

Building and Managing the Meta Data Repository



By David Marco

President

Enterprise Warehousing Solutions, Inc.



Enterprise Warehousing Solutions



Enterprise Warehousing Solutions, Inc. (EWS) is headquartered in Chicago and is a strategic consulting partner and systems integrator providing clients with best-in-class, knowledge-based solutions using data warehousing and meta data repository technologies.

EWS provides both strategic consulting and full-service implementation services. This combination affords our clients a full range of services for any size data warehouse/business intelligence and/or meta data repository project. EWS' strategic consulting and implementation services are dedicated to a single purpose: ensuring that our clients have successful data warehousing and meta data repository projects. Our client list includes:

Agilent Technologies
Alltel Information Services
American Management Systems
Arizona Supreme Court
Branch Banking & Trust (BB&T)
California DMV
Corning Cable Systems
Driehaus Capital Management
Eli Lilly and Company
Fair Isaac and Company

Ford Motor Company
GlaxoSmithKline
Harris Trust and Savings Bank
Harvard Pilgrim HealthCare
Informix
Janus Mutual Funds
Key Bank
Mars Candy
Microsoft
National City Bank

Neighborhood Health Plan
Pillsbury
The Regence Group
Singapore Defense Science & Technology
Target Corporation
Thomson Multimedia (RCA)
United State Air Force - Wright Patterson AFB
United States Transportation Command
Wells Fargo
Wisconsin Department of Transportation

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 708.233.9551 or email us at Info@EWSolutions.com



www.EWSolutions.com

Enterprise Warehousing Solutions, Inc. ©
2002 All Rights Reserved



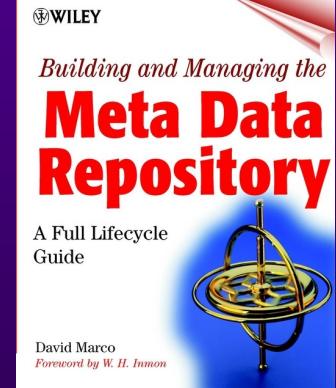
Professional Profile/Contact Information

Mr. Marco is an internationally recognized expert in the field of data warehousing, Capability Maturity Model (CMM), XML, business intelligence, and is the world's foremost authority on meta data. Mr. Marco is the author of the widely acclaimed book "**Building and Managing the Meta Data Repository: A Full Life-Cycle Guide**" (Wiley, 2000). This groundbreaking book has been broadly endorsed by many of the largest software companies in the industry and by several major magazines. In addition, he is coauthor of the soon-to-be-released "**Impossible Data Warehouse Situations and Solutions From The Experts**" (Addison-Wesley).

- ◆ Worked on many data warehousing and meta data repository projects
- ◆ Co-sponsors with **Penn State University** a series of certified data warehousing courses
- ◆ Judged over 16 industry awards in meta data repository and data warehouse development
- ◆ Provided strategic IT consulting for numerous corporations
- ◆ Teaches at the **University of Chicago**
- ◆ Published over 80 articles on information technology
- ◆ Presented over 100 keynotes/seminars on information technology
- ◆ 2001 finalist for the DAMA Individual IT achievement award

Direct: 708.233.6330
Email: DMarco@EWSolutions.com

London 2002 - 3



www.EWSolutions.com

Enterprise Warehousing Solutions, Inc. ©
2002 All Rights Reserved

Acknowledgment

Session materials adapted from...

Building and Managing the Meta Data Repository: A Full Life-Cycle Guide

John Wiley and Sons, Inc.
ISBN 0471-355232

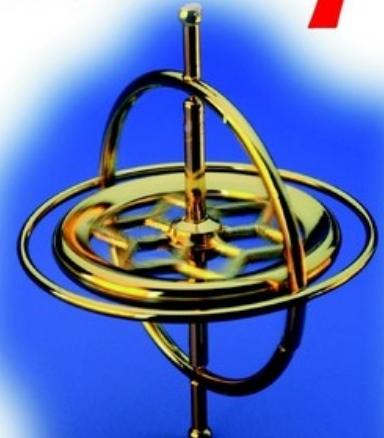
London 2002 - 4



Building and Managing the **Meta Data Repository**

A Full Lifecycle
Guide

David Marco
Foreword by W. H. Inmon



www.EWSolutions.com

Enterprise Warehousing Solutions, Inc. ©
2002 All Rights Reserved



Protocol

Turn pagers and cellular phones to vibrate
Please don't take cellular calls in the class room
Keep side conversations to lunch and breaks
Be polite
Have fun!





Agenda

The State of the Meta Data Marketplace

Meta Data ROI

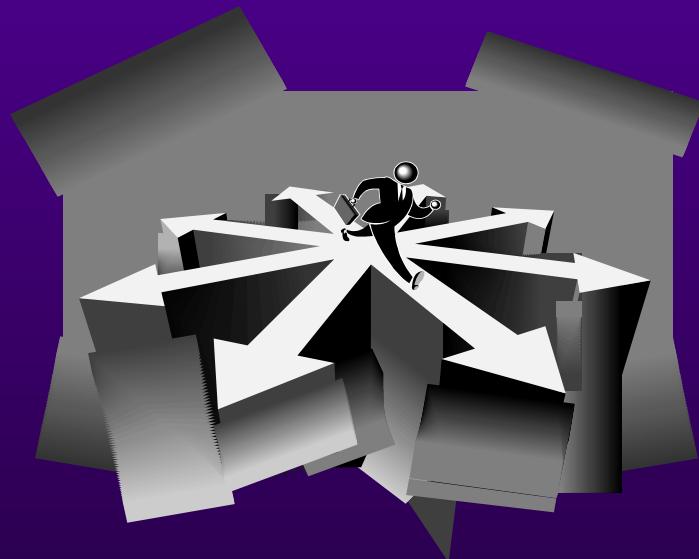
Meta Data Fundamentals

Meta Data Tools

Evaluating Meta Data Vendors

Generic Meta Modeling

Real-World Case Studies



www.EWSolutions.com

Enterprise Warehousing Solutions, Inc. ©
2002 All Rights Reserved

The State of Meta Data Today

Increased need for meta data
Misleading marketing
Complex integration architectures





Increased Need For Meta Data

Inflexible Systems

Increasing IT Demands

Unfulfilled Business User Needs



www.EWSolutions.com

Enterprise Warehousing Solutions, Inc. ©
2002 All Rights Reserved

Increased Need For Meta Data

Inflexible Systems

Business continues to change more rapidly

Build it and they will come doesn't work anymore

Current systems are too inflexible and non-integrated

Need for enterprise impact analysis



www.EWSolutions.com

Enterprise Warehousing Solutions, Inc. ©
2002 All Rights Reserved



Increased Need For Meta Data

Increasing IT Demand

Integrated systems

Corporate data typically doubles every three years



Increased Need For Meta Data

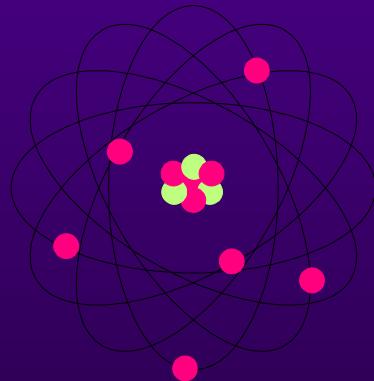
Unfulfilled Business Users Needs

We are not meeting the needs of our business users
We do not speak to them in their language



Misleading Marketing

Both from consulting firms and software firms
Occurs because meta data is a “hot” (profitable) area





Misleading Marketing

Cutting Through the Hype

Vision: Vendors (consulting & software) should have a vision as to what meta data can do, and where the industry is heading

Methodology: Vendors must know what they can deliver today and how it can be expanded, over time to provide an enterprise solution

Look out for the “**Quick Fix**”



www.EWSolutions.com

Enterprise Warehousing Solutions, Inc. ©
2002 All Rights Reserved



Complex Integration Architectures

Repositories commonly source data from the following tools:

- ETL (extraction, transformation, & load)
- OLAP (on-line analytical processing) & Portals
- Modeling
- Databases
- Spreadsheets
- Word processors
- Data quality
- Data profiling

Best-of-breed market creates non integrated tools





Meta Data

What is Meta Data?

Meta Data

By definition meta data is
(3) “Data about data”.
(4) “Everything that data is not”.

So What!!





What is Meta Data?

Meta Data Full 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.

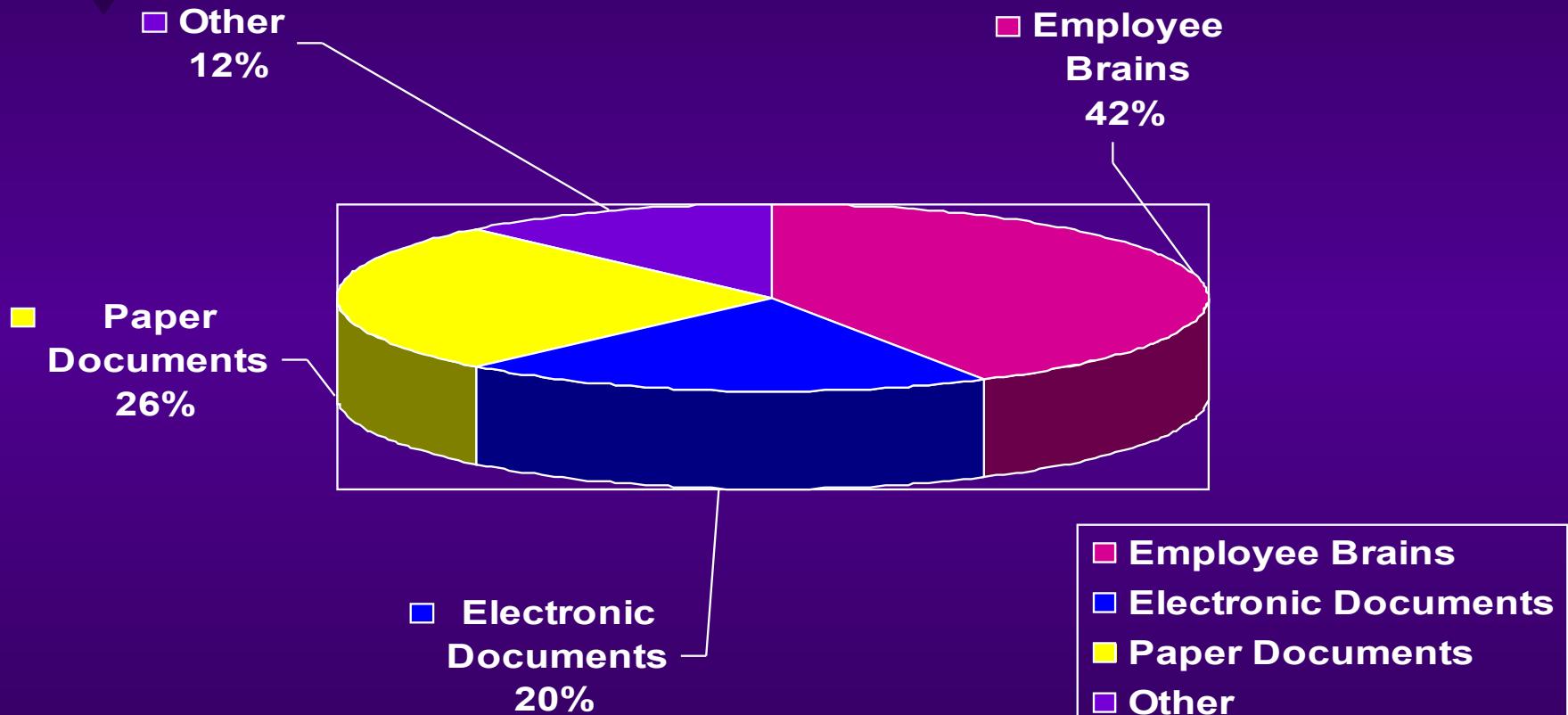
Meta Data Is Knowledge



www.EWSolutions.com

Enterprise Warehousing Solutions, Inc. ©
2002 All Rights Reserved

Where is Corporate Knowledge Stored?



