Design Tip #135 Conformed Dimensions as the Foundation for Agile Data Warehousing

By Margy
Ross

Some clients and students lament that while they want to deliver and share consistently-defined master conformed dimensions in their data warehouse/business intelligence (DW/BI) environments, it's "just not feasible." They explain that they would if they could, but with senior management focused on using agile development techniques to deliver DW/BI solutions, it's "impossible" to take the time to get organizational agreement on conformed dimensions. I want to turn this argument upside down by challenging that conformed dimensions enable agile DW/BI development, along with agile decision making.

Whether a Kimball specialist or not, many of you are conceptually familiar with conformed dimensions. A conformed dimension is descriptive master reference data that's referenced in multiple dimensional models. Conformed dimensions are a fundamental element of the Kimball approach. Conformed dimensions allow DW/BI users to consistently slice-and-dice performance metrics from multiple business process data sources. Conformed dimensions allow data from different sources to be integrated based on common, unified dimension attributes. Finally, conformed dimensions allow a dimension table to be built and maintained once rather than recreating slightly different versions during each development cycle.

Reusing conformed dimensions across projects is where you get the leverage for more agile DW/BI development. As you flesh out your portfolio of master conformed dimensions, the development crank starts turning faster and faster. The time-to-market for a new business process data source shrinks as developers reuse existing conformed dimensions. Ultimately, new ETL development will focus almost exclusively on delivering more fact tables as the associated dimension tables are already sitting on the shelf ready to go.

Defining a conformed dimension requires organizational consensus and commitment to data stewardship. You don't need to get everyone to agree on every attribute in every dimension table. At a minimum, you should identify a subset of attributes that have significance across the enterprise. These commonly-referenced descriptive characteristics become the starter set of conformed attributes, enabling drill across integration.

Over time, you can expand from this minimalistic starting point iteratively. Any special dimensional attributes unique to a single business process do not have to be discarded during the conforming process, so there are very few serious compromises. And finally, if your organization has already implemented a master data management (MDM) capability to support the transactional source systems, the work to create conformed dimensions for the enterprise data warehouse just got much easier.

If you fail to focus on conformed dimensions because you're under pressure to deliver something yesterday, the various departmental analytic data silos will likely have inconsistent categorizations and labels. Even more troubling, data sets may look like they can be compared and integrated due to similar labels, but the underlying business rules may be slightly different. Business users waste inordinate amounts of time trying to reconcile and resolve these data inconsistencies which negatively impacts their ability to be agile decision makers.

The senior IT managers who are demanding agile systems development practices should be exerting even greater

organizational pressure, in conjunction with their peers in the business, on the development of consistent conformed dimensions if they're interested in both long-term development efficiencies and long-term decision making effectiveness across the enterprise.