

Kimball Design Tip #7: Getting Your Data Warehouse Back On Track

By Margy Ross and Bob Becker

During the past year, we've repeatedly observed a pattern with maturing data warehouses. Despite significant effort and investment, some data warehouses have fallen off course. Project teams (or their user communities) are dissatisfied with the warehouse deliverables - the data's too confusing, it's not consistent, queries are too slow, etc. Teams have devoured the data warehousing best sellers and periodicals, but are still unsure how to right the situation (short of jumping ship and finding new employment).

If this situation sounds familiar, take the following self-check test to determine if the four leading culprits are undermining your data warehouse. Consider each question carefully to honestly critique your warehouse situation. In terms of corrective action, we recommend tackling these fundamental concerns in sequential order, if possible.

1. Have you proactively gathered requirements for each iteration of the data warehouse from business users and aligned the data warehouse implementation with their top priorities?

This is the most prevalent problem for aging data warehouses. Somewhere along the line, perhaps while overly focused on data or technology, the project lost sight of the real goal to serve the information needs of business users.

As a project team, you must always focus on the users' gain. If the team's activities don't provide benefit to the business users, the data warehouse will continue to drift. If you're not actively engaged in implementing solutions to support users' key business requirements and priorities, why not? Revisit your plans to determine and then focus on delivering to the users' most critical needs.

2. Have you developed a Data Warehouse Bus Matrix?

The Matrix is one of the data warehouse team's most powerful tools. Use it to clarify your thinking, communicate critical points of conformance, establish the overall data roadmap, and assess your current progress against the long-term plan. If you're unfamiliar with the Matrix, refer to Ralph's 12/7/99 Intelligent Enterprise article at <http://www.intelligententerprise.com/990712/webhouse.shtml>.

3. Is management committed to using standardized conformed dimensions?

Conformed dimensions are absolutely critical to the long-term viability of a data warehouse. We find many warehouse teams are reluctant to take on the socio-political challenges of defining conformed dimensions. In all honesty, it's extremely difficult for a data warehouse team to establish and develop conformed dimensions on its own. Yet the team can't ignore the issue and hope it will resolve itself. You'll need management support for conformed dimensions to help navigate the organizational difficulties inherent in the effort.

4. Have you provided atomic data in dimensional models to users?

Data shortcomings are often at the root of data warehouse course adjustments - it's either the wrong data, inappropriately structured or prematurely summarized. Focusing on business requirements will help determine the right data; then the key is to deliver the most atomic data dimensionally. Unfortunately, it's tough to gracefully migrate from data chaos to this nirvana. In most cases, it's best to bite the bullet and redeploy. Teams sometimes resort to the seemingly less drastic approach of sourcing from the current quagmire, however, the costs are inevitably higher in the long run. Often the granularity of the existing data precludes this alternative due to premature summarization.

In summary, if your data warehouse has fall off course, it won't magically right itself. You'll need to revisit the basic tenets of data warehousing. Listen to users to determine your target destination, get a map, establish a route, and then follow the rules of the road to get your data warehouse back on track.