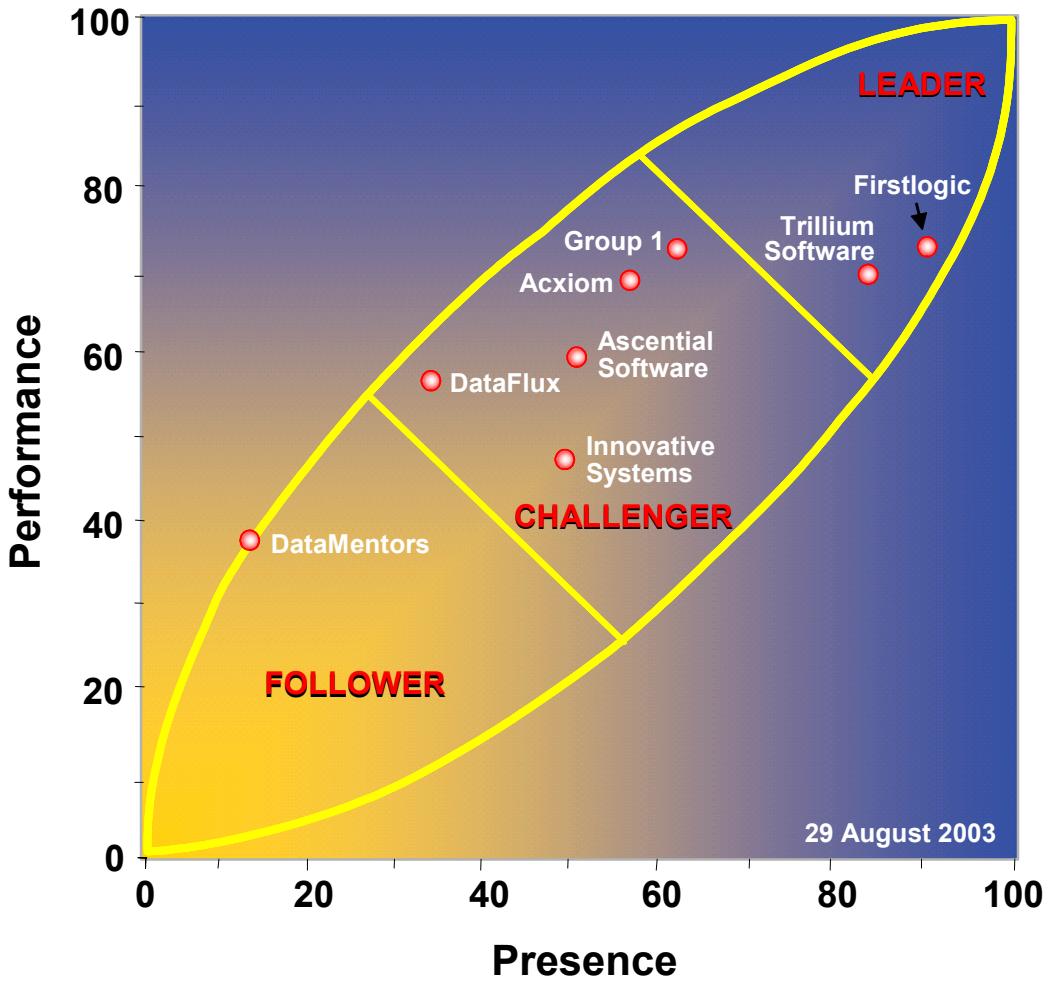
METAspectrum℠
Market Summary

29 August 2003
METAspectrum Evaluation Service, Enterprise Analytics Strategies, Application Delivery Strategies

METAspectrum 36.0
Doug Laney, Kevin Waddle

Data Quality Tools

METAspectrum℠ Evaluation



METAGROUP

Market Definition

The data quality (DQ) software market comprises vendors that offer solutions used in the identification and remediation problems with an enterprise's information assets. Specifically, these vendors offer various technologies for profiling, standardizing, matching, cleansing, and enriching data to help ensure its accuracy, completeness, and integrity.

