Magic Quadrant for Data Integration Tools

Published: 3 August 2017   **ID:** G00314940

Analyst(s): Mark A. Beyer, Eric Thoo, Mei Yang Selvage, Ehtisham Zaidi

The data integration tool market has established a focus on transformational technologies and approaches demanded by data and analytics leaders. The presence of legacy, resilient systems and innovation all in the market together requires robust, consistent delivery of highly developed practices.

Market Definition/Description

The market for data integration tools includes vendors that offer software products to enable the construction and implementation of data access and data delivery infrastructure for a variety of data integration scenarios. These include:

- **Data acquisition for business intelligence (BI), analytics and data warehousing —** Extracting data from operational systems, transforming and merging that data, and delivering it to integrated data structures for analytics purposes. The variety of data and context for analytics is expanding as emergent environments — such as nonrelational and Hadoop distributions for supporting discovery, predictive modeling, in-memory DBMSs, logical data warehouse architectures and end-user capability to integrate data (as part of data preparation) — increasingly become part of the information infrastructure. With the increased demand to integrate machine data and support Internet of Things (IoT) and digital business ecosystem needs for analytics, data integration challenges intensify.

- **Sourcing and delivery of application and master data in support of application data management and master data management (MDM) —** Enabling the connectivity and integration of the data representing critical business entities such as customers, products and employees. Data integration tools can be used to build the data access and synchronization processes to support application data management and also MDM initiatives.

- **Data consistency between operational applications —** Data integration tools provide the ability to ensure database-level consistency across applications, both on an internal and an interenterprise basis (for example, involving data structures for SaaS applications or cloud-resident data sources), and in a bidirectional or unidirectional manner. The IoT is specifically exerting influence and pressure here. Data consistency has become critical with new functionality in DBMS offerings — hinting that the battle for data integration is heating up to include the traditional data management vendors.

- **Interenterprise data sharing —** Organizations are increasingly required to provide data to, and receive data from, external trading partners (customers, suppliers, business partners and

others). Data integration tools are relevant for addressing these challenges, which often consist of the same types of data access, transformation and movement components found in other common use cases.

- **Populating and managing data in a data lake.** The emerging concept of a "data lake" — where data is continuously collected and stored in a semantically consistent approach similar to a traditional DBMS, or with an expectation that data processing efforts will refine the semantics of a nontraditional DBMS (such as nonrelational data stores) to support data usage. The need for integrating nonrelational structures and distributing computing workloads to parallelized processes (such as in Hadoop and alternative NoSQL repositories) elevates data integration challenges. At the same time, it also provides opportunities to assist in the application of schemas at data read time, if needed, and to deliver data to business users, processes or applications, or to use data iteratively. In addition, the differing structure of IoT or machine data is introducing new integration needs.

- **Data migration.** Previously considered a data integration style in its own right, data migration is more of a task that can be done with a variety of tools or techniques. The primary feature of data migration is moving data to a new platform or to an update of an existing data management platform. It can also include moving data from one application to a new application or to an upgraded version of an application.

The usage of data integration tools may display characteristics that are not unique to just one of these individual scenarios. Technologies in this market are required to execute many of the core functions of data integration, which can apply to any of the above scenarios. Examples of the resulting characteristics include:

- Increasingly, data integration tools are expected to collect, audit and monitor information regarding the deployed data integration services and processes in the organization. This ranges from use cases for simple reporting and manual analysis to the inclusion of recommendations and even automated performance optimization. While primarily focused on management tasks, the ability to profile new data assets and recognize their similar nature and use cases as compared to other data currently integrated is growing in importance. Small devices that roam and attach to data portals will also become prevalent. The requirement for metadata capabilities will become the center of all integration approaches.

- Interoperating with application integration technology in a single solution architecture to; for instance, expose extraction, transformation and loading (ETL) processes that extract data from sources as a service to be provisioned via an enterprise service bus (ESB). Increasingly, there is a demand for analyzing and integrating data during "business moments" when events demand an in-process operational change based upon data-driven decisions.

- Enabling data services as an architectural technique in a service-oriented architecture (SOA) context. Rather than the use of data integration per se, this represents an emerging trend for data integration capabilities to play a role and to be implemented within software-defined architecture for application services.

- Integrating a combination of data residing on-premises and in SaaS applications or other cloud-based data stores and services, to fulfill requirements such as cloud service integration.

Organizations are also seeking the capability for pivoting between cloud and on-premises — in enabling a hybrid integration platform (HIP).

- Connecting to, and enabling the delivery of data to — and the access of data from — platforms typically associated with big data initiatives such as Hadoop, nonrelational and cloud-based data stores. These platforms provide opportunities for distributing data integration workloads to external parallelized processes.

Magic Quadrant

Figure 1. Magic Quadrant for Data Integration Tools

Vendor Strengths and Cautions

Actian

Based in Palo Alto, California, U.S., Actian offers data integration capabilities via Actian DataConnect and Actian DataCloud. Actian's customer base for data integration tools is estimated to be approximately 7,000 organizations.

Strengths

- **Market presence:** By focusing on targeted aspects of the overall data integration market — through messaging-style solutions, bulk data movement, and alignment to Actian's B2B solutions — a relatively large market reach as a company provides Actian with leverage to create opportunities for its data integration tools.

- **Performance and manageability:** Good performance and throughput for integrating data and centralized management of integration processes are attractive propositions for organizations emphasizing these requirements. Additional functionality planned on Apache Spark Streaming sets out to extend Actian's support for big data to include metadata analysis for the stream.

- **Synergistic product strategy:** Provisioning of functionality through a portfolio of complementary capabilities for integrated use is cited as a key value by Actian customers and implementation partners. Actian plans to align key integration products in its portfolio into a unified platform: including DataCloud, Business Xchange and DataConnect.

Cautions

- **Upgrade complexity and metadata support:** Reference customers for Actian expressed difficulties with version upgrades, the technical complexity of migrating between major releases, and the quality of documentation. They also cited metadata management and modeling functionalities as a relative weakness. Actian's latest release of DataConnect, version 11, offers direct and automatic import of mappings and other artifacts from version 9 to version 11, to ease upgrades from older versions.

- **Limited guidance and support for implementation:** The availability of skilled implementers and guidance for common and best practices are growing concerns among Actian's reference customers, who desire readily accessible self-help resources for implementation approach and issue resolution.

- **Low appeal to diversifying integration roles:** Actian has its roots in technical offerings that are aligned well to communities in the IT arena. This runs contrary to current market momentum that tends strongly toward the needs of multiple personas and self-service integration options for less-technical and nontechnical users. Actian is addressing this through a browser-based, graphical user interface to support the needs of business roles to perform basic and intuitive integration tasks.

Adeptia

Based in Chicago, Illinois, U.S., Adeptia offers the Adeptia Integration Suite (AIS) and Adeptia Connect. Adeptia's customer base for this product set is estimated to be more than 550 organizations.

Strengths

- **Delivers core competency within an ESB:** Adeptia supports the core requirements of bulk/batch data delivery and granular data capture and propagation with a combination of its data integration capability, application integration, ESB, B2B integration and trading partner management.

- **Attractive pricing and flexibility:** Reference customers view Adeptia's products as being attractively priced relative to its competitors; they also value its flexible subscription licensing options. Adeptia's ability to promote the interoperation of data integration functionalities with capabilities for ESB and business process management (BPM) is greatly appreciated according to reference customers and Gartner inquiry discussions.

- **Performance, usability and integration platform as a service support:** Adeptia Integration Suite offers integration platform as a service (iPaaS) capabilities, which enable B2B integration and interenterprise data sharing use cases. Its products support integration of on-premises endpoints, cloud endpoints and a combination of the two for integration patterns that support "pervasive integration." Reference customers also cite ease of use, good performance and throughput as strengths, which are particularly relevant for enabling citizen integrators and capitalizing on business moments.

