Gartner

Research

Publication Date: 4 June 2008 **ID Number: G00157464**

Magic Quadrant for Data Quality Tools

Ted Friedman, Andreas Bitterer

Growth, innovation and volatility (via mergers and acquisitions) continue to shape the market for data quality tools. Investment on the part of buyers and vendors is increasing as organizations recognize the value of these tools in master data management and information governance initiatives.

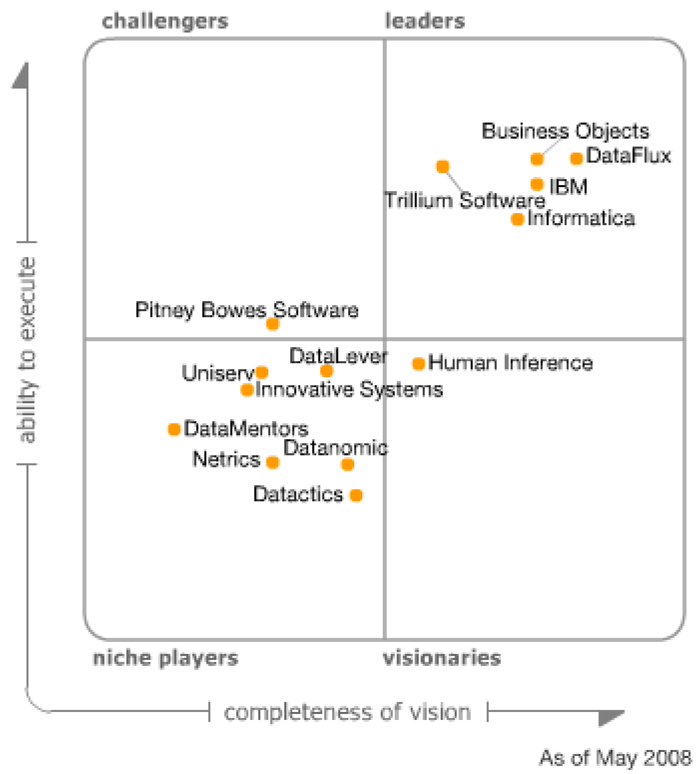
WHAT YOU NEED TO KNOW

The market for data quality tools continues to enjoy significant growth, but experiences ongoing volatility in the form of acquisitions (both direct acquisitions of stand-alone vendors in this market, as well as the acquisition of larger vendors for which this market represents one of many competitive fronts). Most vendors have evolved to full-function data quality tool suites that address a broad range of data quality requirements. This is a clear indication of the blending of data profiling, data-cleansing operations and domain-specific management. Specialist vendors, with a focus on a single functional competence, provide narrow functionality at a lower cost but are increasingly pressured to expand capabilities as more consolidation occurs. A macro trend of convergence of the data quality tools market and the related market for data integration tools continues, as organizations recognize that data integration activities must provide more than simply data delivery — they must ensure the quality of the data being delivered enhances the value of data integration investments.

When evaluating offerings in this market, organizations must consider the breadth of functional capabilities (for example, data profiling, parsing, standardization, matching, monitoring and enrichment) relative to their requirements. Other key criteria include the degree of integration of these capabilities into a single architecture and product — specifically, integration at the metadata level, for example, a single unified metadata repository or the ability to apply findings from one toolset to create inference outcomes in another. Finally, consider nontechnology characteristics, such as the availability of preferable deployment and pricing models, and the size, viability and partnerships of the vendors.