Cutter Consortium



www.EWSolutions.com

Enterprise Warehousing Solutions, Inc. ©
2002 All Rights Reserved



Meta Data Repository ROI

Meta Data Repository ROI

“The key to your company’s prosperity is how well you gather, retain and disseminate knowledge”

“Meta data repositories are the key to gathering, retaining and disseminating knowledge”



www.EWSolutions.com

Enterprise Warehousing Solutions, Inc. ©
2002 All Rights Reserved

Meta Data Repository ROI

Meta Data for the Business (business meta data)

Meta Data for the IT Department (technical meta data)



www.EWSolutions.com

Enterprise Warehousing Solutions, Inc. ©
2002 All Rights Reserved



Business Meta Data



Meta Data Repository ROI

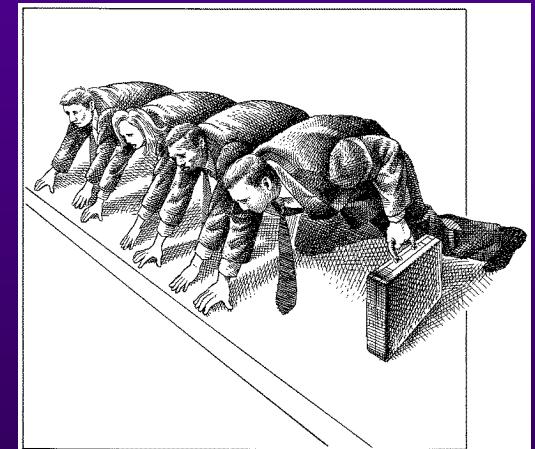
**“The Greatest Problem In Communication Is
The Illusion That It Has Been Accomplished.”**

Daniel Davenport

Meta Data Repository ROI

Meta Data for the Business (**business meta data**)

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



Meta Data for the Business

Reduces training costs

Makes the data warehouse/data mart information much more valuable as it aids analysts in making more profitable decisions

Limits incorrect decisions

Assists business analysts in finding the information they need, in a timely manner

Increases confidence in the IT system data

META DATA



Technical Users
(Developers & Analysts)



Business Users
(Executives & Business Analysts)



Project Manager



www.EWSolutions.com

Enterprise Warehousing Solutions, Inc. ©
2002 All Rights Reserved



Business Meta Data In Action



Meta Data Providing the Semantic Layer

Search



Customer Relationship & Sales Management Reports



Marketing Reports



Finance Reports



Logistics Reports



Ad-Hoc Reporting



Search

Corporate Information Access

Enter Your Search Terms Below

"Monthly Product Sales"

SEARCH

CLEAR

MORE INFORMATION ON SEARCHING

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.

NEXT 10 DOCUMENTS

1 of 4 pages, 17 total documents found



www.EWSolutions.com

Enterprise Warehousing Solutions, Inc. ©
2002 All Rights Reserved



Meta Data for the Business

2000 Monthly Global Sales Report

February 7, 2002

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

U.S. sales includes the United States, Canada, and Mexico, but does not subtract sales dollars from returned orders





Meta Data Makes for Better Decisions



Meta Data for the Business

2000 Monthly Global Sales Report

February 7, 2002

Month	Product Category	Sales \$ (in thousands) U.S	Sales \$ (in thousands) International	Sales \$ (in thousands) Total
December	TV	22,101	10,200	32,301
	VCR	11,190	4,300	15,490
	Cellular Phone	12,190	7,193	19,383
	Digital	4,002	1,301	5,303
	Miscellaneous	1,209	870	2,079
November	TV	42,000	22,200	64,200
	VCR	21,190	9,878	31,068
	Cellular Phone	28,193	12,193	40,386
	Digital	8,901	2,901	11,802
	Miscellaneous	2,730	1,530	4,260
October	TV	70,100	32,950	103,050
	VCR	31,900	178	46,778
	Cellular Phone	41,700	4,100	59,250
	Digital	20,000	4,100	24,100
	Miscellaneous	4,850	2,850	7,700

Data Quality Tracking Statistics

8.4% of the dollar value were not loaded
 1.7% of the records were not loaded

www.EWSolutions.comEnterprise Warehousing Solutions, Inc. ©
2002 All Rights Reserved



Break

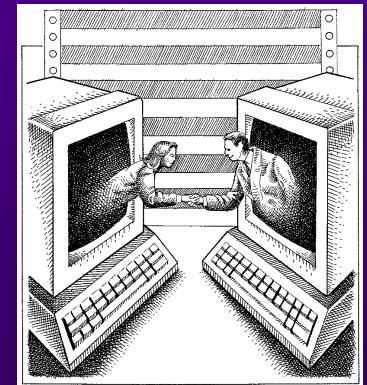


Technical Meta Data

Meta Data Repository ROI

Meta Data for the IT Department (technical meta data)

IT departments need to better manage their systems as business' demands are increasing



www.EWSolutions.com

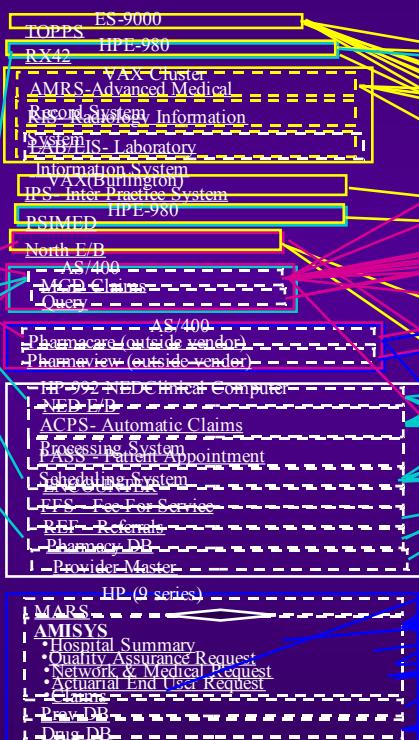
Enterprise Warehousing Solutions, Inc. ©
2002 All Rights Reserved



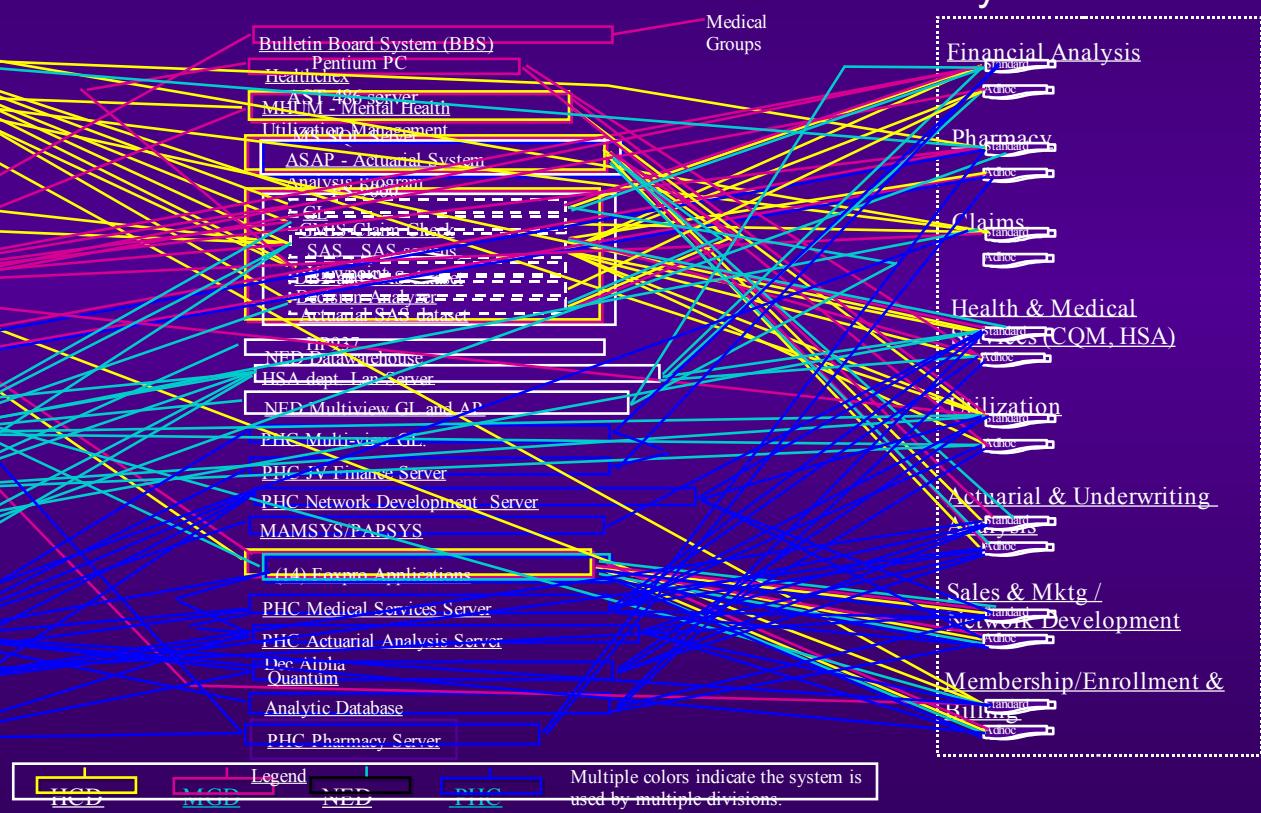
What Do Our Current IT Systems Look Like?

Islands of Data

Operational Applications



Islands of Data



Meta Data for the IT Department

IT Project Failure

1998 the Standish Group surveyed 7,000 IT projects

26% Successful

28% Failed

46% Significant Cost Overruns

1999 Standish Group IT Survey

17% Successful

34% Failed

49% Significant Cost Overruns

2000 Standish Group IT Survey

28% Successful

23% Failed

49% Significant Cost Overruns



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

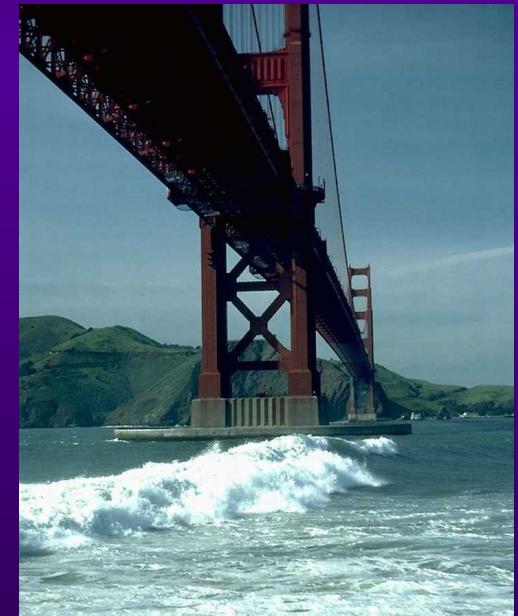
Leverage work done by other teams

Reduced rework

Research time

Reduce unproductive work

Lowers the impact of staff turnover





Technical Meta Data In Action



Meta Data for the IT Department

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

Impact Analysis Report

January 7, 2002

Source System	Source Table	Impact Field	Program Impacted	Tables/Files Impacted	Table Type	Fields Impacted
Order Entry	CUST	Customer_Name	CUSTOMER_PR02	DW_CUSTOMER	T	Cust_Name_First
						Cust_Name_Middle
						Cust_Name_Last
			CUSTOMER_PR01	I02_CUSTOMER	I	Cust_Name_First
						Cust_Name_Middle
						Cust_Name_Last
		Customer_Addr	CUSTOMER_PR02	DW_CUSTOMER	T	Cust_Name_Address
						Cust_Name_City
						Cust_Name_State
						Cust_Name_Zip
			CUSTOMER_PR01	I02_CUSTOMER	I	Cust_Name_Address
						Cust_Name_City
						Cust_Name_State
						Cust_Name_Zip