Cautions

- **Weakness supporting big data initiatives:** Adeptia's current implementations and competitive bids indicate extremely limited traction in support of big data initiatives, although connector capability is available to plug in external components such Apache Spark, Hive and Pig. Since data integration into big data stores (such as Hadoop and nonrelational) is increasingly being emphasized in the market, pressures for enabling upcoming and popular use cases (enabling a data lake, for example) will grow.

- **Narrow market functionality coverage:** Reference customers appreciate Adeptia's support for traditional use cases involving bulk/batch, operational and BPM scenarios. However, its product roadmap does not include incorporating other comprehensive styles of data integration (data virtualization, for example) into its AIS platform for exploiting the entire breadth of data integration use cases, which poses a challenge in competitive situations.

- **Limited capability to integrate with other data management solutions:** Reference customers expressed a desire for better interoperability and integrated usage with related technologies for data management and application integration (including data quality, data governance, metadata management and MDM). Adeptia continues to address these concerns by actively expanding its network of technology partners.

Attunity

Based in Burlington, Massachusetts, U.S., Attunity offers Attunity Replicate, Attunity Compose and Attunity Visibility. Attunity's customer base for this product set is estimated to be approximately 2,500 organizations globally.

Strengths

- **Strength and stability of targeted functionality:** Attunity offers strong data integration functionality in the key areas of data replication and synchronization technology applied to heterogeneous data types, with a historical strength in addressing mainframe data. Customers favor Attunity's business longevity in supporting data consistency while also addressing the modern requirements of cloud and big data scenarios. Attunity has also targeted cloud data integration and is listed as a preferred data integration partner for both Amazon Web Services and Microsoft Azure.

- **Time to value:** Attunity's tooling supports data integration specialists as well as, increasingly, less technically skilled personnel in order to make data available across applications and data structures for BI and analytics. The role-based interface allows citizen integrators to integrate data quickly for their purposes. Design and administrative tools are used by enterprise architects and designers when more robust solutions are required. The result is a mix of time-to-value of delivery that matches many different business needs.

- **Aligned with modern data integration approaches:** Integration activities involving Apache Kafka, Spark and Hadoop in cloud environments extend Attunity's established experience in supporting analytics and data warehousing to meet the increasing challenges found in event-driven data requirements and a mix of cloud and on-premises integration scenarios.

Cautions

- **User experience focused on targeted data integration styles:** Adoption of Attunity's product set is predominantly for its change data capture (CDC)/replication capability. The continuing shift of buyer demands toward comprehensive data delivery styles, and of integrated usage between data integration activities with related technologies for extensive metadata support, application integration, and information governance needs, poses challenges for Attunity in competitive situations.

- **Specific administrative operations being addressed:** Some reference customers have reported specific issues with restart and recovery. Attunity has released operations management capabilities for addressing these requirements and enhancements in operational control and administration and is developing metadata and operations management to address these specific enterprise demands.

- **Demand for documentation and low availability of skills:** As deployments of Attunity increase in complexity, reference customers have expressed their concern about the availability of skilled resources (which are difficult to find). Improved documentation is also desired for administrative operation and implementation guidance.

Cisco

Based in San Jose, California, U.S., Cisco offers the Cisco Information Server (CIS). Cisco's customer base for this product set is estimated to be around 400 organizations.

Strengths

- **Strong roadmap for IoT integration use case:** Leveraging a unique opportunity that has arisen from its leadership in network technology, Cisco has concentrated its data integration investments to support the IoT use case, specifically on operational technology (OT)/IT convergence. As a result, Cisco is introducing new and innovative capabilities for IoT. Currently, this is more aspirational than actual, but combines Cisco's historic strength in networks with data virtualization capabilities that have the potential to be "cutting edge."

- **Leverages brand and market presence on global scale:** Cisco's brand is well-known worldwide for network capabilities. By adding data integration along its communication backbone, it will be able to expand quickly into developing markets where networks and data integration can proceed together (for example, in Eastern Europe and Asia).

- **Established capability in data virtualization:** CIS has been a leading brand in data virtualization for more than 10 years. Reference customers praise the product's stability, broad connectors, data catalog and optimization engine, as well as its ability when serving as an enterprise semantic layer. It is important to note that Cisco is one of only two data-virtualization-focused vendors present on the Magic Quadrant — even beating out some smaller, broad-based tools for the spots.

Cautions

- **Isolating on data virtualization may limit other data integration styles:** With Cisco's tight focus on IoT use cases, its innovations for other data virtualization use cases (such as application data access and consumption) and other data integration styles (such as batch, data replication and messaging) have taken a back seat. CIS could eventually become absorbed by Cisco's IoT platform and therefore cease to remain a truly independent data virtualization product.

- **High price and complicated license models:** Because data virtualization needs to be complemented by other data integration styles, end-user organizations often have limited budgets for data virtualization tools. Increasing numbers of Gartner inquiries have reported that Cisco's pricing and complicated licensing models have prevented them from adopting CIS or have created procurement challenges.

- **Lack of skilled resources and inadequate knowledge transfer:** Cisco reference customers as well as Gartner inquiry clients report challenges in finding qualified developers in the market and cite inadequate knowledge transfer after the product "went live." Cisco is addressing these concerns in the product roadmap through establishing a knowledge base, communities, enhanced self-service and expanding its implementation partner base.

Denodo

Based in Palo Alto, California, U.S., Denodo offers the Denodo Platform as its data virtualization offering. Denodo's customer base for this product set is estimated to be around 380 organizations.

Strengths

- **Strong momentum, growth and mind share within the data virtualization submarket:** The Denodo Platform is mentioned in almost all competitive situations involving data virtualization under Gartner's contract review service — a significant increase during the past year that is highlighted by expanding partnerships with technology and service providers. Denodo has added offerings on the AWS marketplace for subscription and pay-as-you-go-based licensing options. Denodo is gaining momentum in 2017, and is one of only two data-virtualization-focused vendors present on the Magic Quadrant.

- **Broad connectivity support and streaming support:** All available connectors are included within the Denodo Platform's cost. Denodo has also increased its connectivity support for streaming data (with support for Kafka message queues, Apache Storm, Spark Streaming, and so on); and cloud services on the Amazon Web Services (AWS) and Azure marketplaces; as well as partnerships with database platform as a service (dbPaaS) vendors such as AWS Redshift and Snowflake. Denodo can also interoperate with Docker technology.

- **Mitigation for traditional virtualization issues:** Denodo Platform is a mature data virtualization offering that incorporates dynamic query optimization as a key value point. This capability includes support for cost-based optimization specifically for high data volume and complexity; cloud environments; predicate optimization techniques, including full and partial aggregation pushdown below joins; and partitioned unions for logical data warehouse and data lake requirements. It has also added an in-memory data grid with Massively Parallel Processing (MPP) architecture to its platform in supporting incremental caching of large datasets, reusing complex transformations and persistence of federated data stores.

Cautions

- **Challenged to support multiple data delivery styles:** Denodo has limited support for ETL, CDC/replication and messaging. Denodo is incorporating features for data quality and data governance into its platform, but its mind share in these related markets continues to be lower than that of other incumbent vendors with these capabilities.

- **Pricing issues and contract flexibility:** Some of Denodo's existing customers (reference customers, reference interviews and in Gartner events' one-on-one discussions) report the total cost of ownership (TCO) and high price points (particularly for its enterprise unlimited licensing option) as barriers to adoption. Reference customers also indicate that the average maintenance and support paid on a yearly basis is high.

- **Various user experience issues due to growth:** Denodo customers indicate good support and responsiveness to customer queries; however, a small but significant number report their desire for better documentation, training (particularly on new features), customer support (largely

driven by log reviews) and user experience. A recent release of new web-based documentation and enhancements of the integrated development environment (IDE) are among Denodo's focus on initiatives for customer experience.