Publication Date: 4 June 2008/ID Number: G00157464

Figure 1. Magic Quadrant for Data Quality Tools



Source: Gartner (May 2008)

Market Overview

Organizations of all sizes and in all industries are recognizing the importance of high-quality data and the critical role of data quality in information governance and stewardship driven by broader enterprise information management initiatives (see "Key Issues for Implementing an Enterprisewide Data Quality Improvement Project, 2008," "Key Issues for Enterprise Information Management Initiatives, 2008" and "Key Issues for Establishing Information Governance Policies, Processes and Organization, 2008"). As a result, their interest in the role of tools and technology for data quality improvement continues to grow. Fueled by a market of purpose-built, packaged tools for addressing various dimensions of the data quality discipline, data quality functionality is readily available from a variety of providers, both large and small. Data quality functionality is also being recognized as a fundamental component of offerings in many related software markets, such as data integration tools, master data management (MDM) solutions and business intelligence (BI) platforms.

The vendors in this market offer a broad range of data quality functionality, ranging from data quality analysis and profiling, to data-cleansing operations such as parsing, standardization and matching, through to data enrichment. Much convergence of technology has occurred, and today vendors offer more functionality within a smaller number of discrete products — most vendors have consolidated the bulk of their core data quality functionality into a single data-cleansing platform, with data profiling remaining the only major functional component commonly sold as a separate product. However, specialized add-on capabilities (such as global name and address support, application-specific knowledgebases and dashboards for data quality metrics) for their core platforms persist, and even grow in number, as vendors adapt their packaging and pricing models to suit a wider range of potential buyers.

One of the most significant trends in this market is the continued expansion of the tools' capabilities beyond the basic data quality operations of parsing, standardization and matching of structured data assets in a narrow set of data domains (for example, customer data only). Increasingly, both new entrants and longtime competitors are delivering technology with a focus on data quality analysis, pervasive deployment of data quality controls, ongoing data quality monitoring and flexibility to address a range of data subject areas. The technology is evolving rapidly in various ways, including:

Data quality assessment and monitoring technology is coming to the fore. One of the most significant areas of R&D investment and innovation in the data quality tools market relates to technology to help organizations measure and monitor levels of data quality. Data profiling tools were a first step toward these goals, but more holistic solutions are emerging.

Data quality capabilities will be deployed and consumed as services. As service-oriented architecture (SOA) and alternative delivery models such as software as a service (SaaS) gain traction, data quality capabilities will be increasingly deployed as services, both internal and external to the organization. Many vendors of data quality tools that have traditionally sold technology only for on-premises deployment by their customers will start to offer hosted solutions for certain types of data quality operations.

Domain-agnostic technology is beginning to dominate. As organizations continue to view data quality in multiple domains (beyond the traditional arena of customer contact data), vendors are building new products or adapting existing ones so that they can address the new range of demands. Vendors with optimized technology for a specific data domain may retain lucrative niches in the market, but will find limited success in enterprise-class scenarios.

Data quality concepts and approaches will emerge for less-structured data. As with data integration tools, data quality technology will expand to address quality assurance of data types beyond the traditional structured variety. The concepts and approaches for applying data quality techniques to these nontraditional data types will evolve in the next couple of years.

In time, vendors that do not act on these trends — and instead continue to focus solely on customer data and traditional approaches to data quality improvement — will fall behind their competition and lose market share.

The market for data quality tools is modest in size (approximately $365 million in software revenue) but will grow at a compound annual rate of 17% or more between 2006 and 2011 (see "Forecast: Data Quality Tools, Worldwide, 2006-2011"), which is stronger than the growth of many other software markets. Much of the innovation continues to come from outside the United States. As a result, the veteran data quality tool vendors are being challenged by entrants that have a more significant international focus. Many new entrants focus on domain-agnostic data quality services (stand-alone or embedded in applications), based on a centrally managed set of business rules. However, with the increasing trend toward embedding data quality capabilities in business applications, data integration tools and other software offerings from larger vendors, these small competitors will face significant challenges as they attempt to survive and grow. Also, acquisition activity in this and related markets continues to change the competitive landscape for data quality tools. For example, SAP's acquisition of Business Objects (see "SAP's Planned Business Objects Buy Signals Strategic Shift") brings significant data quality tools into the SAP portfolio for the first time, while Informatica's acquisition of Identity Systems (see "Informatica

Expands Data Quality With Identity Systems Buy") continues the trend of small data quality specialists being subsumed into the portfolios of larger players in this market.