***Legend**

“T” = Target

“I” = Intermediate

“S” = Source

www.EWSolutions.comEnterprise Warehousing Solutions, Inc. ©
2002 All Rights Reserved



Meta Data for the IT Department

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

Impact Analysis Report

January 7, 2002

Field	System	Tables/Files	Fields	Domain
Customer Name	Order Entry	CUSTOMER_BILL_TO	CUST_NAME	Alphanumeric 20
		CUSTOMER_SELL_TO	CUST_NAME	Alphanumeric 20
		CUSTOMER_SHIP_TO	CUST_NAME	Alphanumeric 20
		ORDER_HEADER	CUST_NAME	Alphanumeric 20
		ORDER_DETAIL	CUST_NAME	Alphanumeric 20
	General Ledger	CUSTOMER	Cust_Name	Alphanumeric 35
		EXPENSES	Cust_Name	Alphanumeric 35
		CUST_ACCOUNTS	Cust_Name	Alphanumeric 35
	Data Warehouse	DW_CUSTOMER	Cust_Name	Alphanumeric 20
		I01_CUSTOMER	Cust_Name	Alphanumeric 20
		I02_CUSTOMER	Cust_Name	Alphanumeric 20
		I03_CUSTOMER	Cust_Name	Alphanumeric 20
Data Mart - Marketing	Data Mart - Marketing	DM_CUSTOMER	DM_Cust_Name	Alphanumeric 20
		I01_DM_CUSTOMER	DM_Cust_Name	Alphanumeric 20
		I02_DM_CUSTOMER	DM_Cust_Name	Alphanumeric 20


www.EWSolutions.com



Meta Data Repository ROI

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

**“A Meta Data Repository Is A System That
Manages Our Systems.”**

David Marco



Meta Data Fundamentals

Meta Data Consumers

Business Users

Less technical

Use business meta data

Need field definitions, report definitions and data warehouse navigation



Power Business Users

More technical

Need technical meta data for ad-hoc reporting

Technical Users

Acquisition & access developers, analysts, data modelers, architects

Need users access patterns & frequency

Need technical meta data to manage systems

External Users/Customers

Meta data consumers that are outside your company

Tend to need business meta data (definitions, descriptions, etc.)

Could be Web users

Often times are customers



Technical vs. Business Meta Data

Technical Meta Data

Technical meta data is used by the IT staff (both operational and decision support) to aid in the development and maintenance of the company's systems.



Business Meta Data

Business meta data is designed to aid the business person in understanding the data in a corporation's IT systems.

Now that we've got this warehouse how do I find the information I need?





Examples of Technical Meta Data

User access patterns, frequency, and execution time of reports/queries

Audit controls, data quality, and balancing information

Systems of record feeding the data warehouse

Identification of source system fields

Mappings & transformations from the system of record to the data warehouse/operational systems

Encoding/reference table conversions

Physical and logical data models (relational, multidimensional, object, etc.)

Entities, attributes, keys and indexes

Subject areas

History of extracts

Data warehouse archiving

Job and data dependencies

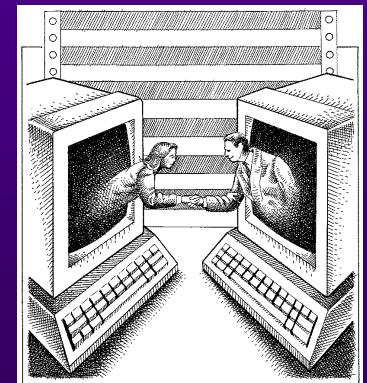
Program names and descriptions

Version maintenance

Technical domain values

Security rules and administration

Purge criteria



Examples of Technical Meta Data for IT Portfolio Management

Hardware assets (mainframes, Unix, servers, PCs, etc.)

Hardware configurations (disk space, memory, processor types, etc.)

Hardware locations

Hardware costs (purchase price, leasing fees, maintenance fees, etc.)

Software licenses

Software license expiration dates

Software installations (purchase price, leasing fees, maintenance fees, etc.)

Software costs

Installed software patches

System listings

System technology rankings (E = evolving, S = stable, A = aging, O = obsolete/unsupported)

System purposes

System inputs/outputs

Project listings

Project costs (estimates and actuals)(internal, consulting, hardware, and software)

Project success rates

Project staffing (internal, vendors, temporary, etc.)

Project estimated date of completion, actual date of completion, etc.

Project business justification and scope

Project status

Network links

Contact person





Examples of Business Meta Data

The structure of data as known to the business analyst

Common routines for the access of information

Subject areas

Business table names and **business definitions**

Business attribute names and **business definitions**

Data warehouse field mappings, transformations, and summarizations in business terms

Rules for drill down, drill up, drill across, etc.

Domain values

Units of measure

Data stewardship

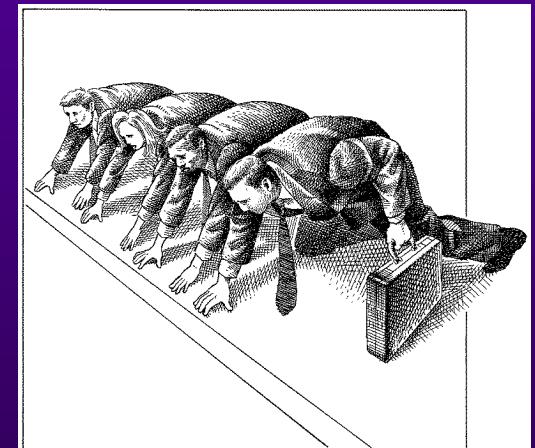
Audit trails

Data location

Data warehouse refresh dates

Data security policies

Organization hierarchies and contact trees





Data Heritage and Data Lineage

Data Heritage

The original source of the data. For example, a sales person types in the customer name in the Sales system or CUST_ID is a sequential number assigned by the Sales system

Data Lineage

Everything that has “happened” to the data. Whether it was moved from one system to another, transformed, aggregated, etc. ETL (extraction, transformation, and load) tools look to electronically capture this meta data





Meta Data and External Data

External Data

Data sources not controlled by the enterprise

Types of External Data

Federal organizations

Vendor systems

Customer systems

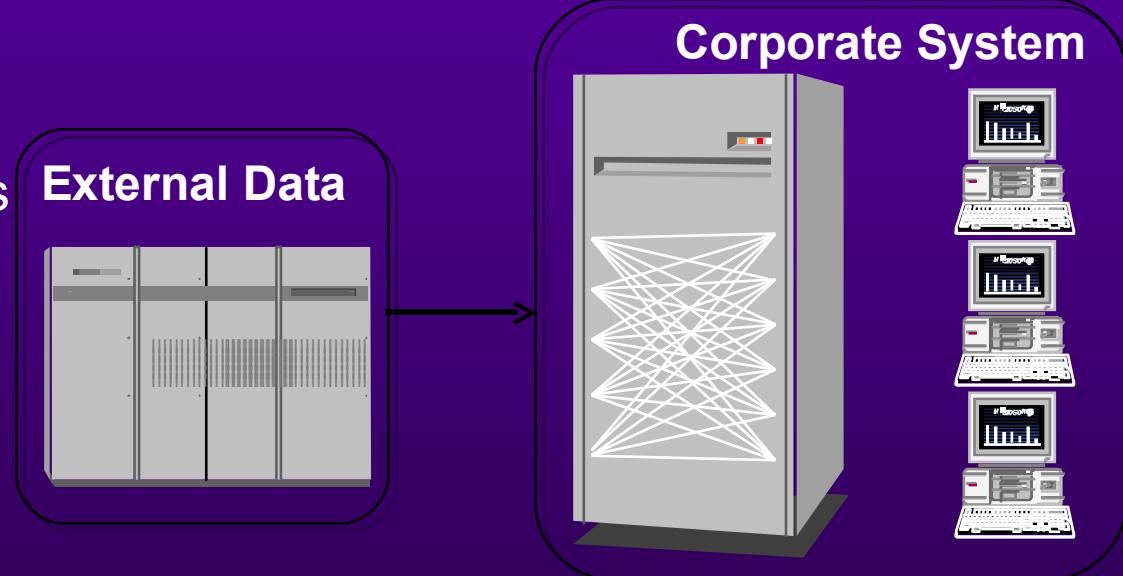
Websites

Dun & Bradstreet reports

Dow Jones reports

White papers

Articles in magazines





Meta Data and External Data

External Meta Data

Document id

Date and time of entry

Source of the external data

Classification of the external data

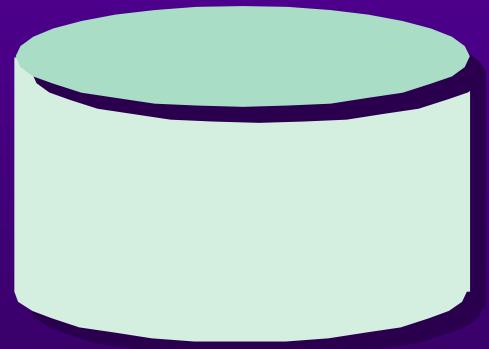
Errors received

Index words

Purge date

Physical location reference

Length of the external data





What To Avoid When Building A Meta Data Repository



What To Avoid

Top 10 List

- #10 Meta data repository team creates standards none of the supporting teams can follow

- #9 Not creating a meta data repository team and/or trivializing the meta data repository effort

- #8 Failing to have an experienced meta data project manager/architect leading the project

- #7 Neglecting to establish good business policies



www.EWSolutions.com

Enterprise Warehousing Solutions, Inc. ©
2002 All Rights Reserved





What To Avoid

Top 10 List

- #6 The meta data repository team reporting to the project manager of the data warehouse team
- #5 Too many manual meta data integration processes
- #4 Not providing EASY access to the meta data
- #3 Selecting a repository tool without conducting an evaluation or defining requirements
- #2 Trying to do a “Big Bang” (waterfall) implementation...**Boiling the Ocean**



www.EWSolutions.com

Enterprise Warehousing Solutions, Inc. ©
2002 All Rights Reserved



What To Avoid

Number #1 Reason

Not defining tangible business/technical objectives of the meta data repository!!



www.EWSolutions.com

Enterprise Warehousing Solutions, Inc. ©
2002 All Rights Reserved



Lunch



Architecture Approaches

Centralized

Decentralized

Distributed



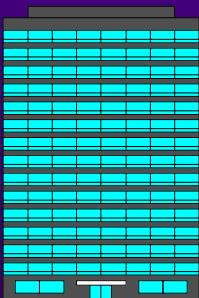
www.EWSolutions.com

Enterprise Warehousing Solutions, Inc. ©
2002 All Rights Reserved

Centralized Architecture

Centralized Approach: All meta data is stored and shared in this one centralized repository

