

Enterprise Data Warehouse Architecture

More complicated architectures combine:

- A dedicated 'landing zone' for the extracted data (more on this later)
- A Data Cleansing Layer
- A Data Integration Layer
- The Central Data Warehouse Database
- An Aggregation Layer
- User tables, views, or cubes

Enterprise Data Warehouse Architecture

Landing Zone

- A layer where the source system data remains unchanged
- This area provides the starting point for the data that is transformed in the Data Warehouse process supporting the audit trail
- Also provides a place to restart from

Enterprise Data Warehouse Architecture

Leverage ETL tools

- Using ETL tools can speed not only development but increase turn-around time on changes or coding corrections.
- During testing a mapping rule can be changed in a matter of minutes.
- There are certain complex processes that may need to be developed outside the tool.

Enterprise Data Warehouse Architecture

Source to Target Mappings (STTM)s

- A document describing the mapping or transformations between a source and target element.
- Applies between each layer of the Data Warehouse Architecture.
- Becomes the main reference document for the project.
- Transformation rules can be business-like and technical.

Enterprise Data Warehouse Architecture

Provide user views upon a detailed data model

- The central Data Warehouse tables should be modeled to store the most granular (detailed) data.
- Providing user tables or views should insulate power-users from potentially complex queries.
- These conceptually can be considered data marts.

Business Intelligence

The second aspect for a successful Enterprise Data Warehouse is Business Intelligence (BI). The BI product is what the majority of the Data Warehouse users will see. The biggest issue I have encountered is inconsistent views of information. This issue has primarily been caused by inadequate meta data.

Also, the Data Warehouse Data Architecture play a major role in successful BI.

Enterprise Data Warehouse Data Architecture

Enterprise Data Warehouse Data Architecture

- Uncover data issues
 - research feedback from users
 - inconsistent analysis / report results
 - difficulty accessing data
 - data audits
 - reconciliation of reports
 - consistency of code values
 - data design reviews
 - data models
 - physical structures
 - availability of meta data
 - data elements
 - ETL / transformations
 - report definitions

Enterprise Data Warehouse Data Architecture

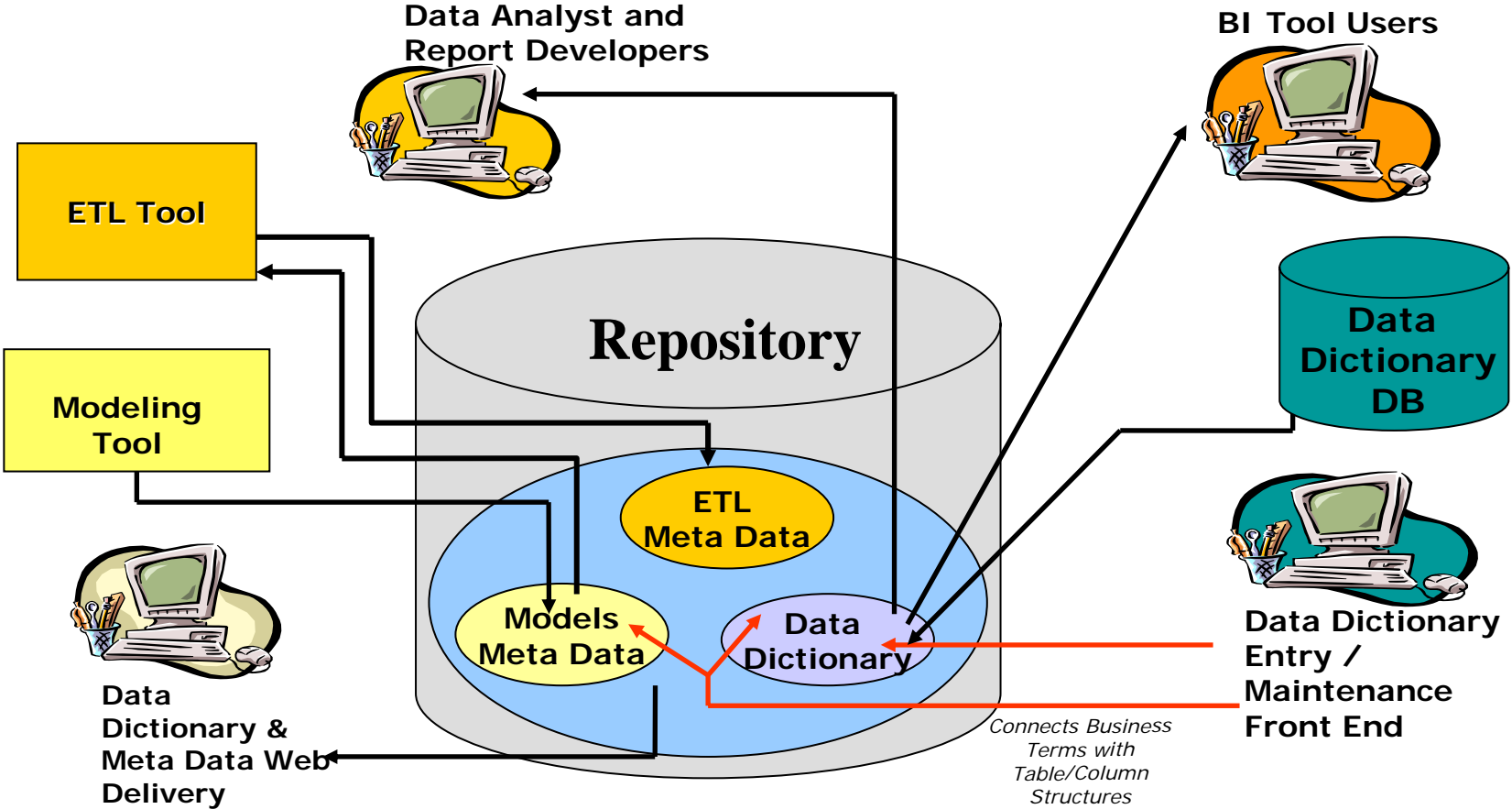
- Development practices
 - understand the Data Warehouse development process
 - how are requirements gathered and documented
 - data modeling
 - report development
 - how and what meta data is collected