Market Definition/Description

The data quality tools market comprises vendors that offer stand-alone software products for addressing the core functional requirements of the data quality discipline:

- **Profiling**: Analysis of data to capture statistics (metadata) that provide insight into the quality of the data and aid in the identification of data quality issues.
- **Parsing and standardization**: Decomposition of text fields into component parts and formatting of values into consistent layouts based on industry standards, local standards (for example, postal authority standards for address data), user-defined business rules and knowledge bases of values and patterns.
- **Generalized "cleansing"**: Modification of data values to meet domain restrictions, integrity constraints or other business rules that define sufficient data quality for the organization.
- **Matching**: Identification, linking or merging related entries within or across sets of data.
- **Monitoring**: Deployment of controls to ensure ongoing conformance of data to business rules that define data quality for the organization.
- **Enrichment**: Enhancing the value of internally held data by appending related attributes from external sources (for example, consumer demographic attributes or geographic descriptors).

In addition, these products provide a range of related functional capabilities that are not unique to this market but are required for executing many of the data quality core functions, or for specific data quality applications:

- **Connectivity/adapters** — Ability to interact with a range of different data structure types.
- **Subject-area-specific support** — Standardization capabilities for specific data subject areas.
- **International support** — Relevance for data quality operations on a global basis.
- **Metadata management** — Ability to capture, reconcile and interoperate metadata related to the data quality process.
- **Configuration environment** — Capabilities for creation, management and deployment of data quality rules.
- **Operations and administration** — Facilities for supporting, managing and controlling data quality processes.
- **Service-enablement** — Service-oriented characteristics and support for SOA deployments.

The tools provided by vendors in this market are generally consumed by technology users for internal deployment in their IT infrastructure, although hosted data quality solutions are continuing to evolve and grow in popularity.

Inclusion and Exclusion Criteria

For vendors to be included in the Magic Quadrant, they must meet the following criteria:

- Offer stand-alone (not only embedded in, or dependent on, other products and services) packaged tools that are positioned, marketed and sold specifically for data quality applications.
- Deliver functionality that addresses, at a minimum, profiling, parsing, standardization, cleansing and matching. Vendors offering only narrow functionality (for example, only address cleansing and validation or only matching) are excluded because they do not provide complete data quality tool suites.
- Support this functionality for data in more than one language and specific to more than one country (in the case of address standardization).
- Maintain an installed base of at least 50 production customers for their data quality products.
- Demonstrate, via customer references, use of the tools at an enterprise (cross-departmental or multiapplication) level.

A vendor that does not meet the above criteria may be considered for inclusion if it is a new entrant that is demonstrably different from established vendors and represents a future direction for data quality tools.

There are many data quality tools vendors but most do not meet the above criteria and are, therefore, not included in the Magic Quadrant. Many vendors provide products that address one very specific data quality problem, such as address cleansing and validation, but cannot support other types of applications, or lack the full breadth of functionality expected in today's data quality solutions. Others provide a range of functionality, but operate only in a single country or support only narrow, departmental implementations. Others may meet all the functional, deployment and geographic requirements but are at a very early stage in their "life span" and, therefore, have few, if any, production customers. The following vendors may be considered by Gartner clients alongside those appearing in the Magic Quadrant when deployment needs are aligned with their specific capabilities, or are newer entrants beginning to gain visibility in the market but lacking a significant customer base:

- **AddressDoctor**, Maxdorf, Germany, www.addressdoctor.com — specializes in international address standardization and validation, supporting 240 countries and territories.
- **AMB Dataminers**, Chicago, Illinois, www.payasyougodataquality.com — provides profiling, standardization and cleansing functionality for deployment in Windows environments.
- **Anchor Software,** Plano, Texas, www.anchorcomputersoftware.com — provides a range of data quality utilities supporting common customer list management operations such as file splitting, deduplication and suppression.
- **BackOffice Associates**, Harwich, Massachusetts, www.boaweb.com — offers services and technology for governance of master data within SAP applications.
- **BCC Software** (a division of Bowe Bell + Howell), Rochester, New York, www.bccsoftware.com — provides a range of data quality utilities supporting common

customer list management operations such as address validation, change of address, deduplication and suppression.