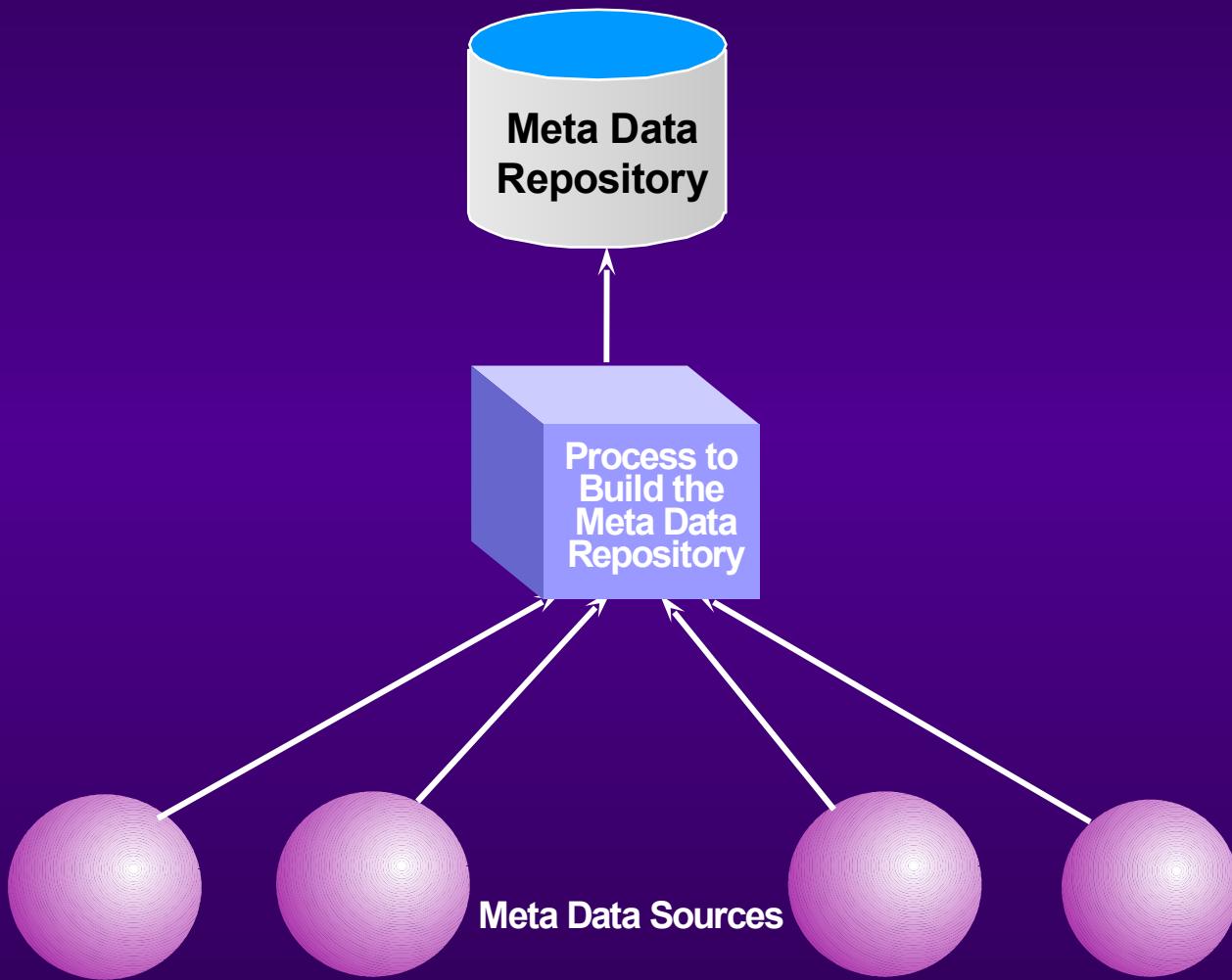
Very popular for companies that do not have very distinct lines of business



www.EWSolutions.com

Enterprise Warehousing Solutions, Inc. ©
2002 All Rights Reserved

Centralized Architecture





Decentralized Architecture

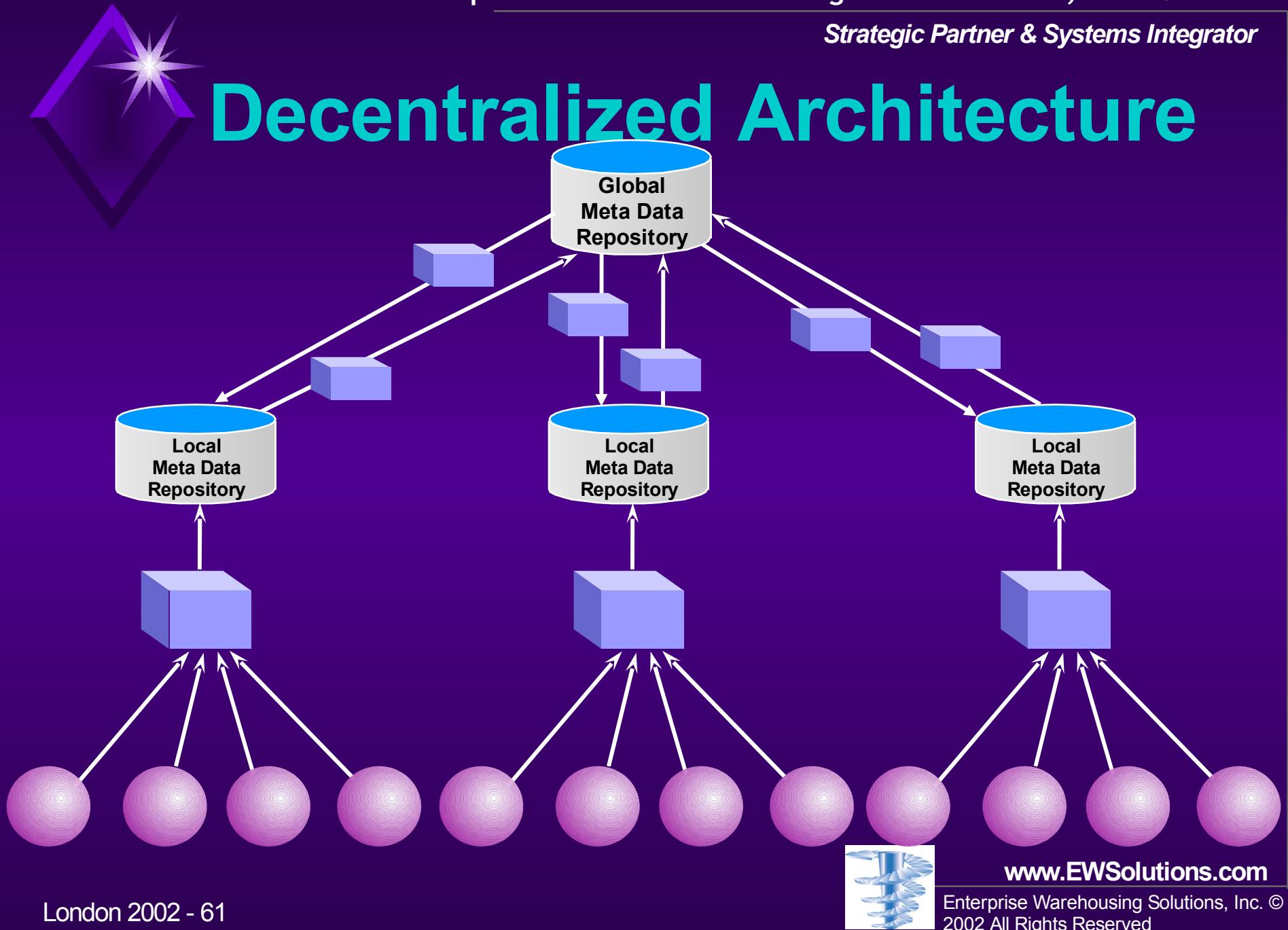
Decentralized Approach: Creates a global (central) meta data repository that shares meta data elements that also appear in local meta data repositories
Meta data is shared and re-used between the different local repositories by first going through the global repository

The global repository is a subset of the local

Sharing and access to the local meta data is independent of the global repository
Allows each local repository to be autonomous for its own content and administration requirements



Decentralized Architecture





Distributed Architecture

Distributed Approach: The meta data is captured and presented to the user at the time of the request

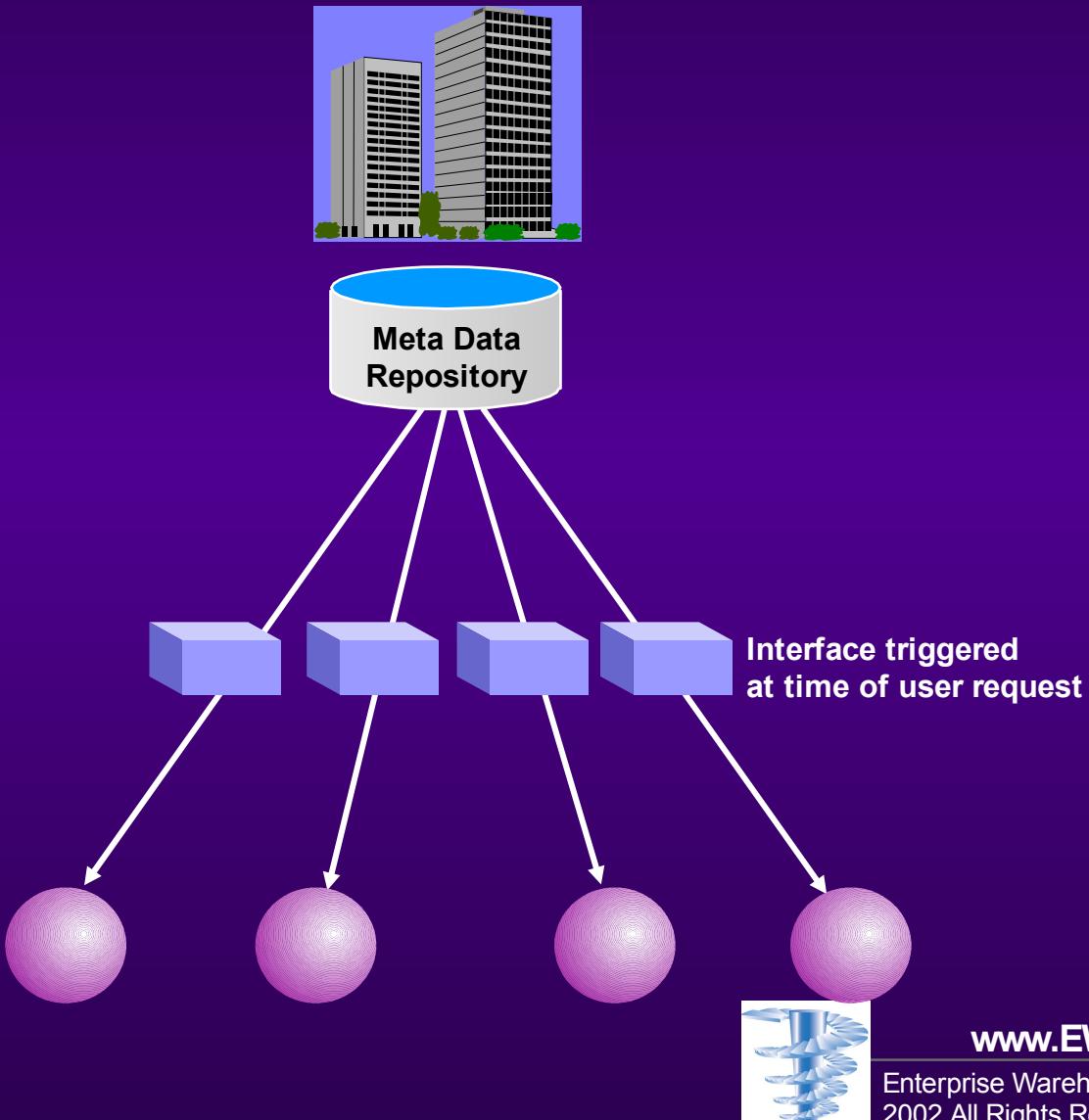
Newer concept that is fairly unproven

Cannot be used for all meta data requirements

Can be used in conjunction with centralized and decentralized



Distributed Architecture





Meta Data Tools



www.EWSolutions.com

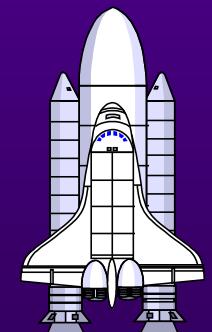
Enterprise Warehousing Solutions, Inc. ©
2002 All Rights Reserved