IBM

Based in Armonk, New York, U.S., IBM offers the following data integration products: IBM InfoSphere Information Server Enterprise Edition, IBM InfoSphere Information Server Enterprise Hypervisor Edition, IBM InfoSphere Federation Server, IBM InfoSphere Data Replication, IBM Data Integration for Enterprise, and IBM Data Integration for Hadoop, IBM Big Insights BigIntegrate, IBM Streams and IBM Bluemix Data Connect (previously DataWorks). IBM's customer base for this product set is estimated to be more than 10,700 organizations.

Strengths

- **Innovation aligned with traditional and emerging market trends:** IBM's stack includes all integration styles for enterprise data integration projects requiring a mix of data granularity and latencies. In 2017, the solution offers expanded capabilities for Avro, Parquet, Kafka, Hive, Ambari, Kerberos and other open-source solutions. For cloud, IBM offers CDC for Hadoop via its WebHDFS (that is, Hadoop Distributed File System) component, and Bluemix for hybrid integration scenarios.

- **Global presence and mind share:** IBM continues to successfully draw on its global presence and extensive experience and mind share in data integration and management. It is frequently mentioned by users of Gartner's inquiry service and often features in competitive evaluations on our contract and RFP reviews. It continues to gain traction due to its significant ecosystem of partners, value-added resellers, consultants and external service providers.

- **Comprehensive portfolio meets integration demands across diverse personas:** IBM continues to invest in data quality, data governance, MDM and application integration (through rapid growth in iPaaS). Its investments in a design module that has separate interfaces for users with varying levels of integration skills (for example, integration specialists, citizen integrators) replaces the aging Information Analyzer workbench, along with Bluemix Data Connect (for self-service data preparation), with Watson Analytics.

Cautions

- **Confusion over integration across product set and product messaging:** Reference customers cited confusion about IBM's varied portfolio of offerings — a perennial issue. Difficulties in understanding integrated deployments as more products are added (including integration of IBM's data integration tools alongside other IBM products), are a challenge amid the growing usage scenarios in data integration. They also expressed confusion over IBM's frequent renaming of its solutions, which could lead to redundancy and shelfware.

- **User issues regarding installation, migrations and upgrades:** Although IBM continues to address migration complexity through in-place migrations and significant improvements to its user experience, some reference customers still reported continuing upgrade difficulties with

some versions, indicating a need for IBM to further improve the user experience offered by its data integration tools.

- **High costs for licensing models and perceived TCO:** Gartner inquiry clients regularly cite high TCO as one of primary reasons for overlooking IBM in competitive situations. This, coupled with the complexity of its licensing options and confusion regarding new cost models, was identified as a major inhibitor of adoption. While IBM's provision of varied licensing approaches — such as processor value unit (PVU), node-, workgroup-, bundle-, subscription (monthly and yearly)- and perpetual-based models — is intended to provide more choice, it has reportedly also confused customers.

Informatica

Based in Redwood City, California, U.S., Informatica offers the following data integration products: Informatica Platform (including PowerCenter, PowerExchange, Data Replication, Advanced Data Transformation, Ultra Messaging, B2B Data Transformation, B2B Data Exchange, Data Integration Hub), Informatica Data Services, Informatica Intelligent Cloud Services, Cloud Integration Hub, Big Data Management, Big Data Integration Hub, Informatica Intelligent Streaming, Informatica Intelligent Data Lake and Informatica Data Preparation. Informatica's customer base for this product set is estimated to be more than 7,000 organizations.

Strengths

- **Strong vision for product strategy across cloud and big data management:** Informatica continues to deliver on its vision of a unified integrated platform for all data delivery styles, for a broad range of use cases. Strong interoperability and synergies between Informatica's data integration tools and other data management technologies encourage usage as an enterprise standard that links with data quality, MDM, metadata management, data governance, data hub, data lake, big data analytics, data security and iPaaS technologies.

- **Roadmap aligned with emerging trends and multipersona demand:** Informatica's Cloud-scale AI-powered Real-time Engine (CLAIRE) technology for metadata-driven AI along with strong integration with its Enterprise Information Catalog (its metadata management tool) delivers on the emerging trends vision with an effective metadata management and data governance approach for data in cloud, on-premises and big data environments. Informatica also offers self-service data preparation for diverse personas that range from centralized IT to citizen integrators.

- **Broad market presence, global reach and mind share:** Informatica's mind share in the data integration market is as the leading brand appearing in just over 70% of all contract reviews and in the vast majority of competitive situations in our inquiry calls. It has a significant presence in established as well as new or growing geographies for data integration, with a well-established global network of more than 500 partners.

Cautions

▪ **Poor perceptions of pricing:** Informatica's data integration tools are historically perceived as being high-cost, with hardware-based perpetual license models. Informatica is executing a shift toward subscription pricing models, including pay-as-you-go and hourly pricing models, and offering PowerCenter and Big Data Management on the AWS and Azure Marketplaces. Reference customers still cite their concerns over Informatica's strategy to continue charging separately for some targeted connectors.

▪ **Lack of clarity in product messaging and portfolio architecture:** Gartner continues to receive inquiry calls citing confusion about Informatica's overlapping product functionality and features. Reference and inquiry customers seem confused and are buying similar products with overlapping features and capabilities, which often leads to shelfware or redundancy. Informatica hopes its rebranding and positioning of the company, redesign of its website, and new sales and partner enablement programs will address some of these concerns.

▪ **Challenges for new market offerings and global strategy:** Informatica now focusses extensively on hybrid data integration, cloud data integration, big data management and metadata management. It is also expanding its focus on industry-specific offerings, bundled solutions and non-IT user personas. It's still early days, and Informatica would need a multiyear vision and change management execution — complete with significant sales training — in order to execute well on this strategy.

Information Builders

Based in New York City, New York, U.S., Information Builders offers the following data integration products: the iWay Integration Suite (composed of iWay Service Manager and iWay DataMigrator) and iWay Universal Adapter Suite. Information Builders' customer base for this product set is estimated to be more than 840 organizations.

Strengths

▪ **Positive customer experience:** Information Builders' reference customers continue to report overall satisfaction (in the top 25% for this Magic Quadrant) with the product and its balance of features and functionality; the only caveat being the ease of use of the interface (described as "moderate"). The broad connectivity of the overall platform and the strength of the technology are enhanced by the customer experience — for issue resolution and best practices — from its professional services; customers report a feeling of "partnership." Additionally, Information Builders has worked with and promoted interactions in its user community to facilitate peer-to-peer support.

▪ **Diverse data integration capability:** In a continuation of last year's theme, Information Builders is improving on the model for "making big data normal." With the capability to combine bulk/batch with message-oriented data (for real-time) as well as with diverse assets (commonly referred to as "unstructured" data), this vendor has demonstrated some very complex data integration implementations during the past 12 months.

- **Alignment with essential data management and integration needs:** The breadth of the product portfolio and Information Builders' experience in deployments of various data integration styles aligns it well with the contemporary trends of related markets (for example, its Omni-Gen platform aligns data integration with the adjacent technologies of data quality and MDM).

Cautions

- **Lack of appeal for diverse roles:** Information Builders still appeals mainly to technical communities and IT buyers, though the focus on self-service data preparation has begun to gain some traction. That said, this vendor has maintained its position even though the entire market has shifted down and to the left on the Magic Quadrant.

- **Skills hard to find and training materials and documentation need improvement:** Recent reference customer input indicates that last year's documentation weakness persists, and specifically states that finding online materials is difficult. Customers also report that it is very difficult to locate experienced resources, and recommend aligning your team early when choosing Information Builders.

- **Inability to gain mind share:** Gartner's client inquiries indicate that Information Builders is not considered as frequently as its market-leading competitors. This relative lack of mainstream recognition represents a disadvantage that needs addressing.

Microsoft

Based in Redmond, Washington, U.S., Microsoft offers data integration capabilities via SQL Server Integration Services (SSIS), which is included in the SQL Server DBMS license. Microsoft also includes data integration as part of the Azure Data Factory. Microsoft SQL Server deployments are inclusive of SSIS for data integration (Microsoft does not report a specific customer count for SSIS).