- **Business Data Quality**, London, U.K., www.businessdataquality.com — offers products focused on data profiling (BDQ Analysis) and data quality monitoring (BDQ Monitor).
- **Certica Solutions**, Wakefield, Massachusetts, www.certicasolutions.com — provides products focused on validating data against predefined data quality rules.
- **Ciant**, Richardson, Texas, www.ciant.com — provides parsing, standardization and matching functionality for customer data, in support of sales and marketing analytics.
- **Datras**, Munich, Germany, www.datras.de — focuses on the German-speaking markets, providing profiling, standardization and monitoring capabilities.
- **Datiris**, Lakewood, Colorado, www.datiris.com — provides various data profiling techniques for a range of data sources.
- **DQ Global**, Fareham, U.K., www.dqglobal.com — provides matching, deduplication and international address standardization and validation functionality.
- **FinScore**, Renens, Switzerland, www.finscore.com — offers technology for measuring data quality and presenting metrics in a dashboard form.
- **helpIT systems**, Surrey, U.K., www.helpit.com — provides data quality tools oriented toward customer matching, deduplication and suppression operations.
- **Infogix**, Naperville, Illinois, www.infogix.com — provides controls-based technology for auditing and validating the integrity of data within and across systems.
- **Infosolve Technologies**, South Brunswick, New Jersey, www.infosolvetech.com — provides open-source tools (with required service contract) that support profiling, standardization, matching and deduplication operations.
- **InQuera**, Tefen, Israel, www.inquera.com — specializes in technology for standardization, matching and deduplication, with a specific focus on product data.
- **Intelligent Search Technology**, White Plains, New York, www.intelligentsearch.com — develops products for profiling, matching, deduplication and U.S. address correction.
- **Ixsight**, Mumbai, India, www.ixsight.com — offers services for data quality audits, along with products for standardization and deduplication.
- **Melissa Data**, Rancho Santa Margarita, California, www.melissadata.com — supports standardization of names, addresses and phone numbers, and validation of addresses and phone numbers (both via on-premises software and hosted Web services).
- **Omikron**, Pforzheim, Germany, global.omikron.net — provides products for standardization and deduplication of customer name and address data.
- **QAS** (a subsidiary of Experian), London, U.K., www.qas.com — offers global name and address standardization, validation and matching/deduplication functionality.
- **Silver Creek Systems**, Louisville, Colorado, www.silvercreeksystems.com — provides parsing, standardization and matching functionality, with a focus on product data applications.

- **Spad**, Paris, France, eng.spadsoft.com — offers a suite of data quality products for data profiling, monitoring and standardization.
- **SQL Power**, Toronto, Canada, www.sqlpower.ca — provides open-source tools supporting standardization, address validation and deduplication.
- **SRC**, Orange, California, www.extendthereach.com — provides data cleansing in the context of business intelligence applications with a geographic orientation.
- **Stalworth**, San Mateo, California, www.stalworth.com — offers a platform for standardization and cleansing of customer data, including international address validation.
- **TIQ Solutions**, Leipzig, Germany, www.tiq-solutions.de — provides data profiling and data quality dashboards, with a focus on banking, insurance and distribution verticals.
- **Utopia**, Mundelein, Illinois, www.utopiainc.com — offers services and technology for data quality analysis and data standardization, with a focus on product master data.
- **Veda Advantage**, Sydney, Australia, www.vedaadvantage.com — provides software to cleanse and update customer addresses, add phone numbers, merge databases into a single customer view and append segmentation data.
- **WinPure**, Reading, U.K., www.winpure.com — offers low-cost data cleansing, matching and data deduplication software on the Windows platform.
- **Zoomix**, Jerusalem, Israel, www.zoomix.com — delivers technology for adaptive matching and standardization, with a focus on product data.