Meta Data Tools

Meta data integration tools (loads the meta data repository)

Meta data access tools (accesses the meta data repository)



www.EWSolutions.com

Enterprise Warehousing Solutions, Inc. ©
2002 All Rights Reserved



Meta Data Integration Tools

Product	Vendor
Rochade	ASG
Advantage Repository for Distributed Systems (Advantage/DS)	Computer Associates
Advantage Repository for 390 (Advantage/390)	Computer Associates
MetaCenter	Data Advantage Group
MetaStage (part of DataStage)	Ascential Software
MetadataExchange (MX2)(part of PowerCenter)	Informatica

If you are assessing these vendor's products you may be interested in a third-party evaluation. Information on EWS' 146-page comparative study of these products can be found on the EWS website at http://www.ewssolutions.com/research_paper.asp or by emailing JFaulkenberry@EWSolutions.com

Certified Sources

Certified Sources

Meta data that the tool can:

- directly read
- properly interpret the information
- load it into the correct attributes of the meta model

Easily integrated and do not require an extension to the base meta model

Typically no additional programming or analysis required

Examples:

- Modeling tools (technical meta data)
- Database log files
- ETL tools (transformation rules)

Normally a repository tool is certified for several vendor tools in each of these categories





Generic Sources

Generic Sources

Meta data sources that are in a common formation (i.e. tab, space, or comma delimited)

Most tools allow for one or more generic meta data sources

Meta data that the tool can:

- easily read source

- cannot properly interpret the information

- programming is still needed to map source elements to the correct attributes in the meta model

Meta model may require extensions

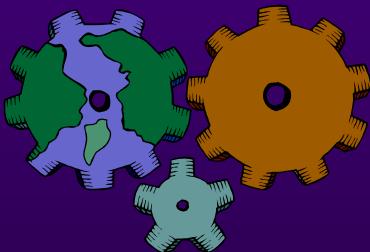
Examples:

- databases (technical meta data)

- spreadsheets (technical and business meta data)

Meta model may require extensions

Programming and analysis ranges from fairly simple to quite complicated



Non-Supported Sources

Non-Supported Sources

Meta data sources that are in a format that the tool cannot read

Meta data that the tool can:

- cannot directly read source (additional programming language required)

- cannot properly interpret the information

- programming needed to map source elements to the correct attributes in the meta model

Meta model may require extensions

Examples:

- closed databases (technical meta data)

- 3rd party applications

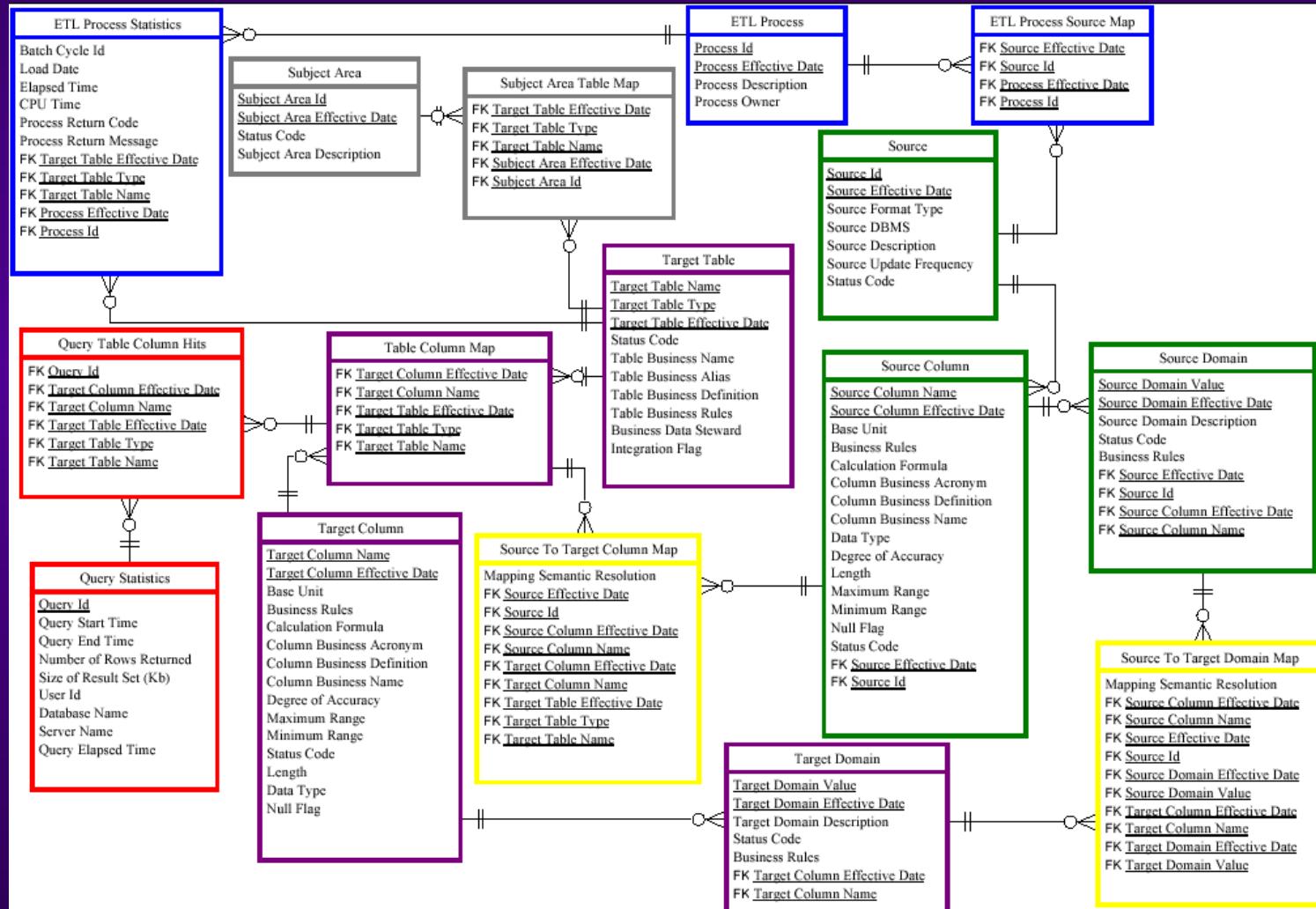
Programming and analysis ranges from fairly simple to quite complicated



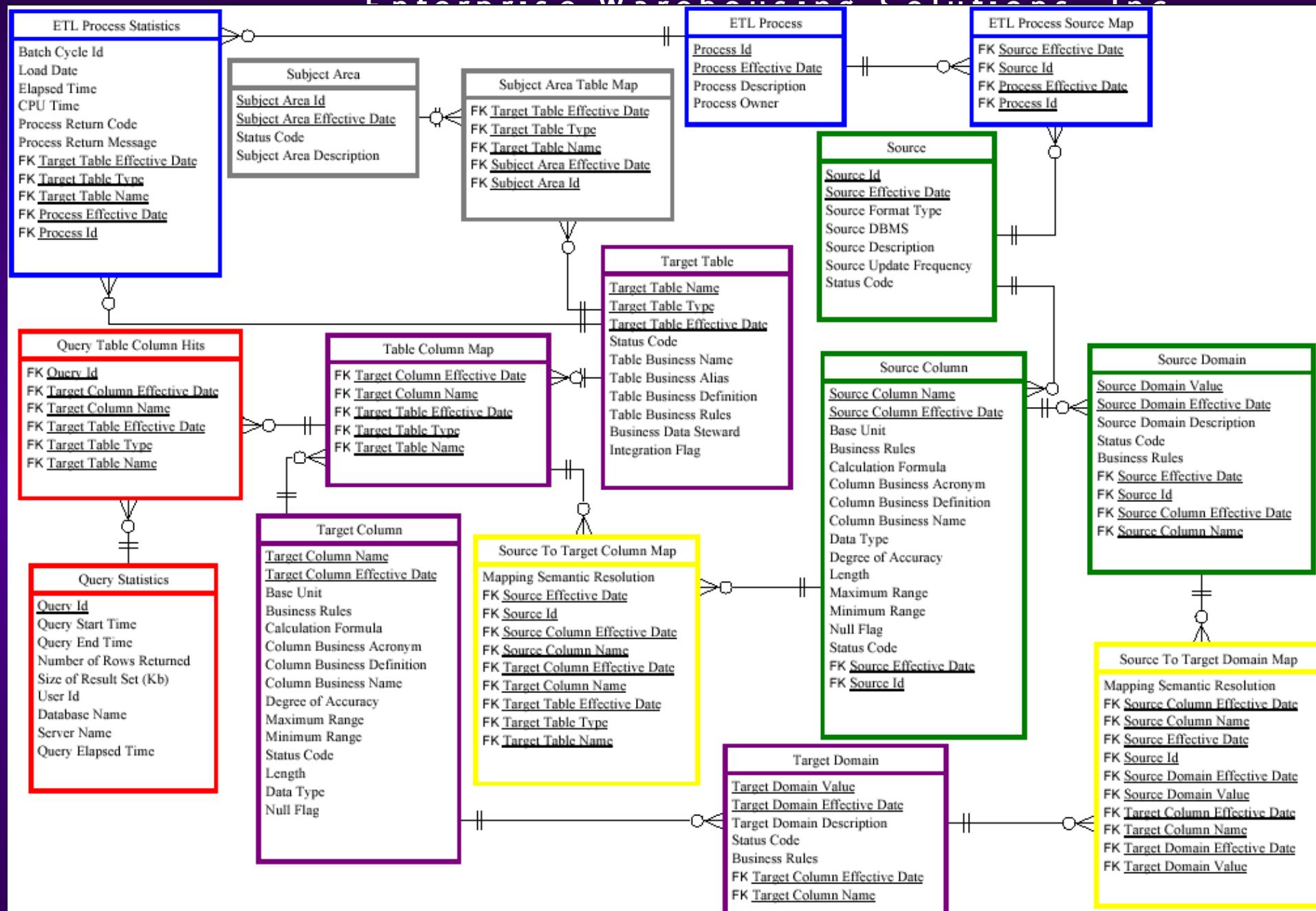


Generic Meta Model Walkthrough

Generic Meta Data Repository Model



Enterprise Warehousing Solutions



Model Provisions

Not a shrink wrap solution, just a starting point

Focused on data warehousing

Includes base components (subject areas) typically required in a warehouse repository

Can be integrated into an internal implementation effort

Does not include sufficient detail or functionality for physical implementation (e.g., versioning or source control)



Possible Model Modifications

Example of potential changes include:

