

Data Management Review

Column Title: Data Warehouse Delivery

Author: Douglas Hackney

Headline: Meta Data Maturity

Issue: March '96

Submitted via: CompuServe

[start copy]

Have you ever seen the movie *The Christmas Story*? It's a classic tale about a little boy named Ralphie who desperately wants a Red Rider BB Gun for Christmas, and his associated adventures, triumphs, and tragedies associated with that fateful holiday. One of the most memorable scenes involves Ralphie's faithful vigil in front of the radio, as Little Orphan Annie finally reveals the end of the secret message he has been collecting. He runs to the bathroom, where he can use his newly acquired secret decoder in private. Surely the message must be of the highest importance and secrecy. After all, it came from the highest of authorities, and she had made it clear that his life would not be complete without its contents. After a few minutes of feverish decoding Ralphie is crushed to learn that the message is 'Drink Your Ovaltine'. His top secret, mission critical message was nothing more than marketing hype. Ralphie leaves the bathroom a saddened, but more mature little boy. In one of his first steps to adulthood his blind faith in his heroes was replaced by the creeping cynicism of the adult consumer.

Life with meta data is much the same: marketing hype hiding some bitter truths. Every vendor in the world is out to sell us on the message that their tools are the way, the truth, and the light to lead us to the promised land of automated meta data creation, maintenance, and access. Unfortunately, our current reality mostly falls short of this utopian vision.

Here are some hard, sobering, and sometimes bitter truths:

- You must have meta data to survive. No data warehouse will exist for long without it. It's as mission critical as it gets when it comes to an effective, robust, and productive data warehouse system for your enterprise.
- You must create and maintain meta data across the entire process flow of your data warehouse system. This means from extract to end user, you must populate and maintain relevant and accurate information about the system, the elements, the sources, the schedules, the algorithms, the business rules, the data structures, the summaries, the detail, the OLAP data sets, the available information resources, the end user access systems, etc.
- The meta data must be accessible. It must be a point and click away. I'm sure your experience has echoed my own, printed documentation is useful primarily for holding down miles of shelves around the world. Your query, reporting, OLAP and other data access tools must provide inherent point and click capability to read your live meta data tables. If they don't, find something that does. Users have two primary seek paths for answers to meta data questions: 1) point and click or 2) call support. With certain types of users, the order is reversed.
- The meta data must be data context sensitive. If a user is looking at objects, then object level information must be a point and click away, if looking at tables, then table level, if field level, then field level. Open ended interfaces that start with an overwhelming view of the entire meta data collection are of marginal value.
- The meta data must be user context sensitive. If the user is a member of the development and support team, then technical meta is the default. If the user is a User user, then end user oriented information is pertinent. End users don't care about what the batch load program name is, or who wrote it, but it is vitally important to the team member trying to discover why it didn't execute last night.

OK, get out your secret decoder rings, it's time to reveal the most painful truth behind the hype:

- Almost all meta data is manual. Regardless of how well the tool vendor's talented and effective marketing staffs dazzle us with tales of mystical automatic meta data, the truth is that almost all valuable meta data is manually created and maintained. Take a look under the covers of the tools and you'll discover that when you're doing mapping and transformation, only the bare minimum meta data fields are required, such as source field, target field, algorithm. All the good stuff, the fields that you'll need to make this system work and be valuable, are optional.

"So what's the big deal?" you ask. Ever lived through a data warehouse implementation? The project is inevitably underestimated, so it's inevitably behind schedule. The pressure is on the management team, who pass it on to the development teams. The first thing to slip off the plate is meta data. After all, "we can always go back and do that", right? About the time you're into the second or third subject area and everyone is starting to wonder why the utilization rate is so low and support costs are so high, you'll be revisiting that thought.

At all costs, avoid starting down this seductive, slippery slope of meta data deferral. Most likely, you'll end up at the bottom of the gorge, being feasted upon by the angry gods of consequences.

What to do?

- Budget resources for meta data creation. Make meta data population a metric for the development teams, regardless of if you code by hand, or use a tool.
- Budget resources for meta data maintenance. As we all know by now, the only constant for a data warehouse is change. Any change requires changes in the meta data. Neglect it at your peril.
- Ensure ease of access. Find and deploy tools that seamlessly integrate meta data.

So now you're all ready to face tomorrow. Saddened no doubt, but older and wiser in the ways of meta data. Next month we'll look into some types of meta data that you'll find most valuable for your team and your users.

[end copy]