Gartner will continue to monitor the status of these vendors for possible inclusion in future updates of the Magic Quadrant for Data Quality Tools.

Dropped

- **Fuzzy Informatik** — This vendor was acquired by Business Objects in 2007 and no longer exists as an independent entity.

Evaluation Criteria

Ability to Execute

We evaluate vendors' ability to execute in the data quality tools market by using the following criteria:

- **Product/Service**: How well the vendor supports the range of data quality functionality required by the market, the manner (architecture) in which this functionality is delivered and the overall usability of the tools. Product capabilities are critical to the success of data quality tool deployments and, therefore, receive a high weighting.
- **Overall Viability**: The magnitude of the vendor's financial resources and the strength of its people and organizational structure.
- **Sales Execution/Pricing**: The effectiveness of the vendor's pricing model and the effectiveness of its direct and indirect sales channels.

- **Market Responsiveness and Track Record**: The degree to which the vendor has demonstrated the ability to respond successfully to market demand for data quality capabilities over an extended period.
- **Marketing Execution**: The overall effectiveness of the vendor's marketing efforts, and the degree of "mind share," market share and account penetration the vendor has achieved as a result.
- **Customer Experience**: The quality of the vendor's general customer service, implementation service and technical support, and customers' perception of overall value.

Table 1. Ability to Execute Evaluation Criteria

Evaluation Criteria	Weighting
Product/Service	high
Overall Viability (Business Unit, Financial, Strategy, Organization)	standard
Sales Execution/Pricing	standard
Market Responsiveness and Track Record	standard
Marketing Execution	standard
Customer Experience	standard
Operations	no rating

Source: Gartner

Completeness of Vision

We assess vendors' completeness of vision for the data quality tools market by using the following criteria:

- **Market Understanding**: The degree to which the vendor leads the market in new directions (technology, product, services or otherwise) and its ability to adapt to significant market changes and disruptions. Given the dynamic nature of this market, this item receives a high weighting.
- **Marketing Strategy**: The degree to which the vendor's marketing approach aligns with and/or exploits emerging trends and the overall direction of the market.
- **Sales Strategy**: The alignment of the vendor's sales model to the way customers' preferred buying approaches will evolve over time.
- **Offering (Product) Strategy**: The degree to which the vendor's product road map reflects demand trends in the market and fills current gaps or weaknesses. We also consider the strength of the vendor's strategy regarding delivery models of different types.
- **Business Model**: The overall approach the vendor takes to execute its strategy for the data quality market. With a reasonably high degree of similarity across the vendors in this market, this item receives a low weighting.
- **Vertical/Industry Strategy**: The level of emphasis the vendor places on vertical solutions, and the vendor's depth of vertical expertise. Given the broad cross-industry

nature of the data quality discipline, vertical strategies are less critical and, therefore, this item receives a low weighting.

- **Innovation**: The degree to which the vendor has demonstrated a willingness to make new investments to support the strategy and enhance product capabilities, the level of investment in R&D directed toward development of the tools and the extent to which the vendor demonstrates creative energy. With rapidly evolving technology requirements — in the face of trends such as SOA — and increased competition in the market from large vendors, this item receives a high weighting.
- **Geographic Strategy**: The global presence of the vendor and the manner in which it is achieved (for example, direct local presence, resellers and distributors) in light of the desire of multinational enterprises to exploit common tools worldwide.