Adherence to internal data modeling practices/standards

Additional meta data requirements



www.EWSolutions.com

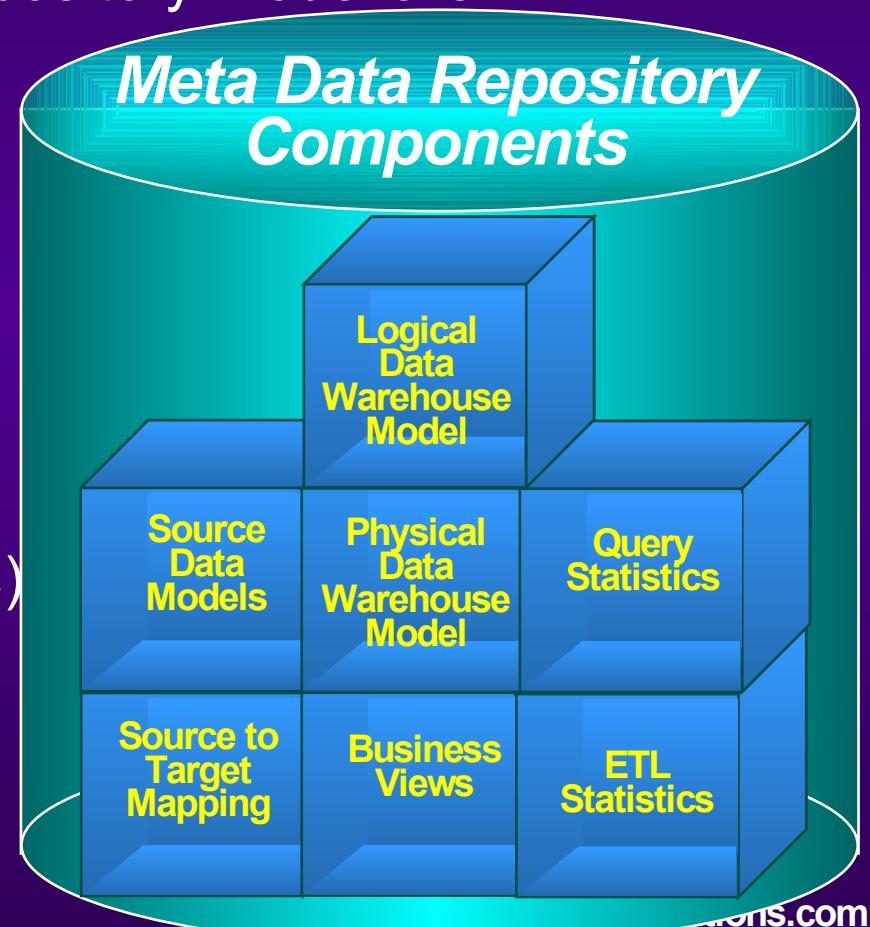
Enterprise Warehousing Solutions, Inc. ©
2002 All Rights Reserved



Model Components

Seven components of the generic repository model are:

- Logical Data Warehouse Model
- Physical Data Warehouse Model
- Source Data Model
- Source-To-Target Mapping
- Logical Grouping (Business Views)
- ETL Statistics
- Query Statistics





Logical DW Model

Contains all business entities for specific business views

Relationships between entities and attributes

Company logical data modeling practices may necessitate additions/changes

Table Column Map
<u>FK Target Column Effective Date</u>
<u>FK Target Column Name</u>
<u>FK Target Table Effective Date</u>
<u>FK Target Table Type</u>
<u>FK Target Table Name</u>

Target Table
<u>Target Table Name</u>
<u>Target Table Type</u>
<u>Target Table Effective Date</u>
Status Code
Table Business Name
Table Business Alias
Table Business Definition
Table Business Rules
Business Data Steward
Integration Flag

Target Column
<u>Target Column Name</u>
<u>Target Column Effective Date</u>
Base Unit
Business Rules
Calculation Formula
Column Business Acronym
Column Business Definition
Column Business Name
Degree of Accuracy
Maximum Range
Minimum Range
Status Code
Length
Data Type
Null Flag

Target Domain
<u>Target Domain Value</u>
<u>Target Domain Effective Date</u>
Target Domain Description
Status Code
Business Rules
<u>FK Target Column Effective Date</u>
<u>FK Target Column Name</u>

* Note: *Logical and Physical Components are differentiated in the model through use of the Target Table Type column (e.g., L = Logical)*





Physical DW Model

Contains detailed information about physical data warehouse target model

Model component may be extended to accommodate your specific environment (e.g., indexes, etc)

Table Column Map
<u>FK Target Column Effective Date</u>
<u>FK Target Column Name</u>
<u>FK Target Table Effective Date</u>
<u>FK Target Table Type</u>
<u>FK Target Table Name</u>

Target Column
<u>Target Column Name</u>
<u>Target Column Effective Date</u>
Base Unit
Business Rules
Calculation Formula
Column Business Acronym
Column Business Definition
Column Business Name
Degree of Accuracy
Maximum Range
Minimum Range
Status Code
Length
Data Type
Null Flag

Target Table
<u>Target Table Name</u>
<u>Target Table Type</u>
<u>Target Table Effective Date</u>
Status Code
Table Business Name
Table Business Alias
Table Business Definition
Table Business Rules
Business Data Steward
Integration Flag

Target Domain
<u>Target Domain Value</u>
<u>Target Domain Effective Date</u>
Target Domain Description
Status Code
Business Rules
<u>FK Target Column Effective Date</u>
<u>FK Target Column Name</u>

* Note: Logical and Physical Components are differentiated in the model through use of the Target Table Type column (e.g., P = Physical)





Source Data

Contains physical information about the various sources of data feeding the data warehouse

Information can originate from databases, file extracts, spreadsheets, Internet, and other formats

Information can be used to alert data warehouse team of potential changes that will impact the DW model, ETL processes, and reports

Information can be utilized by both business and technical analysts for reconciliation and auditing

Source
<u>Source Id</u>
<u>Source Effective Date</u>
Source Format Type
Source DBMS
Source Description
Source Update Frequency
Status Code

Source Column
<u>Source Column Name</u>
<u>Source Column Effective Date</u>
Base Unit
Business Rules
Calculation Formula
Column Business Acronym
Column Business Definition
Column Business Name
Data Type
Degree of Accuracy
Length
Maximum Range
Minimum Range
Null Flag
Status Code
<u>FK Source Effective Date</u>
<u>FK Source Id</u>

Source Domain
<u>Source Domain Value</u>
<u>Source Domain Effective Date</u>
Source Domain Description
Status Code
Business Rules
<u>FK Source Effective Date</u>
<u>FK Source Id</u>
<u>FK Source Column Effective Date</u>
<u>FK Source Column Name</u>





Source-To-Target Mapping

Contains the cross reference mapping and semantic resolution between the source operational systems and the target physical data warehouse model

Key feature of the repository

Level of detail for the semantic resolution description will vary based on firm's goals

Model could be extended here to adopt a more granular approach toward business rules capture (separate component with associative relationships)

Source To Target Column Map

Mapping Semantic Resolution
FK Source Effective Date
FK Source Id
FK Source Column Effective Date
FK Source Column Name
FK Target Column Effective Date
FK Target Column Name
FK Target Table Effective Date
FK Target Table Type
FK Target Table Name

Source To Target Domain Map

Mapping Semantic Resolution
FK Source Column Effective Date
FK Source Column Name
FK Source Effective Date
FK Source Id
FK Source Domain Effective Date
FK Source Domain Value
FK Target Domain Effective Date
FK Target Domain Value



Business Views

Contains the logical grouping (subject areas) of the physical data warehouse tables (e.g., financial, sales, human resources)

Provides the business users with an intuitive method/view of navigating the information stored in the warehouse

Specific business requirements may dictate a more hierarchical method to store and navigate subject area information

Subject Area
<u>Subject Area Id</u>
<u>Subject Area Effective Date</u>
Status Code
Subject Area Description

Subject Area Table Map
<u>FK Target Table Effective Date</u>
<u>FK Target Table Type</u>
<u>FK Target Table Name</u>
<u>FK Subject Area Effective Date</u>
<u>FK Subject Area Id</u>





ETL Statistics

Contains information on individual ETL processes and operational statistics

Used to determine process improvements, database enhancements, fault isolation and other optimization procedures

Use of this component will depend greatly on specific ETL method and tool set(s) capabilities

ETL Process Statistics
Batch Cycle Id
Load Date
Elapsed Time
CPU Time
Process Return Code
Process Return Message
<u>FK Target Table Effective Date</u>
<u>FK Target Table Type</u>
<u>FK Target Table Name</u>
<u>FK Process Effective Date</u>
<u>FK Process Id</u>

ETL Process
<u>Process Id</u>
<u>Process Effective Date</u>
Process Description
Process Owner

ETL Process Source Map
<u>FK Source Effective Date</u>
<u>FK Source Id</u>
<u>FK Process Effective Date</u>
<u>FK Process Id</u>



Query Statistics

Operational information about every query made against the data warehouse database is captured here

Used to determine optimizations that need to be performed against the warehouse based on the operational statistics

Statistics are used to identify dormant data, aggregations and index candidates

Use of component will depend on data monitoring capabilities in the warehouse environment

Query Table Column Hits
<u>FK Query Id</u>
<u>FK Target Column Effective Date</u>
<u>FK Target Column Name</u>
<u>FK Target Table Effective Date</u>
<u>FK Target Table Type</u>
<u>FK Target Table Name</u>

Query Statistics
Query Id
Query Start Time
Query End Time
Number of Rows Returned
Size of Result Set (Kb)
User Id
Database Name
Server Name
Query Elapsed Time



Break



Model Detail

Logical & Physical DW Model Tables

Associative table that simply provides a cross reference of the logical/physical tables to their associated columns

Table Column Map
<u>FK Target Column Effective Date</u>
<u>FK Target Column Name</u>
<u>FK Target Table Effective Date</u>
<u>FK Target Table Type</u>
<u>FK Target Table Name</u>

Entry point table for the component. Stores both the logical and physical views of the warehouse tables through use of the Target Table Type code column.

Target Table
<u>Target Table Name</u>
<u>Target Table Type</u>
<u>Target Table Effective Date</u>
Status Code
Table Business Name
Table Business Alias
Table Business Definition
Table Business Rules
Business Data Steward
Integration Flag

Attribute table that contains both technical and business meta data about the column in the logical/physical table

Target Column
<u>Target Column Name</u>
<u>Target Column Effective Date</u>
Base Unit
Business Rules
Calculation Formula
Column Business Acronym
Column Business Definition
Column Business Name
Degree of Accuracy
Maximum Range
Minimum Range
Status Code
Length
Data Type
Null Flag

Target Domain
<u>Target Domain Value</u>
<u>Target Domain Effective Date</u>
Target Domain Description
Status Code
Business Rules
<u>FK Target Column Effective Date</u>
<u>FK Target Column Name</u>

Attribute table that contains both technical and business meta data about the definition of values for columns containing codes



Target Table Detail

Target Table Name - contains either the logical or the actual physical, database name, of the warehouse table

Target Table Effective Date - provides a mean of distinguishing revisions made to a particular table

Status Code - indicates the current approval status of the warehouse table using domain values such as pending, approved or removed.

Table Business Name - contains the common name of the table that will be presented to users through the front-end access method of the warehouse

Table Business Alias - contains a common alternative name or acronym of the table (e.g., alias for profitability ratio is PR)

Table Business Definition - contains a detailed descriptive explanation of the business process the table plays in the enterprise allowing the end user to make an educated determination about its use

Table Business Rules - is used to denote any conventions or practices the table must adhere to beyond referential integrity constraints

Business Data Steward - lists the individual and/or groups that are responsible in the organization for defining this particular table

Integration Flag - is used as a quick query method to denote which tables obtain source information from more than one system

Target Table
<u>Target Table Name</u>
<u>Target Table Type</u>
<u>Target Table Effective Date</u>
Status Code
Table Business Name
Table Business Alias
Table Business Definition
Table Business Rules
Business Data Steward
Integration Flag





Target Column Detail

Target Column Name - contains either the logical or the actual physical, database name, of the warehouse field

Target Column Effective Date - provides a mean of distinguishing revisions made to a particular field

Base Unit - contains a description or code of the measurement unit used for entries in the field (e.g., dollars, pounds, meters)

Business Rules - used to denote any conventions or practices the column must follow. For example, an amount field must have a corresponding currency value

Calculation Formula - contains a description of the method used to determine the value stored in the field including the names of other tables/columns used

Business Acronym - contains a common acronym coding of the field (e.g., UOM for unit of measure)

Business Definition - contains a detailed descriptive explanation of the business meaning of the column in the context of the enterprise

Target Column
<u>Target Column Name</u>
<u>Target Column Effective Date</u>
Base Unit
Business Rules
Calculation Formula
Column Business Acronym
Column Business Definition
Column Business Name
Degree of Accuracy
Maximum Range
Minimum Range
Status Code
Length
Data Type
Null Flag





Target Column Detail (continued)

Business Name - contains the common name of the column that will be shown to users through the front-end access method of the warehouse

Degree of Accuracy - is used to denote the required number of places after the decimal point that entries in this column are required to contain. For example, currency conversions involving the Euro monetary unit should contain six decimals of precision to accurately convert

Maximum Range and **Minimum Range** - indicate the upper and lower bounds of numeric values for the column

Status Code - indicates the current approval status of the column (e.g., pending, approved, removed)

Length, Data Type and **Null Flag** - provide the basic technical information required to describe the fields into a data modeling tool or to a DBMS. For example, a company name column could be described as having a length of hundred, data type of character or varchar and does not allow null values

Target Column
<u>Target Column Name</u>
<u>Target Column Effective Date</u>
Base Unit
Business Rules
Calculation Formula
Column Business Acronym
Column Business Definition
Column Business Name
Degree of Accuracy
Maximum Range
Minimum Range
Status Code
Length
Data Type
Null Flag





Target Domain Detail

Target Domain Value - code for a column. For example, the column country code has domain values that include USA for United States and CAN for Canada, etc.

