

Hills, peaks and valleys

He who takes the wrong road makes the journey twice. —Proverb



Rob Armstrong
Director of data warehouse
support at Teradata

It has often been said that data warehousing is a journey. When you're on that journey, you will pass through Report-ville, Cube-city, Analytics-burg and Predictive-place and head to Triggered-town on your way to the bustling metropolis of Active Data Warehouse.

Like any journey, when developing an active data warehouse, it helps to chart your path; you'll want a general idea of the route you are going to take through the hills of data consolidation to activation.

Data mart consolidation

“We're all entitled to our own opinions. But none of us can afford to be wrong in our facts.”

—Mort Crim

Many companies start their journeys with data spread across the organization, redundant or inconsistent data—even data that is unaccounted for. Consolidation would allow users to access it with greater ease.

Of course, the real questions are which data should be consolidated and what value does consolidation bring to the corporation.

For the consolidation effort to be successful, it must be based on business

needs. You must first determine which data should be brought together, then find the current locations of that data. Ultimately, you will understand the extent of your data redundancy or inconsistency and begin the process needed to ensure data quality.

Data integration

“A man with one watch knows what time it is. A man with two watches is never sure.” —Segal's Law

Once the data is consolidated the next step, which must be committed to from the start, is integration of the data model. Integration will provide analytical consistency, ease of extract, transform and load (ETL) processes, and time-to-market reduction. The foundation is now laid for sustainable data warehousing—including going active.

Just as you do not automatically jump to a destination, you cannot achieve data integration in a single bound. It is a cyclical journey, driven by business need, data availability and return on investment (ROI).

Data expansion

“The road to success is always under construction.” —Lily Tomlin

The real question of what data comes next is based on two key factors: cost to acquire the new data elements and their ensuing benefits.

The results of this analysis can help you prioritize the next subject area and phase.

Data acquisition

“It is not so much what you know anymore that counts, it is how fast you learn.” —Robert Kiyosaki

When you look at what new data can be of value, one dimension to consider is the timeliness of that data. Can the benefit increase as the data acquisition becomes more real time?

Timeliness has two aspects: frequency and granularity. The data's business usage and value justify how frequently the same level of data is collected and at what scope of granularity.

Welcome to activation!

Now for a quote of my own:

“If you are not taking action, then stop making decisions.”

The ultimate outcome of data warehousing is that actions are triggered by the data itself without requiring constant intervention. It is only the exceptions that need attention, and the data will alert us to those situations.

As you understand what data is meaningful to the processes and the timeliness those processes require, only then can you move forward.

The important message is that data-driven actions will also drive new data points. And remember, this is an ongoing adventure.



For an expanded version of Rob's insights, visit Fresh Perspectives on TeradataMagazine.com.