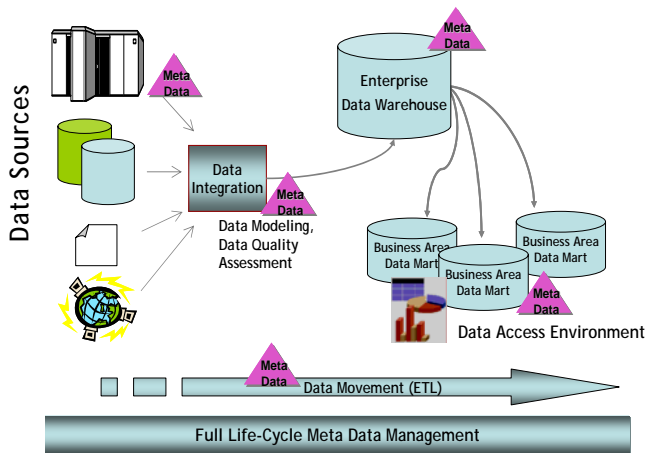


The Importance of Re-use

In order to maximize the return on investment of your data assets, you must be able to effectively re-use its meta-data. That is, meta-data should be defined once, and used as many other places as possible. This will ensure the integrity of meta-data from Sources to Data Warehouse targets. It will increase productivity of developers and business decision-makers. ETL developers will be able to easily get published, standard meta-data definitions to construct their data movement jobs. Data will be defined and consistent across the environment, providing a standard, (or at least an understandable), version of the corporate “truth”, from varying perspectives.

A generic high-level diagram of the data warehouse environment might resemble this:

A Data Warehouse Architecture



The basic process is to separate transaction data systems from the data access environment, bringing the data from multiple disparate transaction sources into an organized, integrated whole and distributing data and its meta-data to the business decision-makers.

Using Tools for Data Warehousing

Many different kinds of tools are used throughout the process. They include, but are not limited to:

- Data modeling tools
- Data extraction, transformation, and loading
- Data Quality Assessment tools
- Database Management Systems
- Gateway products
- Business Intelligence and reporting tools
- Web sites and portals

Meta-data occurs throughout the entire process, and with the use of each kind of tool. An effective meta-data management environment captures the meta-data from all points in the process, organizes it for analysis, and makes it available to all the processes and people that need it.

Defining a meta-data definition once, in the appropriate “tool of record” combined with a disciplined approach to development can alleviate many of the stumbling blocks on the road to building the Data Warehouse environment.

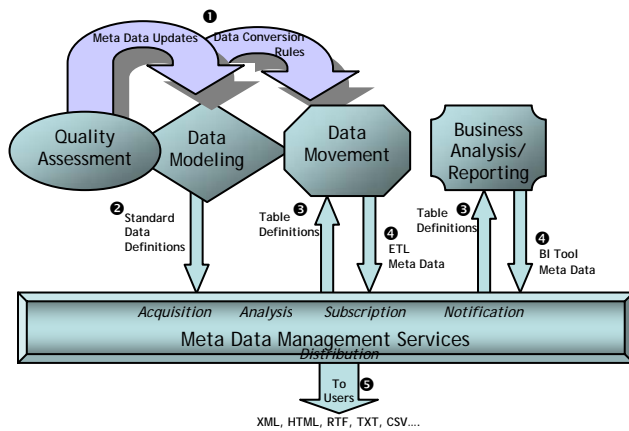
You should choose a meta-data management tool that provides you with the services to support your effort. The meta-data solution should allow you to collect, analyze, organize, and share meta-data in your environment.

Taming Meta-Data Chaos

The diagram on the next page depicts a way to use best of breed tools together effectively.

- Use Data Quality Assessment tools to examine the data itself, exposing data and meta-data anomalies before they hit the ETL Environment.





- Model your Standard data definitions for the Data Warehouse and push them to your ETL and BI tools. Standard data definitions can be reverse-engineered in data modeling tools from many physical data structures to document a set of standard table definitions for source systems.
- Collect meta-data from the ETL process as they are completed. ETL processes introduce new meta-data about the detailed data transformations that occur between the source and target systems.
- BI tools introduce their own meta-data to the process. They may introduce the use of familiar business terms, and may further aggregate and summarize the data from the Data Warehouse for a particular set of business decision-makers.
- Lastly, but most importantly, the meta-data must be made available to its owners and constituents in a variety of easy-to-access formats.

In Control When Change Occurs

By instituting processes that adopt this kind of meta-data flow among tools, you create an integrated set of meta-data across the Data Warehouse development process. The benefits of that integrated meta-data are illustrated dramatically as change occurs over time. Impact Analysis and

change notification are the powerful meta-data management tools that keep current information flowing in the business.

Y2K remediation was a particularly dramatic change phenomenon because it affected every information system in the world, at the same time. Yet, there are changes as dramatic to individual businesses that occur all the time:

- We're running out of Order Numbers – the 5 character field will need to expand...
- We just bought a competitor – their product IDs don't conform to our standards...
- There's a new industry regulation – airport ID codes will expand to 4 characters and flight numbers to five...
- The vendor just changed the schema of one of our Warehouse source systems...
- We're implementing SAP...!

To illustrate how Impact Analysis would work in one of these scenarios, let's describe one of the bulleted examples in more detail – a schema change to one of the packaged source applications.

A particular bank has a world-wide operation and a Data Warehouse environment that integrates source systems from around the globe. It distributes data to a network of Warehouses and Marts so that business analysts can access the data themselves with business intelligence tools, like Business Objects and Cognos PowerPlay.

Once or twice a year, the source system vendors each change the schema of their applications to incorporate maintenance requirements and enhancements. The bank must incorporate these changes throughout the world-wide operation so that analytical systems continue to work, using the most up-to-date version of the meta-data.



Impact analysis allows the bank to ascertain the following about its operation:

This particular column on this Table in the application schema, is used on these 87 specific links in these 60 ETL transformation jobs, in the following 4 ETL Projects, related to these 10 attributes in these 4 Warehouse data models which reside in ERwin and PowerDesigner, and is used in these 17 Business Objects Universes, and in these 20 Cognos Catalogues.

This information can be published to the data modelers, ETL job designers, and BI tool administrators via an internal web site, and used by the overall project managers to build a plan with realistic time estimates. The change gets completely and efficiently implemented throughout the environment. The systems are cut over to use the new source meta-data and Warehouse processing doesn't skip a beat. There are no angry business analysts calling "the morning after" because their systems are blowing up.

This is the payback from a well-managed meta-data environment. This scenario is based on a true example from a UK bank with world-wide operations. When the project sponsor saw the impact analysis diagram that highlighted each point in the numerous ETL processes where the affected column was used, he replied, "This product just paid for itself!"

In summary, meta-data is important. Manage it with an effective meta-data management strategy (and tool capabilities) and you will optimally manage your Data Warehouse environment.

About Metaview³⁶⁰

Metaview³⁶⁰ consults in the areas of data integration, data governance, data warehousing, meta data management, and information systems architecture. Metaview³⁶⁰ delivers solutions that positively affect your bottom line. Please visit us at www.metaview360.com

©2004 Metaview³⁶⁰, Inc. All Rights Reserved. Metaview³⁶⁰ is a trademark of Metaview³⁶⁰, Inc.