Target Domain Effective Date - provides a mean of distinguishing revisions made to a particular code

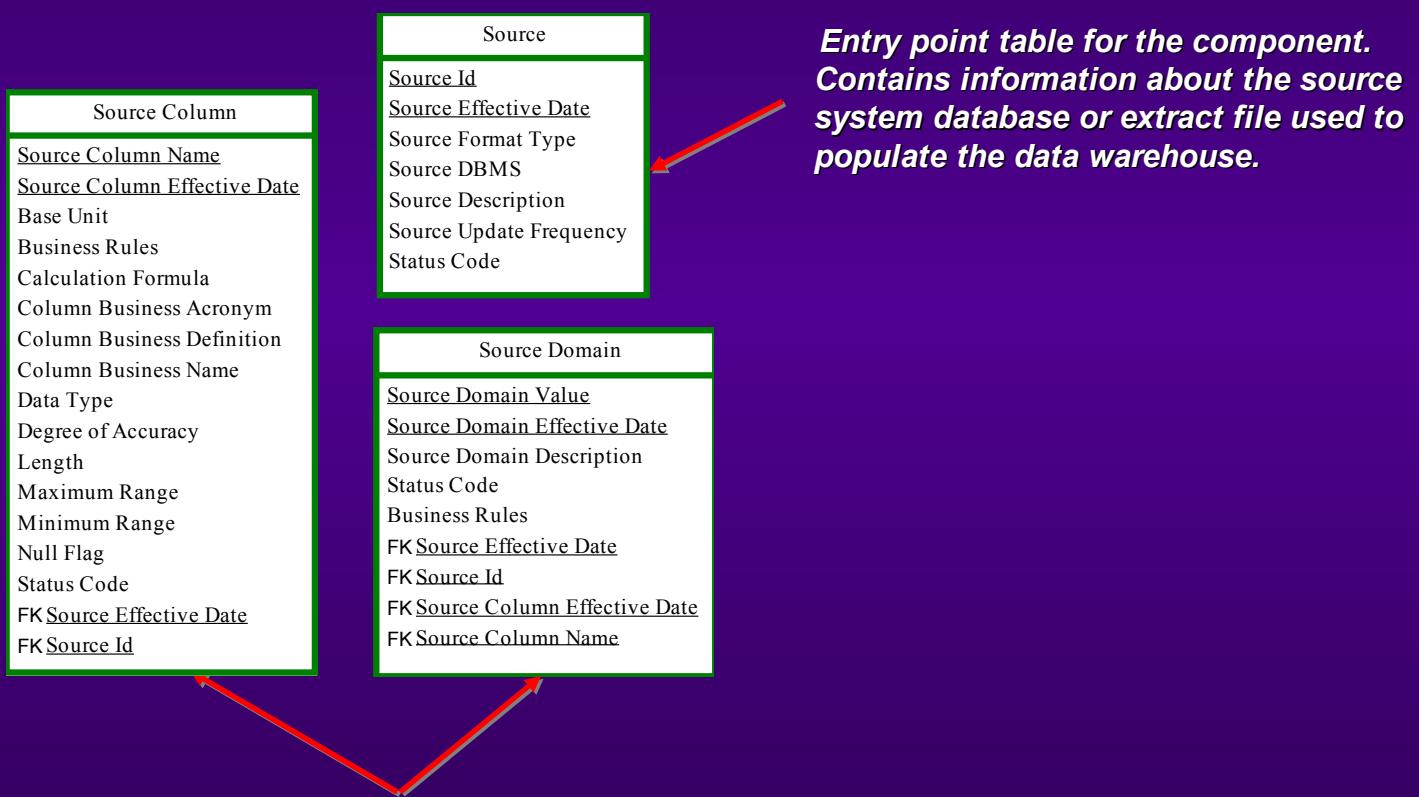
Target Domain Description - contains a detailed descriptive explanation of the business meaning of the code

Status Code - indicates the current approval status of the code (e.g., pending, approved, removed)

Business Rules -used to denote any conventions or practices the code must follow

Target Domain
<u>Target Domain Value</u>
<u>Target Domain Effective Date</u>
Target Domain Description
Status Code
Business Rules
<u>FK Target Column Effective Date</u>
<u>FK Target Column Name</u>

Source Data Tables





Source Detail

Source ID - used to uniquely identify a particular system of record (e.g., ERP, Order Management 1, Trouble Ticket, Client Extract 5)

Source Effective Date - provides a mean of distinguishing revisions made to an operational system or extract file

Source Format Type - provides a means to identify the category of the source information such as a server/database, directory/file or spreadsheet file

Source DBMS - contains the physical names of these items for cases where operational information is being extracted directly from a source database table

Source Description - contains a detailed narrative describing the particular source of information for the warehouse

Source Update Frequency - denotes the cycle at which this source is updated (e.g., daily, weekly, monthly, annually)

Status Code - indicates the current approval status of the source system or extract file (e.g., pending, approved, cancelled)

Source
<u>Source Id</u>
<u>Source Effective Date</u>
Source Format Type
Source DBMS
Source Description
Source Update Frequency
Status Code





Source-To-Target Component Tables

Source To Target Column Map

Mapping Semantic Resolution
FK Source Effective Date
FK Source Id
FK Source Column Effective Date
FK Source Column Name
FK Target Column Effective Date
FK Target Column Name
FK Target Table Effective Date
FK Target Table Type
FK Target Table Name

Source To Target Domain Map

Mapping Semantic Resolution
FK Source Column Effective Date
FK Source Column Name
FK Source Effective Date
FK Source Id
FK Source Domain Effective Date
FK Source Domain Value
FK Target Domain Effective Date
FK Target Domain Value

Mapping Semantic Resolution - for both tables, describes any additional instructions required for cross referencing source to target columns or domains

Tables provide necessary design strategy to build ETL processes by linking the operational sources to the warehouse tables. Associative mapping (linking) of columns and domains primary purpose of these two tables.





Source-To-Target Mapping Report Example

Source/Target Data Map Report											January 17, 2000		
Source Data Model/Extract File Format						Target Data Model							
Source System	Table Name	Column Name	Data Type	Length	Domain(s)	Table Name	Column Name	Data Type	Length & Accuracy	Domain(s)	Semantic Resolution	Creation Date	Update Date
ERP	Customer	CUSTOMER	char	100	N/A	Customer	CUSTOMER NAME	char	100	N/A	Business requirement has the ERP source system as primary source	March 22, 1999	January 3, 2000
Order Processing	GS.ORDERP 01.DWEXTRA CT(0)	CUSTOMER	char	50	N/A								
Sales Force	Client	CLIENT DESCRIPTION	char	30	N/A								
Order Processing	GS.ORDER02 .DWEXTRACT (0)	PRODUCT	char	50	N/A	Product	PRODUCT NAME	char	50	N/A	Business requirement has the Order Processing source system as primary source	January 3, 2000	January 3, 2000
Sales Force	Product	PRODUCT DESCRIPTION	char	30	N/A								
Order Processing	GS.ORDERP 01.DWEXTRA CT(0)	CUSTOMER CODE	char	2	AC (Active Customer)	Customer	CUSTOMER TYPE	char	1	D (Direct)	Business requirement has the Sales Force source system as primary source	January 3, 2000	January 3, 2000
Order Processing	GS.ORDERP 01.DWEXTRA CT(0)	CUSTOMER CODE	char	2	AA (Active Affiliate)					I (Indirect)			
Order Processing	GS.ORDERP 01.DWEXTRA CT(0)	CUSTOMER CODE	char	2	AS (Active Subsidiary)					I (Indirect)			
Sales Force	Client	CLIENT PRIORITY CODE	char	1	I (Indirect)					I (Indirect)			
Sales Force	Client	CLIENT PRIORITY CODE	char	1	D (Direct)					D (Direct)			


www.EWSolutions.com

 Enterprise Warehousing Solutions, Inc. ©
2002 All Rights Reserved

ETL Statistics Tables

Associates the ETL process to the target table needed for particular transformation procedure plus additionally capture operational statistics

ETL Process Statistics	
Batch Cycle Id	
Load Date	
Elapsed Time	
CPU Time	
Process Return Code	
Process Return Message	
<u>FK Target Table Effective Date</u>	
<u>FK Target Table Type</u>	
<u>FK Target Table Name</u>	
<u>FK Process Effective Date</u>	
<u>FK Process Id</u>	

Entry point table for the component. Stores information on the actual each ETL process

ETL Process	
Process Id	
<u>Process Effective Date</u>	
Process Description	
Process Owner	

ETL Process Source Map	
<u>FK Source Effective Date</u>	
<u>FK Source Id</u>	
<u>FK Process Effective Date</u>	
<u>FK Process Id</u>	

Associates the ETL process to the source operational data needed for particular transformation procedure



ETL Process Detail

Process Id - uniquely identify the ETL procedure in the warehouse environment

Process Effective Date - identifies when this particular process was included into the load cycle steps

Process Description - contains a detailed description of the ETL process from a technical and business perspective

Process Owner - lists the individual and/or groups that are technically responsible for this process

ETL Process
<u>Process Id</u>
<u>Process Effective Date</u>
Process Description
Process Owner



ETL Process Statistics Detail

Batch Cycle ID - a sequential identifier assigned during each load cycle to the data warehouse regardless of the refresh frequency (e.g., daily, weekly, monthly, etc.)