Strengths

- **Productivity and time to value:** SSIS supports connectivity to diverse data types and broad deployment in Microsoft-centric environments. Wide use of SSIS by SQL Server customers has resulted in widely available community support, training and third-party documentation on implementation practices and approaches to problem resolution.

- **Synergies among data, applications, business roles and artificial intelligence:** SSIS is often used to put data into SQL Server, to enable analytics, data management and end-user data manipulation using Microsoft's Office tools, particularly Excel. Using SSIS in conjunction with Microsoft's BizTalk and Azure Data Factory platforms enables delivery of data from enterprise business workflows and data preparation. Microsoft builds its data integration customer base by embedding it into use cases for other Microsoft products (for example, data delivery to the Cortana Intelligence Suite to establish synergy between data integration, cloud and cognitive computing capability.

- **Brand awareness and market presence:** Microsoft's size and global presence provide a huge customer base and a distribution model that supports both direct and channel partner sales. Adoptions continue, because Microsoft products are among the most familiar to implementers when considering interfaces, development tools and functionality.

Cautions

- **Integration lacking in the portfolio:** Microsoft is a large company with somewhat isolated product roadmaps and no significant indication that this will change in the near future. Integrated implementation of Microsoft's offerings indicates difficulty for the integrated deployment of an expanding range of offerings and functionality when discerning optimal ways for manipulating and delivering data of interest alongside Azure Data Factory, data quality and governance activities.

- **Evolution perceived as Microsoft-focused:** There are concerns among customers about Microsoft's roadmap becoming tightly linked to Azure- and Cortana-related platforms, somewhat contradictory to its emphasis on the "any to any" heterogeneous needs of data integration. This is nothing new to either Microsoft or its customers and most can proceed as they have in the past.

- **Limited platform supported:** The inability to deploy SSIS workloads in non-Windows environments is a limitation for customers wanting to draw on the processing power of diverse hardware and operating environments. Microsoft has release plans for SQL Server 2017 on Linux.

Oracle

Based in Redwood Shores, California, U.S., Oracle offers the following data integration products: Oracle Data Integrator (ODI), Oracle Data Integrator Cloud Service, Oracle GoldenGate, Oracle GoldenGate Cloud Service, Oracle Data Service Integrator and Oracle Service Bus. Oracle's customer base for this product set is estimated at more than 10,800 organizations.

Strengths

- **Customer base directs product innovations:** Oracle has invested in and rolled out several focused upgrades to its data integration product portfolio during 2016. GoldenGate's "zero footprint" technology means that it doesn't have to be installed on the source or target system. Oracle has ensured tight integration between ODI and GoldenGate via Kafka for the delivery of data in real-time opening deployment options to include Kappa or Lambda architectures for stream analytics. ODI's ability to push down processing to Spark augurs well for big-data-related use cases. Reference customers report that Oracle's knowledge modules are an effective way to decouple business logic from ETL code.

- **Role and self-service support:** Synergy of data integration alongside complementary capabilities in Oracle Big Data Preparation (which is now also available as a cloud service) adds support for natural-language processing and graph features, along with improved integration

with its metadata manager and machine-learning capabilities to empower non-IT roles (particularly citizen integrators) to build, test and then make operational integration flows.

- **Global partner network eases implementation worries:** Oracle uses its large partner network to assist customer implementations in finding implementation service providers for Oracle technologies in data integration. Along with this, Oracle has a diverse portfolio of solutions for data integration and supporting data management technologies — including iPaaS, metadata management, data quality, and data governance — which allow existing customers to expand with Oracle.

Cautions

- **Issues with pricing flexibility and perceived value:** Calls with Gartner customers on our contract review and inquiry services, along with ratings on our Peer Insights, indicate concerns with the cost model for customers desiring more flexibility in license models and pricing options — representing one of the biggest reasons for prospects choosing the competition over Oracle on our contract reviews. Through a range of term-based, subscription-based, name-based, metered-based pricing options offered (even hourly on GoldenGate and ODI), Oracle is trying to address this concern using more flexible packaging.

- **Difficulties with migration and upgrades:** Reference customers continue to cite issues with migration to the newer versions of Oracle's data integration tools. Some existing customers reported significant challenges with bugs in existing versions and upgrades impacted by these bugs — causing Oracle's scores to be lower than average for this customer satisfaction category.

- **Limited appeal to new users for new and extended offerings:** Oracle appears in less than 30% of data integration competitive situations for projects of inquiries from Gartner clients. Gartner recognizes that Oracle does have capable solutions across the breadth of data integration styles; however, its existing and new customers seem unaware of the maturity and relevance of these solutions across the range of modern integration needs.

Pentaho

Pentaho is a Hitachi Group Company with its global headquarters in Orlando, Florida, U.S., and worldwide sales based in San Francisco, California, U.S. Pentaho does not report customer counts for specific products, but has more than 1,500 commercial customers.

Strengths

- **Broadening use-case-agnostic integration solution:** Pentaho Data Integration (PDI) provides data integration across a broad spectrum of relational DBMSs, Java Database Connectivity (JDBC)/Open Database Connectivity (ODBC) access, and cloud-based data management solutions. During the past three and more years, Pentaho has positioned its data integration tool as an agnostic solution that is increasingly capable of delivering against independent targets and enterprise-class demands. PDI includes a large number of prebuilt data access and

preparation components, a rich GUI for data engineers, orchestration of integration components and an integrated scheduler that can interoperate with enterprise system schedulers.

- **Experience in cloud, on-premises and hybrid:** Pentaho's customer reference base includes examples of all three deployment models of data integration, including very large customers across back-office, IoT and machine/sensor data solutions, as well as traditional data integration demands. Loads to Amazon Redshift and integration with Amazon Elastic MapReduce (EMR), and Cloudera, as well as embedded R, Python and Spark Machine Learning library (MLlib) models in the integration stream, capitalize on deployment needs.

- **Market-awareness of open source and roles:** PDI already works well within Apache Spark and other distributed processing environments, and is addressing issues such as load balancing with task isolation to enhance distributed processing operations. Pentaho leverages open-source solutions (such as Kafka) to mix real-time integration with batch/bulk capability. Pentaho's existing capability in BI has been added to the PDI capability that allows users to visualize data integration results in-line and identify data quality problems before moving to production deployments.

Cautions

- **Market mind share is low:** Organizations that are not familiar with PDI still consider Pentaho to be more of an analytics vendor — which limits the interest expressed by the market. In response, Pentaho has improved its marketing as a stand-alone data integration tool provider. PDI pull-through is exhibited in the ability of Pentaho to sell a complete analytic platform to customers after PDI is in place. The Hitachi brand is also showing pull-through, especially for major global brands.

- **Development environment can be improved:** Error handling in job execution and a need for extensive developer feedback are reported by Pentaho's customer references. As an open-source-based solution this is to be expected, in part, and the user documentation is reported as being above expectations (for open source), which also helps.

- **Focus on big data detracts from other use cases:** With good capabilities for Hadoop and other big-data-related solutions, as well as satisfactory throughput in batch processes (Hadoop is effectively also batch), it is important to keep the use cases aligned with the tools' capabilities. The redistribution and isolation of workloads to improve performance exhibit some limitations, which Pentaho plans to resolve in its next release.

SAP

Based in Walldorf, Germany, SAP offers the following data integration products: SAP Data Services, SAP Replication Server, SAP Landscape Transformation Replication Server, SAP Remote Data Sync, SAP Data Hub, SAP Hana platform, SAP Cloud Platform Integration and SAP Event Stream Processor. SAP's customer base for this product set is estimated at more than 27,000 organizations.

Strengths

- **Solution strategy leverages product breadth:** SAP continues to deliver strong data integration functionality in a broad range of use cases to its customers. A mix of granularity, latency, and physical and virtualized data delivery supported alongside complementary offerings of iPaaS, self-service data preparation, information governance and MDM, all combine to position SAP data integration solutions for complex problems in the SAP ecosystems. Introduction of a "try and buy" licensing approach enables organizations to incrementally gain familiarity with role-based functionality prior to purchase.

