META Delta

**Delta 2744** **2 February 2004**

# Data Management Vendors: The Ones to Watch in 2004

**Enterprise Analytics Strategies, Application Delivery Strategies**

Doug Laney

**During 2H03, a resurgence took place in data warehouse and data quality initiatives. The mainstream drive toward pan-enterprise analytic solutions is considerable, and vendors in this market have begun to respond.**

Enterprise analytic solutions are more than just departmental analytic solutions on steroids. They require quantum leaps in scalability, robustness, usability, compatibility, and manageability — and at least a modicum of additional functionality. Enterprise-class solutions must scale to support a vast volume/velocity/variety of data and number of users, demonstrate mission-critical robustness, and accommodate a spectrum of user classes (e.g., technical users to business users, casual users to power users). They must be compatible with an array of platform technology (e.g., DBMS, servers), standards (e.g., Web services), languages/geographies, and other complementary technologies through messaging protocols and metadata interchange. Also, enterprise-upgraded solutions must include centrally managed/locally developed repositories (i.e., for business rules) and more sophisticated administration features (e.g., scheduling, job control, monitoring). In addition, analytic solution providers will become competitive fodder without some degree of core functional enhancement on top of their "newly enterprised" solutions.

Indeed, since major providers in the data integration, business intelligence, and data quality markets intend to engage one another at the high end through 2004/05, many upstart vendors aptly sense niche or midtier opportunities. Low-end analytics is still the stuff of Microsoft Excel/Access "implementations" and hand-rolled data integration/quality "applications," but the vast midtier and specialty markets will continue to be largely underserved into 2006/07. Through 2005/06, the more vocational midtier IT brethren will remain a bit confounded by how to architect an analytic technology stack. Regardless, as businesses and government organizations tread the 2004 analytics landscape, they should consider the following oases of enterprise-class or otherwise notable solutions in key submarkets.

**Data Integration.** Ascential and Informatica will continue to dominate the extract/transform/load (ETL) market, though SAS makes a strong showing internationally in highly heterogeneous data environments and where annuity pricing is not an issue. Ascential's product architecture may be more apropos in mainframe-laden environments, but Informatica's Striva acquisition lends it increased cache on "big iron." Still, savvy enterprises realize the price/performance benefits of ETL products that leverage DBMS engines. Business Objects and Hummingbird are notable here, and both sport strong new user interfaces.

***META Trend: In 2003/04, as data warehouse initiatives continue to unify, enterprises will gravitate toward data integration solutions with the broadest array of data source adapters (e.g., for databases, commercial business applications, message queues, Web services, information service providers). Fueled by a need for infrastructure/development/operations cost savings and real-time access to heterogeneous data sources, the market for data federating technology (e.g., enterprise information integration [EII]) will approach the demand for batch-oriented extract/transform/load (ETL) by 2005/06. Through 2007/08, all classes of data integration solutions (e.g., ETL, EII, EAI) will evolve to accommodate unstructured content sources.***

Others to watch include DataMirror and Golden Gate — those that did real time before real time became fashionable; the cloak-and-dagger (sans dagger) Ab Initio, the data integration vendor to the rich and famous that excels in high-volume, high-performance, high-complexity environments; and providers such as Data Junction (Pervasive) and Sunopsis that span the ETL/EAI continuum.

The re-emerged and now recontracted enterprise information integration market offers technologies for


Copyright © 2004 META Group, Inc. All rights reserved.

federating (or virtually integrating) data across heterogeneous environments. MetaMatrix and Avaki are the standout standalone vendors, flanked by giants IBM and BEA.

**Data Quality.** Ultimately recognizing that data is a corporate asset and fuel for business performance, enterprises made data quality one of the hottest topics in 2003. We expect this heat to continue through 2005/06. Product road maps show Firstlogic and Group 1 poised to make the greatest near-term leaps at the high end of the data quality market. Through 2004/05, Trillium will remain a major force, particularly internationally; DataFlux (SAS) will make the largest market share leap; and Innovative Systems will get strong play in the ancillary customer data integration and USA PATRIOT Act compliance markets. Other vendors to watch include Similarity Systems, which has an intent focus on data quality problems beyond contact data, and Language Analysis Systems, which can generate alternative name spellings to help catch “bad guys” (via customer databases). We still expect syndicated data providers to show extremely well through 2005/06. Currently, they seem to be reeling from privacy mandates cramping their style, but “infovendors” such as Acxiom, Chordiant, Experian, First Data, and infoUSA should be key suppliers of supplemented customer information in most leading North American enterprises by 2006/07.

**Metadata.** Increasingly related to data quality is the progressively prevalent class of technology known as data profiling. These tools help with the analysis of data to determine and quantify data quality issues and data integration opportunities. For data/metadata administrators, such tools can readily expose issues with current sources of metadata (e.g., data models, data dictionary entries). Most forms of metadata are manually created and infrequently updated or synchronized, so introspecting the actual data periodically can be beneficial to keeping it clean, complete, and readily associated. By 2005/06, data profiling technologies will be a mainstay in the data management technology stack. Vendors to watch include Avellino, which requires profiled data to reside in its repository but has very strong profiling functions; Evoke, which has appropriately retrenched and repriced its pioneering tool; and Firstlogic and DataFlux (SAS), which are the first of the data quality vendors to offer this important complementary offering.

Recognizing that active metadata is the key enabler of adaptive data management, vendors such as DataFlux, Sypherlink, and DB Wizards are demonstrating a nascent ability to generate data integration or data quality business rules by analyzing the data first. In the traditional metadata space, interest in megarepository products and ETL-centric metadata management is waning, but adoption of metadata brokers (e.g., Meta Integration Technology) and federated repository services (e.g., Data Advantage Group) will be on the rise through 2004/05.

**Analytic Infrastructure.** Although Oracle and IBM continue to augment their transaction-oriented DBMS products to accommodate more sophisticated analytic needs, Teradata will maintain its position as the elite analytic platform of choice through 2006/07. In cases where its one-two hardware/DBMS punch is not palatable for reasons of budget or commitment, targeted analytic-specific DBMSs (e.g., from Sand Technology or Sybase) will fit the bill. (However, users should beware of similar solutions that choke on SQL and render them incompatible with ubiquitous query and reporting tools.) The one to watch in the analytic infrastructure space is Netezza, with its change-the-rules approach to extreme analytic performance using proprietary hardware and heavily tweaked open source DBMS.

**Professional Services.** Analytic parity prevails among professional services providers. Pockets of prowess persist, but there are no clear performance leaders. Through 2004/05, vendors to watch include BearingPoint, which (on the planning side at least) has demonstrated a knack for articulating the complexities of data warehouse projects and clearly presenting options and recommendations to its clients; Knightsbridge, the go-to consultant for big-data projects, especially data warehouses; and PricewaterhouseCoopers, whose data management practice (particularly in the area of data quality) leads in methods and technology competence. We suggest keeping an eye on the offshore outsourcing space, but we do not envision this being a viable option for full-scale analytic initiatives until 2006/07.

## *Bottom Line*

**Leading providers in all data management submarkets will persist through 2004, augmenting their solutions to be more broadly deployable within an enterprise. This leaves room for emerging solutions to fill gaps or combine technologies in unique ways that will be viable for some enterprises.**

***Business Impact: At an enterprise level, business performance management requires an appropriately enterprise-quality analytic technology stack.***


Copyright © 2004 META Group, Inc. All rights reserved.