Table 2. Completeness of Vision Evaluation Criteria

Evaluation Criteria	Weighting
Market Understanding	high
Marketing Strategy	standard
Sales Strategy	standard
Offering (Product) Strategy	standard
Business Model	low
Vertical/Industry Strategy	low
Innovation	high
Geographic Strategy	standard

Source: Gartner

Leaders

Leaders in the market demonstrate strength across a complete range of data quality functionality, including profiling, parsing, standardization, matching, validation and enrichment. They exhibit a clear understanding and vision of where the market is headed, including recognition of noncustomer data quality issues and the delivery of enterprise-level data quality implementations. Leaders have an established market presence, significant size and a multinational presence (directly or as a result of a parent company).

Challengers

Challengers in the market provide strong product capabilities but may not have the same breadth of offering as Leaders. For example, they may lack several functional capabilities of a complete data quality solution. Challengers have an established presence, credibility and viability, but may demonstrate strength only in a specific domain (for example, only customer name and address cleansing) and/or may not demonstrate a significant degree of thought leadership and innovation.

Visionaries

Visionaries in the market demonstrate a strong understanding of current and future market trends and directions, such as the importance of ongoing monitoring of data quality, engagement of business subject matter experts and delivery of data quality services. They exhibit capabilities aligned with these trends, but may lack the market presence, brand recognition, customer base and resources of larger vendors.

Niche Players

Niche Players often have limited breadth of functional capabilities and may lack strength in rapidly evolving functional areas such as data profiling and international support. In addition, they may focus solely on a specific market segment (such as midsize businesses), limited geographic areas or a single domain (such as customer data), as opposed to positioning toward broader use. Niche Players may have good functional breadth but may have an early-stage presence in the market, with a small customer base and limited resources. Niche Players that specialize in a particular geographic area or data domain may have very strong offerings for their chosen focus area and deliver substantial value for their customers in that segment.

DataMentors

- DataMentors specializes in customer data quality applications, providing matching, linking, standardization and cleansing operations via its DataFuse product, and data profiling capabilities via ValiData. Its partnership with smartFocus enables the vendor to offer campaign management, analytics and mapping capabilities (branded as DataMentors PinPoint). The vendor's roots are in database marketing, with the management team having been involved in large-scale applications of this type for more than 20 years.

- Customer references cite accuracy of matching, ease of use and attractive pricing relative to that of some of the more prominent vendors in the market as key strengths and reasons for their selection of DataMentors' technology. Forthcoming versions of DataFuse will introduce further advancements in ease of use, parallel processing and data quality monitoring functionality. While all the installed base is using the technology in the customer data domain, some customer references indicate use in the product data domain as well.

- The vendor's customer base reflects a higher percentage of hosted (SaaS) implementations than is seen for any other vendor in this market. DataMentors estimates that half its customers are using its technology in a hosted manner and this is reflected in the vendor's customer references.

RECOMMENDED READING

"Magic Quadrants and MarketScopes: How Gartner Evaluates Vendors Within a Market"

"Toolkit: RFP Template for Data Quality Tools"

"Key Issues for Implementing an Enterprisewide Data Quality Improvement Project, 2008"

Acronym Key and Glossary Terms

BI business intelligence

CDQ Customer Data Quality

ETL extraction, transformation and loading

Publication Date: 4 June 2008/ID Number: G00157464

ISV independent software vendor

MDM master data management

SaaS software as a service

SI system integrator

SOA service-oriented architecture

UDC Universal Data Cleanse

VAR value-added reseller

Vendors Added or Dropped

We review and adjust our inclusion criteria for Magic Quadrants and MarketScopes as markets change. As a result of these adjustments, the mix of vendors in any Magic Quadrant or MarketScope may change over time. A vendor appearing in a Magic Quadrant or MarketScope one year and not the next does not necessarily indicate that we have changed our opinion of that vendor. This may be a reflection of a change in the market and, therefore, changed evaluation criteria, or a change of focus by a vendor.