Load Date - provides a date/time stamp of when the specific ETL process ran

Elapsed Time and CPU Time - provide information on the computer resources that the ETL process used during this batch cycle

Process Return Code - captures the completion, warning or error return code value from the ETL process for examination by the data acquisition developer

Process Return Code Message - contains a narrative explanation of warning or error return codes for the particular operating system (if available)

ETL Process Statistics
Batch Cycle Id
Load Date
Elapsed Time
CPU Time
Process Return Code
Process Return Message
FK <u>Target Table Effective Date</u>
FK <u>Target Table Type</u>
FK <u>Target Table Name</u>
FK <u>Process Effective Date</u>
FK <u>Process Id</u>





ETL Statistics Report Example

ETL Statistics Report										January 17, 2000	
Target Table	ETL Process	Target Table Type	Source System(s)	Processing Time	Load Cycle	Number of Records Inserted	Number of Records Updated	Total Number of Records	Load Date	Load Time	
Customer	cs0001	Dimension	ERP	0:43:12	3	223	196	1,194	March 2, 2000	5:18:10	
Customer	cs0002	Dimension	Order Processing	0:27:39	3	22	5	1,194	March 2, 2000	5:18:10	
Customer	cs0003	Dimension	Sales Force	0:18:45	3	51	17	1,194	March 2, 2000	5:18:10	
Product	pd001	Dimension	Order Processing	0:29:28	3	41	28	112	March 2, 2000	5:18:10	
Employee	ee001	Dimension	ERP	0:37:55	3	23	15	276	March 2, 2000	5:18:10	
Time	tm001	Dimension	External Source	0:05:41	3	1	0	3	March 2, 2000	5:18:10	
Sales	sl001	Fact	Order Processing	2:35:30	3	36,908,928	0	58,095,382	March 2, 2000	5:18:10	
Customer	cs0001	Dimension	ERP	0:41:45	2	287	201	898	February 1, 2000	3:36:38	
Customer	cs0002	Dimension	Order Processing	0:25:32	2	15	1	898	February 1, 2000	3:36:38	
Customer	cs0003	Dimension	Sales Force	0:16:48	2	78	24	898	February 1, 2000	3:36:38	
Product	pd001	Dimension	Order Processing	0:28:34	2	31	24	71	February 1, 2000	3:36:38	
Employee	ee001	Dimension	ERP	0:34:59	2	9	2	253	February 1, 2000	3:36:38	
Time	tm001	Dimension	External Source	0:04:13	2	1	0	2	February 1, 2000	3:36:38	
Sales	sl001	Fact	Order Processing	1:04:47	2	16,130,774	0	21,186,454	February 1, 2000	3:36:38	
Customer	cs0001	Dimension	ERP	0:49:42	1	334	0	518	January 3, 2000	3:47:35	
Customer	cs0002	Dimension	Order Processing	0:28:21	1	37	0	518	January 3, 2000	3:47:35	
Customer	cs0003	Dimension	Sales Force	0:31:09	1	147	0	518	January 3, 2000	3:47:35	
Product	pd001	Dimension	Order Processing	0:45:33	1	40	0	40	January 3, 2000	3:47:35	
Employee	ee001	Dimension	ERP	0:50:18	1	244	0	244	January 3, 2000	3:47:35	
Time	tm001	Dimension	External Source	0:02:21	1	1	0	1	January 3, 2000	3:47:35	
Sales	sl001	Fact	Order Processing	0:20:11	1	5,055,680	0	5,055,680	January 3, 2000	3:47:35	





Business View Tables

Entry point table for this component. Stores logical grouping or subject area view of the physical data warehouse tables.

Subject Area
Subject Area Id
Subject Area Effective Date
Status Code
Subject Area Description

Subject Area Table Map
FK Target Table Effective Date
FK Target Table Type
FK Target Table Name
FK Subject Area Effective Date
FK Subject Area Id

Associates the subject area view to the actual target warehouse table (logical or physical)



Subject Area Detail

Subject Area Id - uniquely identifies the group of tables within the enterprise from a business perspective

Subject Area Effective Date - identifies when this particular business view was last updated or created

Status Code - indicates the current approval status of the grouping (e.g., pending, approved, cancelled)

Subject Area Description - contains a detailed description of the business grouping

Subject Area
<u>Subject Area Id</u>
<u>Subject Area Effective Date</u>
Status Code
Subject Area Description



Query Statistics Tables

Associates the query to the physical target warehouse table and column



Query Table Column Hits
FK Query Id
FK Target Column Effective Date
FK Target Column Name
FK Target Table Effective Date
FK Target Table Type
FK Target Table Name

Query Statistics
Query Id
Query Start Time
Query End Time
Number of Rows Returned
Size of Result Set (Kb)
User Id
Database Name
Server Name
Query Elapsed Time



Entry point table for this component. Table contains various information about query results made against the warehouse



Query Statistics Detail

Query ID - is a sequential identifier assigned by the DBMS or the data monitoring tool to uniquely identify a query request

Query Start and **End Time** - capture the date/time of when the query was initiated and completed

Number Of Rows Returned - indicates just that the count of rows successfully gathered by the query

Size of the Result Set - contains the size in kilobytes of the row returned in the query. This can be useful in diagnosing report problems with query reporting tools

User Id - contains the operating system identifier of the individual or group who requested the query. This column can be useful for identifying business needs from users who have frequent requests or long running queries

Server and **Database Name** - identify which version of the warehouse database is being queried such as development, QA, production or training

Query Elapsed Time – clock time that it took the query to run

Query Statistics
<u>Query Id</u>
Query Start Time
Query End Time
Number of Rows Returned
Size of Result Set (Kb)
User Id
Database Name
Server Name
Query Elapsed Time





Query Statistics Report Example

Query Statistics Report						January 17, 2000	
Month	User	Number of Queries	Average Time Per Query	Maximum Time of a Query	Average Number of Rows Returned Per Query	Maximum Number of Rows Returned by a Query	Tables Accessed
March	John Smith	30	0:03:32	0:23:17	304,280	1,136,639	Customer, Product, Employee, Time, Sales
	Mary Brown	14	0:10:51	0:14:21	1,676,368	8,277,267	Customer, Product, Employee, Time, Sales
	James Jones	58	0:02:06	0:03:33	275,237	571,277	Customer, Product, Time, Sales
	Jane McDoyle	9	0:13:15	0:14:15	15,629,295	15,629,295	Customer, Product, Employee, Time, Sales, Receivables
February	John Smith	26	0:03:01	0:19:54	260,069	971,486	Customer, Product, Employee, Time, Sales
	Mary Brown	12	0:09:16	0:12:16	1,432,794	7,074,587	Customer, Product, Employee, Time, Sales
	James Jones	50	0:01:48	0:03:02	235,246	488,271	Customer, Product, Time, Sales
	Jane McDoyle	8	0:11:19	0:12:10	13,358,372	13,358,372	Customer, Product, Employee, Time, Sales, Receivables
January	John Smith	23	02:43.0	17:56.0	234,296	875,213	Customer, Product, Employee, Time, Sales
	Mary Brown	11	08:21.0	11:03.0	1,290,805	6,373,502	Customer, Product, Employee, Time, Sales
	James Jones	45	01:37.0	02:44.0	211,933	439,884	Customer, Product, Time, Sales
	Jane McDoyle	7	10:12.0	10:58.0	12,034,569	12,034,569	Customer, Product, Employee, Time, Sales, Receivables





Real-World Meta Data Cases



Scoring

Value

- 1 = You've just earned a promotion
- 2 = Highly valuable
- 3 = Moderate value
- 4 = Low to moderate value
- 5 = Little to no value

Development Effort

- 1 = Trivial
- 2 = Fairly small
- 3 = Moderate
- 4 = Fairly large and complex
- 5 = Parting the Red Sea



Large Financial Institution

One of the largest financial companies in the U.S. and one of the largest financial institutions in the world

Company is highly respected by CEOs

International company

Meta data repository was the enabling technology of a **Data Assets Management** team

CIO was directly involved





Large Financial Institution

Problem/Opportunity

Lower IT costs

Enable data integration

Ensure the success of the data warehousing projects

Meta Data Solution

Manage “code” data

Manage XML schemas

Store and manage ***business meta data***

Manage technical meta data support applications

Enable overseas system support

Entity/attribute naming standards for all new applications





Large Financial Institution

ROI

Business users have grown to depend on the business meta data

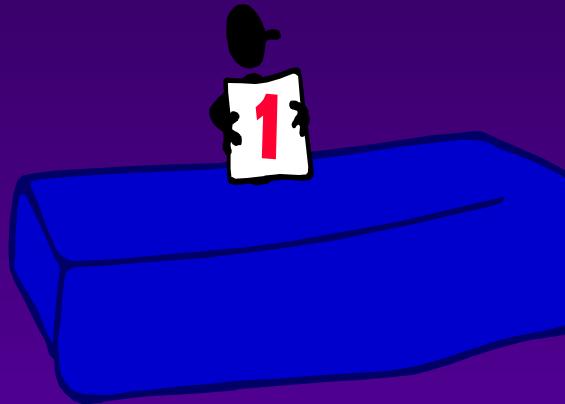
The IT project teams are reporting considerable cost savings

Needed to provide to ***state regulators*** a data trail for information that was provided to policyholders. The repository stored the mappings for a data trail from legacy systems through to the fields on letters that were sent to each policyholder

This satisfied the requirements set forth by the ***public auditors***

Large Financial Institution

Value =



Development Effort =



Information Provider

Large corporation that provides data to banks and other financial institutions

Data is delivered in a data mart-like structure

Competes in a marketplace against other large information providers



LONDON 2002 - 100



www.EWSolutions.com

Enterprise Warehousing Solutions, Inc. ©
2002 All Rights Reserved



Information Provider

Problem/Opportunity

Clients required a great deal of training to understand the data that they have

Meta Data Solution

Integrate business meta data onto the data structures



Information Provider

ROI

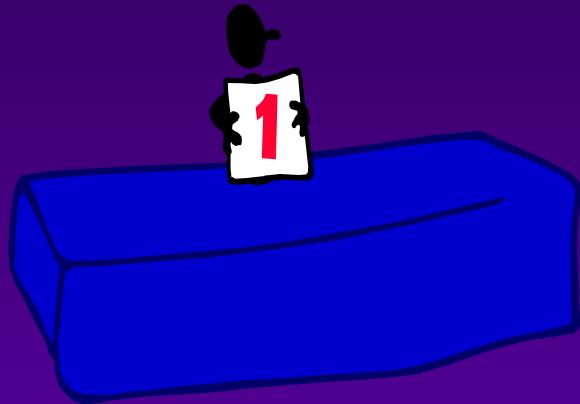
Competition does not have business meta data

Sales team now has a **competitive differentiator** in the marketplace

Charge a premium for the business meta data, so their repository is directly ***generating revenues***

Information Provider

Value =



Development Effort =



Large Bank

Very large bank

Over 55,000 employees serve over 10 million customers
through over 1,600 branches and offices in 35 countries
“Gold Standard” in meta data management

Large Bank

Problem/Opportunity

IT costs were too high (data and application redundancy)

Speed up IT development life-cycles

Meta Data Solution

Integrate technical meta data from across the enterprise

Identification, publication and maintenance of over 4900 batch interfaces between more than 200 applications

Create **enterprise-wide impact analysis**

Managing XML schemas

Support a large data warehouse environment consisting of many databases and many terabytes of data

Entity/attribute naming standards for all new applications



Large Bank

ROI

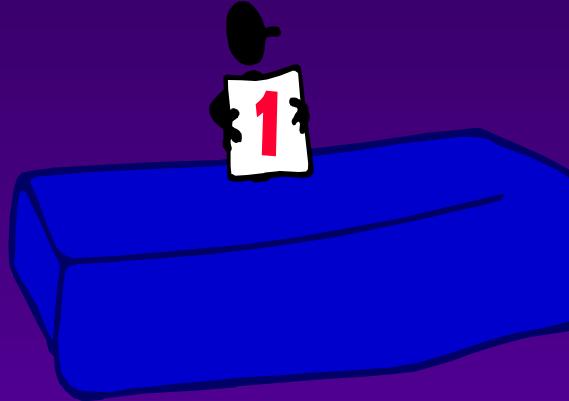
Estimated 28,000 pieces of information were extracted from the meta data repository via the web site. The estimated cost to obtain this information manually (one person talking to another) is over **\$2,500,000**. The actual cost of obtaining this information through the repository is around **\$85,000**.

Impact analysis across 200 applications and can be completed in **hours or days** as opposed to **weeks or months**

Measured that manual impact analysis accounts for 60% to 70% of their application's support and maintenance budget

Large Bank

Value =



Development Effort =



Real-World Meta Data Uses

Don't Limit Yourself



www.EWSolutions.com

Enterprise Warehousing Solutions, Inc. ©
2002 All Rights Reserved

Questions