Enterprise Data Warehouse Data Architecture

- Data management practices
 - logical and physical modeling
 - source system analysis and ETL development
 - documentation (meta data)
- The Data environment
 - Database Inventory
 - Data Tables
 - Central Files
 - Database Technology
 - Data Management Tools
 - Data Standards

Enterprise Data Warehouse Meta Data Architecture

Meta Data Architecture

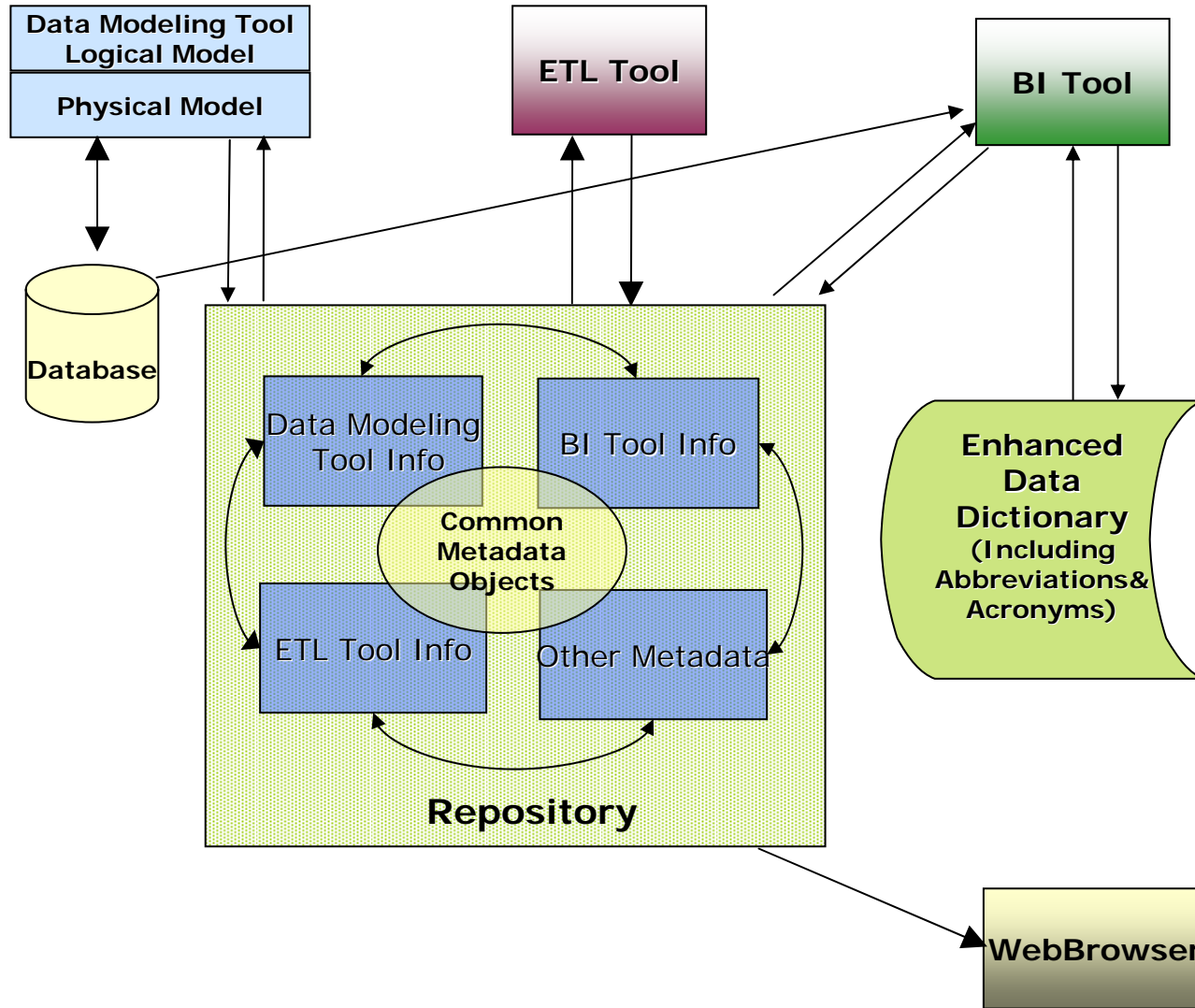


Integrated Tool Architecture

Meta Data “Round Trip”

- Data Modeling Tool
- Repository
- ETL Tool
- Business Intelligence (BI) Tool
- Corporate Data Dictionary

Meta Data “Round Trip”



Meta Data in an Enterprise Data Warehouse Environment

- Definition
 - Data about data....
 - Specifications for data....
 - Collection of things that describe data
 - Encyclopedia of relevant information on data
 - Data the organization needs to manage its information resources
- Categories
 - Each category of meta data is a data model in itself (meta model)
 - Source system data definition
 - Owner and contact information
 - Operating platform and availability
 - Data objects definition (tables, files, etc.)
 - Data element definition
 - Definition of relationships between objects (ER, summaries,...)
 - Allowed values (code tables)

Meta Data in an Enterprise Data Warehouse Environment

- Data extraction and transformation rules
 - Source/Target elements
 - Data extraction programs
 - Data transformation: Filters, Mappings, Transform & Calc rules
 - Administrative Meta Data: Status of jobs, Audits
- Data warehouse / data mart definition
 - Similar to source system Meta Data
 - Owner and contact information
 - Operating platform and availability
 - Data objects definition (tables, columns, etc.)
 - Data transformation and calculations
 - Definition of relationships between objects (ER,summaries,...)
 - Allowed values (code tables)
 - Warehouse usage and purpose

Meta Data in an Enterprise Data Warehouse Environment

- Data quality
 - Data quality metrics (completeness, accuracy, timelines, currency, etc.)
 - Measurement rules
 - Actual measurements (On source systems, On warehouse data)
- End user data access
 - End User Data Access
 - Subscription information
 - Standard queries and reports
 - Security and access authority
- Operational
 - Administrative Meta Data
 - Scheduling of extracts and transformation
 - Status on loads
 - Status on audits
 - Security (user access and privileges)
 - Warehouse usage statistics
 - Warehouse infrastructure

