

Publication: Data Management Review

Column Title: Data Warehouse Delivery

Author: Douglas Hackney

Headline: The March of the Acronyms

Issue: March 1997

Submitted via: e-mail & fax

[begin copy]

We learn early on in this business that technology stops for no one, least of all for those of running as fast as we can to keep up with the breaking news every day in our little corner of the technological world. Sometimes it feels like you're running in place, as the trade press seems to break word of a new technological revolution every couple of days, even while we're trying to understand the potential implications of last month's hot news.

Pity the poor publisher trying to maintain a valid and timely reference book. Just when they get the Java section updated with all the beans, aromas, scripts and grinds, along comes ActiveX, Marimba, and all the other new paradigms flooding out of Silicon Valley, Austin and other hot beds of technological innovation. Even in our little world of data warehousing and data marts, it is a struggle to keep up sometimes.

In the last year we've seen a rapid revolution in the core tools and technologies that are central to our mission of extracting, integrating, scrubbing, transforming, loading, monitoring, accessing and managing information in the enterprise. Along with this evolution has come the inexorable march of acronyms. They keep multiplying, like an army of red ants on the African plain. They move forward, ever forward, relentless, unstop-able, and just like the long lines of ants, inevitably become longer and longer and longer.

It used to be a simple world in data warehousing. We were pretty much focused on a couple of market segments: extract, mapping and transformation; and end user access, since they held the bulk of the tools that were available for direct application in our world. The acronyms were pretty simple too, EMT (Extract, Mapping and Transformation) on the source side and Q&R (Query and Reporting) and EIS (Executive Information Systems) on the access side. Like the soft focus fifties, those uncomplicated times are longed for today by overloaded data warehouse teams resorting to crib sheets to keep the daily onslaught of new acronyms clear. While invaluable in hard fought family scrabble games, the rest of the time the jumble of new letter combinations threatens to overload our overtaxed memories.

On the source side, we've seen the simple 1st generation code generator EMT tools supplanted by the 2nd generation transformation engine powered DTEM (Design, Transformation, Extraction and Management) tools, led by the initial player in this space, Informatica's PowerMart. The DTEM suites, with their server centric design, integrated tools, consistent UI, common, open and extensible meta data repository, and vastly higher throughputs, soon took center stage. Of course, everyone in the technology world noticed that moving from a three letter acronym to a four letter acronym was worth not only major market share, but the chance to establish a new paradigm in the process. There being no higher goal in life for those in the technology industry than establishing a new paradigm, it wasn't long before we had a five letter acronym: DTEMM. Monitoring was added to the transformation engine suites, allowing processes to be tracked from extraction to index. This was good. Very good if you were a user, and very, very good if you were responsible for maintaining a data warehouse or data mart environment.

These were exciting times. Now, you not only had a consistent GUI, smoothly integrated, high throughput, fast & flexible, data pumpin' machine, you had gauges to watch too! The addition of integrated, automatic, meta data driven monitoring to the transformation engine world provided a key step along the evolution of the next possibility. Up to now, these tools were excellent tactical solutions, at a very reasonable price. But they were incapable of scaling to an enterprise level.

Some DTEMM tools could not support multiple, simultaneous population, in which they could perform a transformation once and populate multiple targets simultaneously with the output. Some could not share a single, common repository for a data warehouse and/or multiple data marts. And some even used temporary data base tables in their transformation process with the attendant high overhead of disk I/O thrash, instead of performing all operations in memory and limiting disk I/O to caching operations. Others were incapable of sharing resources such as transformation algorithms among multiple instances of the tool. Our frustration was complete. Just when we thought a new DTEMM enabled world was upon us we discovered we were once again held back by fundamental architectural and design flaws. Undaunted by these challenges, and still in pursuit of the holiest of grails: a new paradigm, the technologists forged onward. They, and they alone could deliver us from this unspeakable horror of non-enterprise scaleable solutions.

They struggled. They searched. They stole each other's ideas. They burned through engineers like a fire through dry brush. They consumed venture capital like a pack of kids vacuuming Halloween candy. But in the end, their fast and furious journey led them to the holiest of holies: they had found a new paradigm. Then, and only then could they add another letter.

Thus was born DTEAMM. A new day dawned. The storm clouds of non-scaleable solutions parted and the bright shining new day revealed our future. An integrated environment that encompasses system design, mapping, data extraction, scrubbing, transformation, loading, monitoring, system management, a common repository for enterprise resources (meta data, queries, reports, analyses, transformation algorithms, etc.), scheduling, delivery, and the crowning glory: End User Access. Yes, now users could tap into the same consistent, all encompassing environment as the technical team. They could share the same meta data, the same reports, the same everything. Yes DTEAMM (Design, Transformation, Extract, Access, Management and Monitoring) was finally upon us, and a new paradigm was born.

Server centric, process distribute-able, thin client enabled, infinitely scaleable, seamlessly integrated, consistently GUI'd, meta data driven and designed and built exclusively for you and me to solve our decision support challenges at price points a fraction of 1st generation EMT tools. Life was good. Gone were the days of millions of lost person years spent desperately patching together tools from disparate vendors. Lost in the dustbin of history were closed, proprietary meta data repositories. DTEAMM brought forth a world of tight, total and seamless integration of all functions, processes and tools in the environment of decision support systems.

Hey, maybe there really is something to this new paradigm thing after all.

And here, in this most warm, dry and comfortable place, the march of the Acronyms has stopped. For now.

[end copy]