- **Relevance for highly connected and distributed digital architecture:** SAP's investment in serverless and in-memory computing, cluster management, and data hub offering to support data integration architecture extends distributed pushdown executions to capitalize on the processing capacity of cloud DBMS, cloud storage and IoT infrastructures. SAP's roadmap in this market includes rule-based autoscaling and self-adjusting optimization for data integration workstreams, in order to capitalize on machine learning.

- **Synergy across data and application integration, governance and analytics solutions:** Assimilating diverse data integration tooling in SAP's unified digital data platform vision enables components to share metadata, one design environment and administrative tooling, and to operate as a hub that supports data management, analytics and application requirements. A huge global customer base using its diverse data management infrastructures and applications has led to extensive adoption of SAP's data integration tools.

Cautions

- **Overcoming market perception of "SAP focus":** SAP's approach to product development and delivery emphasizes differentiation through its strategic SAP Hana platform and is well-positioned for customers using or moving to Hana; however, it is perceived less favorably by prospective buyers whose data doesn't predominantly reside in the SAP ecosystem.

- **Too many entry points into the organization:** Implementations by reference clients cite difficulties with the integrated deployment of SAP's offerings across its portfolio as more products are added. This often takes place when the enterprise begins to address the growing scale and complexity of usage scenarios in data integration activities that have been deployed independently across the organization. SAP hopes to address the concern of too many entry points with its newly released SAP Data Hub.

- **Latent concerns regarding customer support, service experience and skills:** While customer experience for SAP has improved overall, reference customer feedback indicates concerns about both the processes for obtaining product support and the quality and consistency of support services and skill availability, as areas needing further progress.

SAS

Based in Cary, North Carolina, U.S., SAS offers the following data integration products: SAS Data Management (including SAS Data Integration Server and SAS Data Quality), SAS Federation Server,

SAS/Access interfaces, SAS Data Loader for Hadoop and SAS Event Stream Processing. SAS's customer base for this product set is estimated to be 14,000 organizations.

Strengths

- **Continued breadth (and integrated nature) of offerings:** SAS leverages a strong metadata strategy to link its MDM and data quality offerings with its data integration capabilities (which are broad enough to include migration, bulk/batch, virtualization, message-based, synchronization and streams processing). This breadth is particularly useful when processing, managing and enriching data for analytics.

- **Embedded data integration for operational decisions:** SAS Decision Manager is specifically called out for its ability to create data correlations that flow into operational decision models. This is a subtle advantage that is not entirely unique to SAS, but where SAS has strong credentials to enhance its credibility. If a decision model is developed, it may actually require different data inputs when deployed from many different sources — and, based upon the business process model, may occur at different points along the process stream. SAS can restate the metadata used to render the source inputs, and allow for a true "develop once, use many" model that is traceable throughout the organization.

- **Interactive development:** Most modern data integration tools include highly interactive development interfaces that allow users to manipulate and work with data; in SAS's case, another leverage point for metadata is the rapid integration of new sources. The introduction of machine learning through SAS Viya has created a link between enterprise data integration and data discovery. The close relationship between data auditing and profiling, and how they are integrated with data integration development, is the key to quickly adding new analytics data sources.

Cautions

- **Market mind share constrained at times to SAS installed base:** SAS data integration products are still considered by many users as being specific add-on products for supporting SAS analytics solutions. However, beginning before 2014 and continuing into 2017, SAS data integration solutions have expanded their use cases beyond analytics only. SAS has maintained a strategy of enhancing data integration to support analytics throughout the duration of its participation in this market. It may simply not be possible (or even desirable) to add a significant number of customers beyond SAS's core market, which is already a substantial customer base.

- **Some inconsistent usage experiences:** SAS Data Integration Studio is a component of many other SAS products; as such, some customers express increasing difficulty to ensure full cross-platform and complex-use-case compatibility. SAS continues to chase unexpected issues (previously, fixes have been needed for ODS Graphics procedures, temporary file and directory issues, inconsistent migration between product versions, inconsistent database references and more) that are often related to isolated metadata integration needs and detract from the otherwise significant metadata capabilities of its products.

- **Installation, upgrade, and migration concerns:** Version upgrade difficulties are cited as concerns by SAS's reference customers, highlighting the need to improve ease of installation,

reduce product complexity, and increase self-guided support for simplifying migration. The roadmap of SAS Viya sets out to improve the installation and upgrade experience of customers.

Syncsort

Based in Pearl River, New York, U.S., Syncsort offers DMX, DMX-h and Ironstream. Syncsort's customer base for this product set is estimated to be around 2,000 organizations.

Strengths

- **Attractive cost and low TCO:** Syncsort's reference customers praise its competitive pricing and low TCO compared with other data integration price leading vendors. Overall, low TCO is often cited by its customers as one of main reasons for choosing Syncsort.

- **High performance ETL products:** Syncsort builds its reputation on its high-performance ETL and big data integration tools. Enabling on-premises and in-cloud deployments, which Syncsort refers as its "design once, deploy anywhere" architecture, focuses on extending the flexibility and scalability of data integration processing. In recent years, it has built strong partnerships with well-known brands in big data such as Cloudera, Hortonworks and Splunk. Many customers leverage Syncsort to integrate mainframe data and Hadoop.

- **Improved position within the broader data management market:** The acquisition of Trillium (a data quality tool leader) in November 2016, has positioned Syncsort to offer a more comprehensive data integration and data quality solution. Syncsort has improved its ability to support business-centric use cases such as customer 360-degree view, fraud detection, and data governance.

Cautions

- **Lack of mind share:** Syncsort's data integration marketing and sales strategies are rooted in its expertise and strengths in accessing and integrating mainframe data. This has worked well for building a niche in the market with competitive differentiation, and a loyal customer base, but the same tactic also works against Syncsort as it expands into new markets. Gartner inquiry customers report little awareness of the Syncsort brand beyond the mainframe market, even though it has expanded its offerings to embrace Hadoop and the cloud; for example, accessing the data warehouse and integrating data into Hadoop data lakes.

- **Anticipated slow integration following the Trillium acquisition:** Although Syncsort has crafted a new product offering strategy based on the combined solutions of Syncsort and Trillium, and has taken initial steps toward integrating them for data lake governance and customer 360-degree use cases, we do not yet have a clear idea of how well these two sets of products will integrate with each other on an architectural level (that is, shared metadata and common UIs). Without this deeper level of integration, there would be fewer synergistic benefits for customers.

- **Limited range of data integration capabilities:** Although Syncsort's functionality continues to be extended (for example, ingestion of the mainframe log data stream), its predominant data

integration capabilities remain ETL-centric. Although this reflects a specialization that Syncsort chooses to focus on in order to better serve its customers, it also presents a competitive disadvantage when data integration requirements include data virtualization. Syncsort is expanding its integration styles through recently added real-time CDC capabilities for populating Hadoop data lakes with changes in mainframe data, and certified support for messaging integration (for Kafka, MapR Technologies MapR Streams, Apache NiFi and connectivity to IBM MQ and Pivotal's RabbitMQ).

Talend

Based in Redwood City, California, U.S., Talend offers Talend Open Studio, Talend Data Fabric, Talend Data Management Platform, Talend Platform for Big Data, Talend Data Services Platform, Talend Integration Cloud and Talend Data Preparation. Talend's paying customer base for this product portfolio is estimated at more than 1,500 organizations.

Strengths

- **Cost model and flexibility:** Through a scalable licensing model based on a per-developer subscription fee, Talend allows customers to start with small, core data integration projects and then grow their portfolio for more advanced data integration projects (such as integration with Hadoop data stores).