Evaluation Criteria Definitions

Ability to Execute

Product/Service: Core goods and services offered by the vendor that compete in/serve the defined market. This includes current product/service capabilities, quality, feature sets and skills, whether offered natively or through OEM agreements/partnerships as defined in the market definition and detailed in the subcriteria.

Overall Viability (Business Unit, Financial, Strategy, Organization): Viability includes an assessment of the overall organization's financial health, the financial and practical success of the business unit, and the likelihood of the individual business unit to continue investing in the product, to continue offering the product and to advance the state of the art within the organization's portfolio of products.

Sales Execution/Pricing: The vendor's capabilities in all presales activities and the structure that supports them. This includes deal management, pricing and negotiation, presales support and the overall effectiveness of the sales channel.

Market Responsiveness and Track Record: Ability to respond, change direction, be flexible and achieve competitive success as opportunities develop, competitors act, customer needs evolve and market dynamics change. This criterion also considers the vendor's history of responsiveness.

Marketing Execution: The clarity, quality, creativity and efficacy of programs designed to deliver the organization's message to influence the market, promote the brand and business, increase awareness of the products, and establish a positive identification with the product/brand and organization in the minds of buyers. This mind share can be driven by a combination of publicity, promotional, thought leadership, word-of-mouth and sales activities.

Customer Experience: Relationships, products and services/programs that enable clients to be successful with the products evaluated. Specifically, this includes the ways customers receive technical support or account support. This can also include ancillary tools, customer support programs (and the quality thereof), availability of user groups and service-level agreements.

Operations: The ability of the organization to meet its goals and commitments. Factors include the quality of the organizational structure including skills, experiences, programs, systems and other vehicles that enable the organization to operate effectively and efficiently on an ongoing basis.

Completeness of Vision

Market Understanding: Ability of the vendor to understand buyers' wants and needs and to translate those into products and services. Vendors that show the highest degree of vision listen and understand buyers' wants and needs, and can shape or enhance those with their added vision.

Marketing Strategy: A clear, differentiated set of messages consistently communicated throughout the organization and externalized through the Web site, advertising, customer programs and positioning statements.

Sales Strategy: The strategy for selling product that uses the appropriate network of direct and indirect sales, marketing, service and communication affiliates that extend the scope and depth of market reach, skills, expertise, technologies, services and the customer base.

Offering (Product) Strategy: The vendor's approach to product development and delivery that emphasizes differentiation, functionality, methodology and feature set as they map to current and future requirements.

Business Model: The soundness and logic of the vendor's underlying business proposition.

Vertical/Industry Strategy: The vendor's strategy to direct resources, skills and offerings to meet the specific needs of individual market segments, including verticals.

Innovation: Direct, related, complementary and synergistic layouts of resources, expertise or capital for investment, consolidation, defensive or pre-emptive purposes.

Geographic Strategy: The vendor's strategy to direct resources, skills and offerings to meet the specific needs of geographies outside the "home" or native geography, either directly or through partners, channels and subsidiaries as appropriate for that geography and market.

REGIONAL HEADQUARTERS

Corporate Headquarters
56 Top Gallant Road
Stamford, CT 06902-7700
U.S.A.
+1 203 964 0096

European Headquarters
Tamesis
The Glanty
Egham
Surrey, TW20 9AW
UNITED KINGDOM
+44 1784 431611

Asia/Pacific Headquarters
Gartner Australasia Pty. Ltd.
Level 9, 141 Walker Street
North Sydney
New South Wales 2060
AUSTRALIA
+61 2 9459 4600

Japan Headquarters
Gartner Japan Ltd.
Aobadai Hills, 6F
7-7, Aobadai, 4-chome
Meguro-ku, Tokyo 153-0042
JAPAN
+81 3 3481 3670

Latin America Headquarters
Gartner do Brazil
Av. das Nações Unidas, 12551
9° andar—World Trade Center
04578-903—São Paulo SP
BRAZIL
+55 11 3443 1509