Market Forecast

Data quality (DQ) adoption rates will expand 20%-30% annually during the next three to five years, with the services component (not addressed in this evaluation) growing at a commensurate pace. Moreover, the market for DQ-related services will assume a more strategic status, particularly among leading enterprises. The significant breadth and depth of business rules used as a basis for data quality offerings are significant barriers to entry, preventing vendors proffering ancillary data management technology (e.g., DBMS vendors, extract/transform/load vendors) from building competitive offerings from scratch. In addition, given the complexity and continuous maturation of these offerings, vendors that choose to acquire DQ vendors will find their offerings lagging without serious commitment to them. A sustained expansion of the DQ market growth will come about in 2004, fueled by an enterprise-level interest in information asset management and the recognition of information as part of the IT portfolio. More tactically, CRM, data warehouse, and customer data integration (CDI) initiatives will continue to lean heavily on DQ solutions. Since several of the leading DQ vendors are already subsidiaries of a parent company, users should not expect much consolidation within the DQ market during the next few years. Instead, niche DQ purveyors will come and go, continuing to chip away at the market by focusing on specific core functionality or introducing related niche capabilities.

Key Findings

With slightly decreasing frequency in 2003/04, DQ buyers will demonstrate a preference for market leaders, as legitimate challengers are able to offer core capabilities at attractive prices. Still, price has not entered the selection equation in this space as much as it has for other types of technologies. Market leaders will continue to be those with the strongest presence-oriented strategies (e.g., educational marketing, technology-level partnering) combined with the broadest array of capabilities. We do not believe there is much room for new market leaders to emerge through 2005/06. Essential criteria for evaluating and selecting data quality vendors include the following:

- **Core Capabilities:** Abilities to standardize, cleanse, and identify/match/deduplicate are still paramount in the eyes of buyers.
- **Data Profiling:** Emerging rapidly, the ability to discover data quality issues is becoming a requisite capability, though several standalone data profiling vendors exist.
- **Data Enrichment:** Data completeness is the number-one data quality issue. The ability to fill in data gaps is critical to most enterprises.
- **Compliance:** US and other governments are cracking down hard on enterprises that do business with identified suspect individuals and organizations. Data quality buyers increasingly have compliance on their checklists.
- **International:** DQ solutions for enterprises engaged in international trade that lack the ability to handle international names and addresses are becoming quickly eliminated from contention.
- **Non-Name/Address:** A total of 90% of DQ functionality is aimed at customer contact information, but the ability to introduce custom or packaged logic for financial, product, or other data is increasingly critical.
- **Installed Base:** References still go a long way toward shortlisting DQ technologies, particularly given the typical effort involved in installing, customizing, integrating, and deploying them.

Leaders

Leaders in this market have stable, mature products with a deep, expanding knowledgebase of data matching and cleansing business rules. They also surround their core technology with an array of capabilities such as data profiling, data enrichment, and postal logic. Leaders are also characterized by their breadth of international name-and-address handling and demonstrated integration with major business applications (e.g., SAP, Siebel, PeopleSoft).

Challengers

Challengers in the DQ market are mainly characterized by their focus on a narrower scope of capabilities than the leaders. They are either limited to North American name-and-address data, are restricted to particular platforms, focus on ancillary capabilities, or have emerging/diminishing presence in the market.

Followers

Followers in the DQ market are those that strictly focus on Windows/NT platforms, offer only a small subset of overall DQ capabilities, target small enterprises or departments, and price their offerings to slip beneath typical DQ solution budget radars.

Bottom Line

The data quality market will increasingly intersect with the data integration solution (e.g., ETL, EAI, EII), compliance solution, business application, and e-commerce market, yet we do not foresee DQ vendors being absorbed into these markets. Rather, DQ solutions will mature independently, improving their openness to integration with solutions in these other markets. Buyers should not become fixated on cleansing performance or accuracy as much as other key capabilities, such as data enrichment, data profiling, compliance, non-name-and-address, and international name-and-address.

Business Impact: Information should be considered part of the IT portfolio and data quality solutions should be seen as a means to increase the value of that portfolio, in addition to reducing business risk and identifying business opportunity.