- **Integrated portfolio for data integration and interoperability with complementary technologies:** Talend possesses a comprehensive portfolio of data integration and related technology (including data quality, MDM, ESB, application integration and metadata management), interoperates with Docker, and has recently added iPaaS and data preparation capabilities. Gartner inquiry and reference customers alike report a robust product set, which allows them to build and execute end-to-end data management projects and use cases and to capitalize on data integration use cases that require synergy with their related technologies.

- **Strength in core data integration capabilities and delivery for evolving trends:** Customers and prospects are still drawn to Talend's robust core data integration capabilities (including the bulk/batch movement of data), which continue to draw in a significant proportion of its buyer base. Talend also has products catering to current and evolving market needs, including its iPaaS offering (now supporting AWS, Google Cloud Platform and Microsoft Azure integration) and data preparation; significant investment in data integration operations running natively on Hadoop, and evolving operational uses cases (in the Apache Storm and Apache Spark environments); planned features for data lake governance; and partnerships with Cloudera Navigator and Hortonworks for their integration with Apache Atlas.

Cautions

- **New release stability and implementation support:** Reference customers' adoption experiences have sometimes included problems with the stability and performance of Talend's new releases, and also with finding adequate partners/skilled resources that are adept with Talend design and implementation. Talend has launched a new partner certification program and is working with partners to design new reference architectures.

- **Developer focus:** Talend has its roots in open source, and with the technical community in general. Current market momentum is moving strongly toward enabling multiple personas and self-service integration options for nontechnical users. Talend has started addressing more personas with self-service via Talend Data Preparation, hybrid cloud integration capabilities through iPaaS, and support for information stewardship. Talend is investing in a new partner certification program and training for partners and customers.

- **Lack of market awareness beyond bulk/batch:** While Talend's capabilities resonate well with traditional data delivery styles, reference customer concerns and Gartner inquiries both indicate a need to increase awareness of its support for other data integration styles (particularly replication/synchronization of data for real-time integration and data virtualization). More comprehensive and integrated metadata management support across its product portfolio is also desired.

Vendors Added and Dropped

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

Added

- Pentaho

Dropped

- None

Inclusion and Exclusion Criteria

The inclusion criteria represent the specific attributes that analysts believe are necessary for inclusion in this research.

To be included in this Magic Quadrant, vendors must possess within their technology portfolio the subset of capabilities identified by Gartner as the most critical from within the overall range of capabilities expected of data integration tools. Specifically, vendors must deliver the following functional requirements:

- Data delivery modes support — At least three modes are supported among bulk/batch data movement, federated/virtualized views, message-oriented delivery, data replication, streaming/event data, and synchronization. Vendors whose customer reference base fails to represent, in any mix of their products' use, three of the following seven technical deployment styles will be excluded:

- Bulk/batch includes single or multipass/step processing that includes the entire contents of the data file after an initial input or read of the file is completed from a given source or multiple sources. All processes take place on multiple records within the data integration application before the records are released for any other data-consuming application.

- Message-oriented utilizes a single record in an encapsulated object that may or may not include internally defined structure (XML), externally defined structures (electronic data interchange), a single record or other source that delivers its data for action to the data integration process.

- Virtualization is the utilization of logical views of data, which may or may not be cached in various forms within the data integration application server or systems/memory managed by that application server. Virtualization may or may not include redefinition of the sourced data.

- Replication is a simple copy of data from one location to another, always in a physical repository. Replication can be a basis for all other types of data integration, but specifically does not change the form, structure or content of the data it moves.

- Synchronization can utilize any other form of data integration, but specifically focuses on establishing and maintaining consistency between two separate and independently managed create, read, update, delete (CRUD) instances of a shared, logically consistent data model for an operational data consistency use case (may or may not be on the same data management platform). Synchronization also maintains and resolves instances of data collision with the capability to establish embedded decision rules for resolving such collisions.

- Streaming/event data consists of datasets that follow a consistent content and structure over long periods of time and large numbers of records, and that effectively report status changes for the connected device or application or continuously update records with new values. Streaming/event processing includes the ability to incorporate event models, inferred row-to-row integrity, and variations of either those models or the inferred integrity with alternative outcomes that may or may not be aggregated and/or parsed into separate event streams from the same continuous stream. The logic for this approach is embedded in the data stream processing code.

- Data services bus (SOA) capability is the ability to deploy any of the various data integration styles, but with specific capability to interoperate with application services (logic flows, interfaces, end-user interfaces, and so on) and pass instructions to, and receive instructions from, those other services on the bus. Data services bus includes auditing to assist in service bus management, either internally or by passing audit metadata to another participating service on the bus.

- Data transformation support — At a minimum, packaged capabilities for basic transformations (such as data type conversions, string manipulations and calculations).

- Demonstrably broad range of connectivity/adapter support (sources and targets) — Native access to relational DBMS products, plus access to nonrelational legacy data structures, flat

files, XML and message queues, as well as emerging data asset types (such as JavaScript Object Notation [JSON]).

- Mode of connectivity/adapter support (against a range of sources and targets), support for change detection, leveraging third-party and native connectors, connection and read error detection, and integrated error handling for production operations.

- Metadata and data modeling support — Automated metadata discovery (such as profiling new data sources for consistency with existing sources), lineage and impact analysis reporting, ability to synchronize metadata across multiple instances of the tool, and an open metadata repository, including mechanisms for bidirectional sharing of metadata with other tools.

- User- or role-specific variations in the development interface capable of various workflow enhancement mechanisms, which may include supporting templates, version modification (via internal library management or other mechanism), quality assurance capability via either audit/monitor metadata (manual) or embedded workflows (administrator tools).

- Design and development support — Graphical design/development environment and team development capabilities (such as version control and collaboration). This includes multiple versions running in disparate platforms and multiple instances of services deployments in production environments as well as alternative or collaborating development environments.

- Runtime platform support — Windows, Unix or Linux operating systems, or demonstrated capability to operate on more than one commercially available cloud environment regardless of the platform in operation.

- Service enablement — The ability to deploy functionality as services, including multiple operating platforms. The ability to manage and administer operations on multiple platforms and environments is significantly desired.

- Data governance support — Ability to import, export and directly access metadata with data profiling and/or data quality tools, master data management tools and data discovery tools. Accepting business and data management rule updates from data stewardship workflows and sharing data profiling information with such tools is highly desired. No additional advantage is perceived in data integration for also delivering actual data governance tools — the focus is interoperability.

In addition, vendors had to satisfy the following quantitative requirements regarding their market penetration and customer base. Vendors must:

- Generate at least $25 million of their annual software revenue from data integration tools (perpetual license subscription or maintenance/support), or maintain at least 300 maintenance-paying customers for their data integration tools. Gartner will use as many independent resources for validating this information as possible, specifically to validate provided information.

- Support data integration tool customers in at least two of the following geographic regions: North America, South America, Europe and Asia/Pacific.

- Demonstrated market presence will also be reviewed and can be assessed through internal Gartner search, external search engines, Gartner inquiry interest, technical press presence and activity in user groups or posts. A relative lack of market presence could be determined as a reason to exclude a product/service offering.

Vendors could be excluded if they focus on narrow use cases that are too specific for broader market application. Some vendor/supplier tools were excluded because:

- They focused on only one horizontal data subject area; for example, the integration of customer-identifying data

- They focused only on a single vertical industry

- They served only their own, internally managed data models and/or architectures (this includes tools that only ingest data to a single proprietary data repository) or were used by a single visualization or analytics processing platform

Evaluation Criteria

Ability to Execute

Gartner analysts evaluate technology providers on the quality and efficacy of the processes, systems, methods or procedures that enable IT providers' performance to be competitive, efficient and effective, and to positively affect revenue, retention and reputation. Ultimately, technology providers are judged on their ability to capitalize on their vision, and their success in doing so.