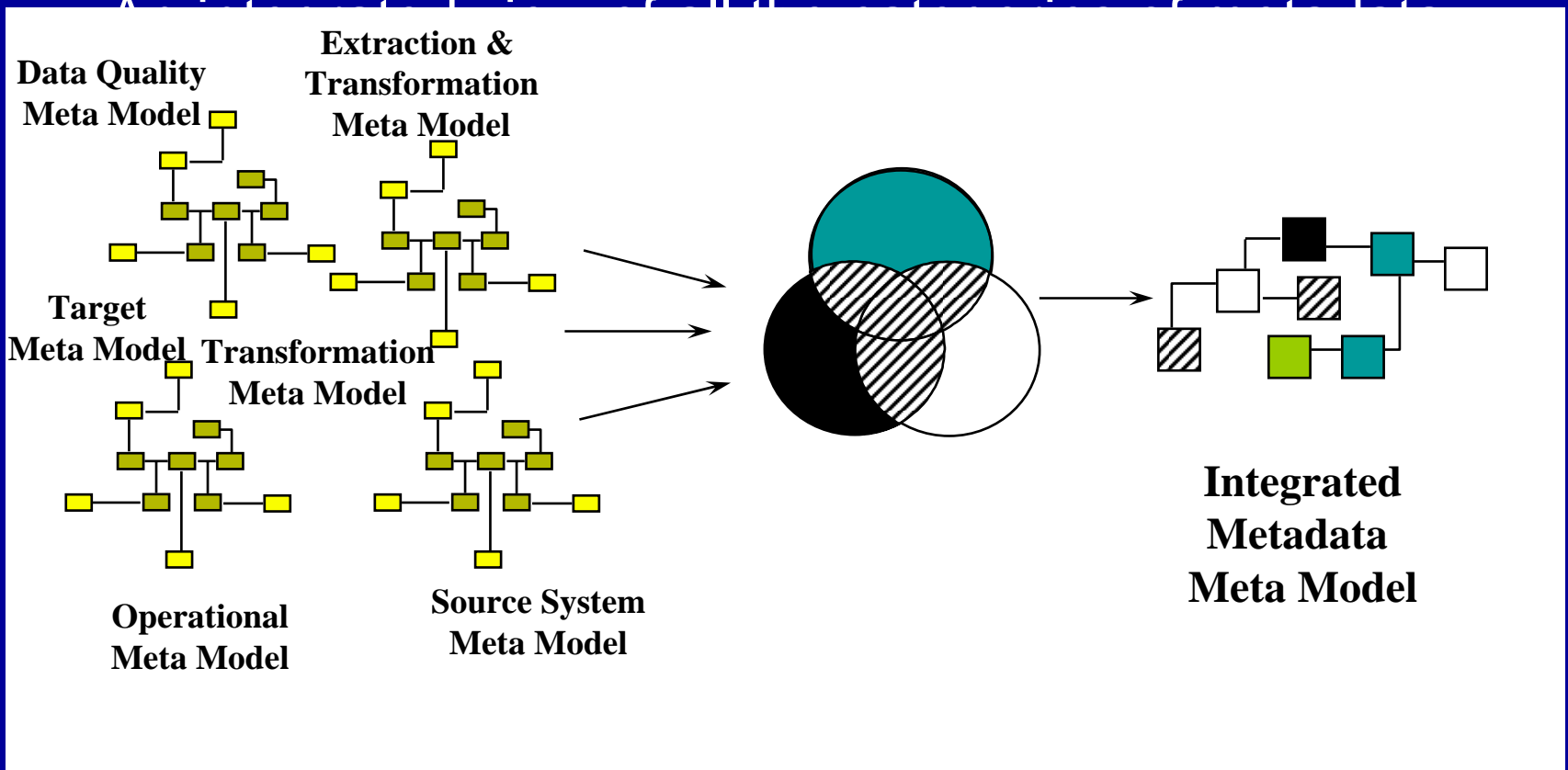
Meta Data in an Enterprise Data Warehouse Environment

Categories (continued)

– For each category identify:

- **Why** meta data is needed (benefits)
- **Who** are the users
- **What** meta data is needed (meta model)
- **Where** is the meta data
- **How** is the meta data gathered, stored and delivered
- **When** outlines the schedule

Repository Model for an Enterprise Data Warehouse



Alignment between the Business and IT/Data Management on Meta Data

Meta Data Issues in an Enterprise Data Warehouse Environment

- Issues
 - Typically meta data is not well defined
 - Lack of / conflicting standards
 - Definition
 - Storage
 - Delivery
 - Different processes to create and maintain meta data
 - Meta data is (conceived to be a) bottleneck to rapid development
 - Repository products experiences are not user friendly

