Publication: Data Management Review

Column Title: Data Warehouse Delivery

Author: Douglas Hackney

Headline: Incremental Data Mart Enterprise Architecture

Issue: May 1997

[begin copy]

To avoid the curse of being tasked with integrating a series of non-integrated LegaMarts, incremental data marts *must* be constructed under an enterprise architecture. An enterprise architecture is designed by the data mart team prior to commencement of the design and construction of the incremental data mart(s). The architecture effort should take as little as possible of the team's time, from an afternoon to a couple of weeks. The enterprise architecture is a high level effort to define the critical factors at whatever level of detail that can be ascertained in the available time frame. Not all answers will be known at the end of this period, but you will know enough to get started, and the rest of the details will be filled in as you progress in your incremental data mart project, and as other incremental data marts follow.

Don't get intimidated by the sound of "enterprise architecture." The only thing "enterprise" about it is the potential to extend it across the enterprise. In the beginning it will only address the specifics of your first incremental data mart. The point of the exercise is to get you to look up from the details of your individual project and take some time to consider the best sources and common elements across all potential areas of the business.

The enterprise architecture must identify:

- **Common sources**

The architecture must identify the best system to obtain each required dimension key and metric used in the anticipated enterprise incremental data mart environment. This involves determining the best source system in the business to obtain "Customer ID" information, for example. Once the source system is identified, the first incremental data mart team will build an extract and transformation process to populate their incremental data mart. This extraction process needs to be stored in a central repositories where other teams can utilize it when they need the same information.

Many times, the different elements of the organization targeted to be served by incremental data marts cannot agree on a common source system. In this case, you must integrate these two sources into a common source for the various incremental data marts in the organization. This usually requires a point solution tool dedicated to the scrubbing, integration and de-duplication mandatory in these circumstances.

- **Common dimensions**

Every business looks at its activities in several ways common to most user groups in the organization. Popular examples include Customer, Product, Time, Geography, Sales Geography, and Employee. A primary goal of the enterprise architecture is to identify common business dimensions that are shared across multiple targeted user groups for multiple incremental data marts in the organization.

- **Common business rules**

Business rules are the algorithms and logic the business uses to calculate metrics and derive classification and structure. It is very common for various business units to use a variety of ways

to calculate standard business metrics such as “sales” or “net profit”. It is your job during the formation of the enterprise architecture to identify as many common business rules as possible. It is also a necessity to identify all the different ways the business uses to calculate the same metrics, classifications and structure. If you are unable to gain consensus on these business rules, you will need to carry individual columns identifying each method of calculation.

- **Common semantics**

Semantic terms are what the business uses to label itself, its elements, its structure, its metrics and its activities. As elsewhere in life, semantic differences between different elements of the organization lead to many challenges. When two different groups use the same term for two different entities, or use two different terms for the same entity, confusion and frustration reign supreme. It is always a tough battle in trying to gain enterprise wide consensus for semantic terms, but one well worth fighting.

- **Common metrics**

Organizations have a variety of ways to measure the operations of the business that are referred to as metrics. Common metrics include units, dollars, hours and other measures of output, throughput or productivity. Metrics are fundamentally what the business is about and how it measures itself.

The creation of the enterprise architecture is an absolute prerequisite for success and its absence will guarantee long term failure of your team’s efforts. It is most tempting when faced with the need for only a single incremental data mart solution to avoid or defer this step. You must not fall prey to this temptation. Your initial single solution will soon lead to a proliferation of “single solutions” that will pop up across the landscape of your business like mushrooms after a spring rain. Without a common enterprise architecture to lay the ground work for common sources, dimensions, metrics, semantics and business rules, you will be faced with multiple versions of the truth, mass confusion and frustration about semantic differences and a looming nightmare of LegaMart integration.

Excerpted from “Understanding and Implementing Successful Data Marts”, a forthcoming book by Douglas Hackney from Addison Wesley Longman Publishing. Further information is available at www.egltd.com.

[end copy]