We evaluate vendors' Ability to Execute in the data integration tool market by using the following criteria:

- **Product/Service:** Core goods and services that compete in and/or serve the defined market. This includes current product and service capabilities, quality, feature sets, skills and so on. This can be offered natively or through OEM agreements/partnerships (as defined in the Market Definition section, or described below). Some consumers are prepared to accept less-capable products from many different suppliers and assemble them together on their own. Connecting data integration activities to data quality and governance-related capabilities (such as master data management) becomes an integral support for all use cases that can share high-quality data as well as lineage and nonlineage metadata, with runtime management and monitoring support. For broader spectrum solutions, the market has de-emphasized the product capability and emphasized the ability to break out pricing and components. How well the vendor supports the range of distinguishing data integration functionalities required by the market, the manner (architecture) in which this functionality is delivered, support for established and emerging deployment models, and the overall usability and consumption of the tools are crucial to the success of data integration tool deployments.

- **Overall Viability:** Viability includes an assessment of the organization's overall financial health as well as the financial and practical success of the business unit. It views the likelihood of the organization to continue to offer and invest in the product, as well as the product's position in

the current portfolio. Overall vendor viability is reviewed and utilized by end-user organizations and developers in determining a supplier's capability to deliver ongoing production support. Importantly, open-source solutions are measured here by the strength of their community and the overall capability of the governing body to guide the roadmap and manage open-source projects. The appropriateness of the vendor's financial resources, the continuity of its people, and its technological consistency affect the practical success of the business unit or organization in generating business results.

- **Sales Execution/Pricing:** The organization'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. Organizations increasingly seek "severability," or the capability to isolate on specifically required functions that are then reflected in their implementation approach and cost allocations. The focus on pricing by verticals — which allows for pricing by use case, role, volumetric and performance metrics; all considered applicable for different market needs — has increased in 2017. In addition, pricing by features and functionality is increasingly sought to allow for flexible use cases within familiar toolsets. The effectiveness of the vendor's pricing model in light of current customer demand trends and spending patterns, and the effectiveness of its direct and indirect sales channels were scored as part of the evaluation.

- **Market Responsiveness/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 to changing market demands. Market track record is itself one measure of market responsiveness, and in this case data integration tools are much like other infrastructure-focused solutions. Often, organizations demand data virtualization, message-oriented data movement, replication and synchronization and streaming/event processing. Traditional bulk-batch processing is still the predominant demand. Not only do most solutions overlap, but the market is demanding a capability to deliver all forms of integration to differently skilled implementers with everything from simple data preparation through self-service data integration to enterprise-class systems. The degree to which the vendor has demonstrated the ability to respond successfully to market demand for data integration capabilities over an extended period and how well the vendor acted on the vision of prior years, are scored as part of the evaluation.

- **Marketing Execution:** The clarity, quality, creativity and efficacy of programs designed to deliver the organization's message in order to influence the market, promote the brand, increase awareness of products and establish a positive identification in the minds of customers. This "mind share" can be driven by a combination of publicity, promotion, thought leadership, social media, referrals and sales activities. Marketing execution was traditionally considered to be the positioning and declarations of a supplier, but now end-user organizations use it frequently as a gauge of how in-tune supplier roadmaps are with overall market demand. Suppliers need to be aware of emerging best practices for data management infrastructure and if they and their customers can specifically benefit from specialized horizontal or vertical capabilities, geographically targeted approaches or partner-supported implementation practices. The overall effectiveness of the vendor's marketing efforts, which impact its mind share, market share and

account penetration, is important. The ability of the vendor to adapt to changing demands in the market by aligning its product message with new trends and end-user interests was scored as part of the evaluation.

- **Customer Experience:** Products and services and/or programs that enable customers to achieve anticipated results with the products evaluated. Specifically, this includes quality supplier/buyer interactions technical support, or account support. This may also include ancillary tools, customer support programs, availability of user groups, service-level agreements, and so on. Data integration has evolved to include a broad range of expectations when it comes to customers' experience. The level of satisfaction expressed by customers with the vendor's product support, professional services, overall relationship with the vendor, and their perceptions of the value of the vendor's data integration tools relative to costs and expectations is part of the evaluation. This criterion retains a weighting of "High" to reflect buyers' scrutiny of these considerations as they seek to derive optimal value from their investments. In 2017, the distinction between advanced use cases and "pedestrian" applications is becoming highly distinct. The evaluation this year will focus on separating success in "traditional" market delivery from "innovative" in reviewing customer experience. Analysis and rating of vendors against this criterion will continue to be driven directly by the results of a customer survey executed during 2017.

- **Operations:** The ability of the organization to meet goals and commitments. Factors include the quality of the organizational structure, skills, experiences, programs, systems and other vehicles that enable the organization to operate effectively and efficiently. Operations are not specifically differentiating to end-user markets — but product management consistency and support/maintenance practices add to the overall customer experience as well as the stability of senior staff. Suppliers need to demonstrate a new balance in their R&D allocation to ensure they are positioned for deploying with greater focus on data services, metadata management and semantic tiers, as well as provide ongoing support for the massive bulk/batch data movement market.

Table 1. Ability to Execute Evaluation Criteria

Evaluation Criteria	Weighting
Product or Service	High
Overall Viability	Medium
Sales Execution/Pricing	High
Market Responsiveness/Record	Medium
Marketing Execution	Low
Customer Experience	High
Operations	Medium

Source: Gartner (August 2017)

Completeness of Vision

Gartner analysts evaluate technology providers on their ability to convincingly articulate logical statements about current and future market direction, innovation, customer needs and competitive forces, as well as how they map to Gartner's position. Ultimately, technology providers are assessed on their understanding of the ways that market forces can be exploited to create opportunities.

We assess vendors' Completeness of Vision for the data integration tool market by using the following criteria:

- **Market Understanding:** Ability to understand customer needs and translate them into products and services. Vendors that show a clear vision of their market will listen, understand customer demands, and can shape or enhance market changes with their added vision. A visionary market understanding recognizes the importance of advanced information management/ integration to support both operational and analytics data use cases. Applications and data management must both address the concept of role-based development. "Citizen" integrators will want rapid access to data without concerns for production optimization, and analytic assistance for data auditing, profiling, qualifying and conformance/alignment will be critical — but will need metadata-driven warnings as well as template library management to support their efforts. The degree to which the vendor leads the market in new directions (in terms of technologies, products, services or otherwise) is key, alongside its ability to adapt to significant market changes and disruptions.

- **Marketing Strategy:** Clear, differentiated messaging consistently communicated internally, externalized through social media, advertising, customer programs, and positioning statements. Marketing is now experience-based and not as susceptible to presentations and collateral development from suppliers. In addition, suppliers must develop a means of converting community "chatter" and excitement to support delivery and go-to-market campaigns. Redesign and redeployment when going into broader implementations is considered suboptimal, so a flow from trial versions into pilot and then production is desired.

- **Sales Strategy:** A sound strategy for selling that uses the appropriate networks including: direct and indirect sales, marketing, service and communication. Also, partners that extend the scope and depth of market reach, expertise, technologies, services and their customer base. This criterion covers the alignment of the vendor's sales model with the ways in which customers' preferred buying approaches will evolve over time. Scaled pricing models are becoming particularly interesting. Suppliers must consider if their internal compensation models incentivize delivery that matches customer demand and implementation profiles. Customers and prospects are less concerned with "positioning" and more concerned with "try then buy" models. This increases the demand in the market for limited freeware versions that can be easily converted to robust solutions once proven.

- **Offering (Product) Strategy:** An approach to product development and delivery that emphasizes market differentiation, functionality, methodology and features as they map to current and future requirements. Existing markets and use cases have begun to weaken in favor of more distributed data integration needs — which increases the demand for self-healing and

wizards/tutors for recognizing new sources and information asset types. Product strategy vision includes the roadmap for continued support of traditional integration needs filling current gaps, weaknesses and opportunities to capitalize on less advanced demand trends in this market. In addition, given the requirement for data integration tools to support diverse environments for data, delivery models and platform-mix perspective, we assess vendors on the degree of openness of their technology and product strategy.