Meta Data Audience for an Enterprise Data Warehouse

- Business Users
 - Business Analysts
 - Data Analysts
 - Report Developers
 - Management
- IT Users
 - Developers
 - Production control
 - Data administrators
 - Database administrators

Enterprise Data Warehouse Business Meta Data Requirements

Glossary Term and Definition

Business Name, Short Name, Abbreviated
Name and Definition

DQ Certification Metrics

Business Rules

Data Steward, Owner

Table/View and Definition

Column, Length, Type, Format and
Definition

Report / Cube Name and Definition

ETL Transformations

Data Lineage (Source to Target)

Data Models

Data Entry / Maintenance
Facility

Automatic / Easy Mapping of
Meta Data

Web Delivery

Connections from the business
data element to the technical
name(s)

Connections from the business
data element to the BI
reporting tool

Abbreviations & Acronyms list

Alignment between Business and IT/Data Management on Meta Data

- Processes and procedures and meta data ownership needs to be agreed upon to move forward.
- Technology challenges, across tools and their interfaces need to be understood.

General Meta Data Issues

- Which organization, business or IT, “owns” which type of meta data?
- Which tool is the “master” keeper of each type of meta data?
- How and when will the meta data be moved among the tools, and by whom?
- What impact analysis is needed of meta data across the tools?
- Which meta data needs to be presented to whom and when?

Meta Data Issue #1

- *Which organization “owns” the business definitions (for data elements)?*
- *Easy Answer: the business representatives*

Meta Data Issue #2

- *Which tool is the “master” keeper of the business definitions?*
- Business definitions may already exist in a Data Dictionary or Glossary
- Many model-oriented definitions may already exist in the Data Modeling tool
- Definitions need to be accessible through BI tool
- *Decision: Data Dictionary, although model-oriented “addendums” to the definitions remain in Data Modeling tool. The model definitions may describe the elements ‘in context’ of the table they are in.*

Meta Data Issue #3

- *Which organization “owns” the naming of data elements?*
- *Data Management designates technical names, using approved abbreviations*
- *Designated business representatives responsible for business data names, long and short for BI reports, also using approved abbreviations*

Meta Data Issue #4

- *How to coordinate data model name abbreviations with business name abbreviations?*
- Some business short names not short enough to be data model/technical names
- *Answer: Combine the IT data model abbreviations list with the business abbreviations & acronyms list adding an IT 'short' abbreviation*

Meta Data Issue #5

- *Which tool is the “master” keeper of the mapping from business names to technical names?*
- Approved business names may already exist in a Data Dictionary
- Data model/technical names exist in the Data Modeling tool
- Some BI tool data elements are derived and do not map to technical names
- Mappings in ETL tool are very detailed and multi-step
- Hard to “match” business names to data model/technical names
- *Decision: Data Dictionary structure needs to support the mappings*

Meta Data Issue #6

- *How and when will the metadata be moved among the tools in order to handle CRUD for data structure changes?*
- Tool-related and work flow issues
 - Movement of definitions and data structure changes between BI tool sections and reports
 - Filtering delta data structure changes
- *Decision: Little movement of meta data away from “master” location*

Meta Data Issue #7

- *Which types of metadata need only one (“production”) version?*

Answer:

- Business definitions and names
 - Approved abbreviations and acronyms
 - BI reporting tool section and report information
 - Connections from business data elements to data model/technical name(s)
 - ETL job metadata
-
- *And which also need a “development” version?*

Answer:

- Data Models
- ETL job metadata

Meta Data Issue #8

- *What is the migration process from development to production for certain meta data?*

Answer:

- *Procedures needed for archiving and versioning of Data Models*
- *A separate production environment should be established for ETL job meta data, using ETL tool utilities for migration*

Meta Data Issue #9

- *Can the SDLC methodology and project management protocols be adjusted such that approved business definitions and business names are provided with business requirements prior to IT data modeling effort?*
- Impact on the workload of the business representatives
- Scheduling coordination
- *Answer: Process changes may be required possibly on a project-by-project basis*

Meta Data Issue #10

- *What further impact analysis is needed of meta data across the tools?*

Answer:

