New Paths to Data Integration: Taming Big Data Helps Address Lingering Issues

database
TRENDS AND APPLICATIONS

March 27, 2013

By Joe McKendrick

Data keeps growing, systems and servers keep sprawling, and users keep clamoring for more real-time access. The result of all this frenzy of activity is pressure for faster, more effective data integration that can deliver more expansive views of information, while still maintaining quality and integrity.

Enterprise data and IT managers are responding in a variety of ways, looking to initiatives such as enterprise mashups, automation, virtualization, and cloud to pursue new paths to data integration. In the process, they are moving beyond the traditional means of integration they have relied on for years to pull data together from various sources to present to decision makers—from connectors to extract, transform, and load (ETL) configurations to data warehouses. Future data integration efforts will "look more like mashups that allow us to pull together all relevant internal and external data for a particular decision," says Frank Diana, principal for digital enterprise solutions at Tata Consultancy Services. Open source tools and frameworks are also coming on the scene as "silo busters," Diana adds.

All this new access presents both opportunities and challenges. Having a much broader base of data available for analysis opens up new possibilities for businesses. However, managing much larger datasets to make them available for business intelligence and analytics also introduces increased complexity as well.

Are today's organizations up to the task? Not yet, according to a recent survey of 298 members of the Independent Oracle Users Group (IOUG), conducted by Unisphere Research, and underwritten by Oracle. Fewer than one out of five data managers and executives are confident that their IT infrastructures and toolsets will be capable of handling the surge of big data types—in terms of volume and variety. (Big Data, Big Challenges, Big Opportunities: 2012 IOUG Big Data Strategies Survey, September 2012.) "Gaining a semantic understanding of what is in data so that it can be integrated with existing data stores continues to be incredibly challenging," John Akred, senior manager and principal data scientist in Accenture's Emerging Technology Innovation group, tells DBTA. "We can train a child to recognize that two columns of data are probably related, but most integration tools fail here."