- Business names used in Data Models
- Data elements used in ETL jobs/steps
- Business data elements (in BI reports) affected by data structure change in data model
- Use of repository to view cross-references among metadata

Meta Data Work Flow Policies

- *Data structures originate in the Data Modeling tool, are imported into the Repository, and then pushed to the ETL tool*
- *Data structures are not brought into the ETL tool directly from the DB*
- *Data structures are brought directly into the BI tool from the DB*

**A Future
Meta Data Management
Vision**

A Future Meta Data Management Vision

Data Standards and Guidelines

- Establish standards for modeling tools, data naming and core data identifiers
- Build consensus for reference data, glossary of terms, common data domains, and structures of shareable data


 **Guidelines and Standards are the foundation of reference and business meta data**

A Future Meta Data Management Vision

Meta Data Management

Establish Repository

- Web Front End Delivery
- Graphical navigation of data models

 **Without Meta Data Management there is no foundation for identifying and resolving Data Issues for the Data Warehouse**

A Future Meta Data Management Vision

Enterprise Data Models

Top-down approach

- Need to support all lines of business
- Reverse engineering can be used to validate integrity
- Divided into data subjects

Multiple levels of data models representing various viewpoints or focuses

- by data subject
- by application view

✓ **Enterprise Data Models are the foundation of logical Meta Data and core data elements.**

A Future Meta Data Management Vision

Center of Excellence

Communication channels for Information sharing and knowledge transfer

- discussion database, intranet site, meetings, etc.

Expertise and leadership

- Use the discussion database to answer data-management related questions
- Constantly research new technologies



Center of Excellence establishes high customer confidence in Data and Metadata Management expertise

A Future Meta Data Management Vision

Community Building

Company-wide DBA / data warehouse / data and meta management meetings

Assist lines of business in establishing local Meta Data Management infrastructure

- Including consultation and resource support

Reach out to peers and professional organizations in the industry

- 📄 **Internally it raises the visibility and awareness of Data and Meta Data Management**
- 📄 **Learning from industry peers helps avoid re-inventing the wheel**

A Future Meta Data Management Vision

Data Warehouse

- Meta Data assists in
 - accessing warehouse data
 - building LOB data warehouses
- Meta Data mapping is critical
 - between enterprise warehouse and LOB warehouses
 - between LOB warehouses and marts
 - between marts and enterprise warehouse
 - between marts



Data Warehouses can provide sponsorship for Meta Data implementation



It helps us prioritize tasks based on business objectives

Meta Data Implementation

Meta Data Implementation

- Gap Analysis
 - Identify Data Opportunities
 - Reference Table Standardization
 - Identifiers for core data
 - Data Naming Standards
 - System of Record
 - Development Processes
 - Requirements
 - Modeling
 - Documentation
- Developing a plan to implement the long-term Meta Data Management vision
 - Prioritize opportunities based on strategic significance and interdependencies
 - Develop High Level Master Plan and divide into Phases
 - Develop Detailed Project Plans for Phase 1

Meta Data Implementation

Meta Data Keys to Success

- ◆ Meta Data must be implemented with “*end goals*” in mind
- ◆ Meta Data **delivery** should be the priority
- ◆ Business should drive Meta Data priorities

Meta Data Implementation

Gaining buy-in

Partner with Application Managers

- support application development and provide value added services



Business Group and Enterprise Meta Data Management Relationship

- Meta Data Management will **coordinate** the LOB's MetaData
 - facilitate **sharing** of information to avoid redundancy between LOBs
 - ensure **consistent** data names, structures, code values across LOBs
 - provide a central point to **communicate Meta Data** Management opportunities and challenges across the Enterprise

Meta Data Implementation

Provide Value Add

- Data and Meta Data Management needs to provide value-added services to projects so as not to be perceived as a bottle-neck to development
- Ways to provide value-add:
 - perform business analysis tasks
 - clearly communicate an understanding of the business requirements correctly model the business requirements
 - design well-performing database structures
 - capture and deliver appropriate meta data

Questions



Contact

**Nick Bonamassa
President, Info Architects, Inc.**

**16 Esther Drive, Suite 101
Freehold, NJ 07728**

**Phone/Fax: (732) 866-4119
eMail: nickb@infoarchitectsinc.com**

www.infoarchitectsinc.com