- **Business Model:** The design, logic and execution of the organization's business proposition to achieve continued success. A visionary business model will balance the emerging (and increasingly stringent) demand for managing internal and external compliance and risk while providing support for existing customers. While broad, all-inclusive models represent one solution approach, it is also expected and reasonable to assume that tightly targeted models for traditional delivery needs can cut delivery cost, increase adoption and deliver specific integration needs to end-user organizations. The overall approach the vendor takes to execute its strategy for the data integration tool market — including diversity of delivery models, packaging and pricing options, and partnership — is important.

- **Vertical/Industry Strategy:** The strategy to direct resources (sales, product, development), skills and products to meet the specific needs of individual market segments, including verticals. This is the degree of emphasis the vendor places on vertical solutions, and the vendor's depth of vertical market expertise.

- **Innovation:** Direct, related, complementary and synergistic layouts of resources, expertise or capital for investment, consolidation, or for defensive or preemptive purposes. The current innovation demands in the market are centered on managing location-agnostic capability. Integration should run on-premises and in the cloud, and switch between them. As data becomes highly distributed, data integration activities are also required to become easily distributable to any data location, or recommending/determining when data needs to be moved for optimal processing. As information management use cases gain in importance to focus on transient data (traditionally the forte of message-oriented technologies), demand for converging data and application integration approaches is rapidly increasing. Important here is the degree to which the vendor demonstrates creative energy in the form of enhancing its practices and product capabilities, as well as introducing thought-leading and differentiating ideas and product plans that have the potential to significantly extend or reshape the market in a way that adds real value for customers. The growing diversity of users indicates a much higher demand for administrative, auditing, monitoring and even governance controls that utilize job audit statistics.

- **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. Data tracing will become a key requirement in the geographic distribution of data. Development platforms must include the ability to monitor where data originates with jurisdictional cognizance and where it is eventually delivered. Violating national laws through data movement must be addressed, and policy-level controls are expected to safeguard the citizen developer and the cloud deployment. The vendor's strategy for expanding its reach into markets beyond its home region/country, and its approach to achieving global presence (for example, its direct

local presence and use of resellers and distributors), is critical for capitalizing on global demands for data integration capabilities and expertise.

Table 2. Completeness of Vision Evaluation Criteria

Evaluation Criteria	Weighting
Market Understanding	High
Marketing Strategy	Low
Sales Strategy	Medium
Offering (Product) Strategy	Medium
Business Model	Medium
Vertical/Industry Strategy	Low
Innovation	High
Geographic Strategy	Medium

Source: Gartner (August 2017)

Additional information regarding the evaluation criteria can be found in Note 1.

Quadrant Descriptions

Leaders

Leaders in the data integration tool market are front-runners in the capability to support a full range of data delivery styles. They have also recognized the growing affinity between data and application integration and are haltingly approaching location-agnostic deployments (those that are not limited to cloud or on-premises, but can be deployed beyond a specific location). They are strong in establishing data integration infrastructure as an enterprise standard and as a critical component of modern information infrastructure. They support both traditional and new data integration patterns in order to capitalize on market demand. Leaders have significant mind share in the market, and resources skilled in their tools are readily available. These vendors recognize, and design to deploy for, emerging and new market demands — often providing new functional capabilities in their products ahead of demand, and by identifying new types of business problem to which data integration tools can bring significant value. Examples of deployments that span multiple projects and types of use case are common among Leaders' customers. Leaders have an established market presence, significant size and a multinational presence (directly or through a parent company).

Challengers

Challengers are well-positioned in light of the key existing practices in the data integration tool market, such as the need to support multiple styles of data delivery. However, they may be limited to specific technical environments or application domains. In addition, their vision may be affected by a lack of coordinated strategy across the various products in their data integration tool portfolio. Challengers generally have substantial customer bases, and an established presence, credibility and viability, although implementations may be of a single-project nature, or reflect multiple projects of a single type (for example, predominantly ETL-oriented use cases). Importantly, more than 80% of all end-user organizations in the world still seek batch/bulk processing (even to and within the cloud) — this means that a highly efficient, but batch-oriented vendor can exhibit high-level execution capabilities without ever crossing to the "right half" of the Magic Quadrant. Traditional Leaders in this market may actually "regress" in Completeness of Vision as they focus on existing demand and maximizing revenue relative to the current market.

Visionaries

Visionaries demonstrate a strong understanding of emerging technology and business trends, or focus on a specific market need that is far outside of common practices while also being aligned with capabilities that are anticipated to grow in demand. They sometimes lack market awareness or credibility beyond their customer base or a single application domain. Visionaries may also fail to provide a comprehensive set of product capabilities — including those that focus on a single data integration style and simply import, export or leverage that primary functionality to create alternative outputs for transformed or integrated data. They may be new entrants lacking the installed base and global presence of larger vendors, although they could also be large, established players in related markets that have only recently placed an emphasis on data integration tools. The growing emphasis on aligning data integration tools with the market's demand for interoperability of delivery styles, integrated deployment of related offerings (such as data integration and data quality tools), metadata modeling, support for emerging information and application infrastructures, and deployment models (among other things) are creating challenges for which vendors must demonstrate vision. With the Leaders now addressing big data and data virtualization needs, focusing on these capabilities is still highly advantageous, but no longer a specific differentiator for the "top half" of the Magic Quadrant (in terms of Ability to Execute).

Niche Players

Niche Players have gaps in either, or both, their Completeness of Vision or Ability to Execute. They often lack key aspects of product functionality and/or exhibit a narrow focus on their own architectures and installed bases. Niche Players may have good functional breadth, but a limited presence and mind share in this market. With a small customer base and limited resources, they are not recognized as proven providers of comprehensive data integration tools for enterprise-class deployments. Many Niche Players have very strong offerings for a specific range of data integration problems (for example, a particular set of technical environments or application domains) and deliver substantial value for their customers in the associated segment. Niche Players often exhibit a particular advantage in pricing or in their small "footprint" that makes them ideal candidates to be a "best fit" solution that complements other technology in the data management infrastructure of an organization. Importantly, Niche Players in this market have demonstrated their capability to

outperform dozens of tool and solution offerings that were considered and eventually excluded from this Magic Quadrant.

Context

In 2016, we stated that data integration requirements had become so onerous that applications should no longer be permitted to be silos first and integrated second. The question now is, "Will this data be used in adjacent use cases?" and the answer is *always* "yes." During the past 18 months, the pace of information capture has continued to grow well beyond the capacity to analyze and use it. During the past three decades, information technology has swung back and forth between abundant capacity for processing, storage, memory and even networks (and being forgiving of poor design) versus data volumes and process demands that overwhelm any planned capacity. The pendulum has begun to swing back to a position where capacity is no longer abundant when compared to data volume and availability. Cloud providers are now encountering the same connectivity and management issues that on-premises providers encountered a decade ago. We have to get smarter.

Enterprises pursuing frictionless sharing of data are increasingly favoring tools that are flexible — in that they can be designed once for delivery across multiple platforms, mixed architectural approaches and broad deployment, without significant rework. The market shift we indicated in 2016 is taking its form in 2017. Most importantly, as "enterprise-class" data integration platforms grow their functionality toward a broadening spectrum, their complexity and pricing goes up. This leaves a gap in the market where specialists with small footprint solutions and a specific focus can become the best at doing "one thing." It has become increasingly difficult for those vendors who are neither Leaders nor Visionaries to maintain their position in the market. The demand for easily deployed basic functionality is accelerating with self-service data preparation (SSDP) demands; some suppliers will be tempted to, and will specifically, pursue "down market" sales that have smaller margins. From a broader data integration perspective, SSDP is considered a submarket that crosses between data integration and BI and advanced analytics tools. As such, we do not evaluate SSDP solutions in this Magic Quadrant (see "Market Guide for Self-Service Data Preparation"). When market demand changes and vendors continue to only enhance established functional requirements, they may position targeted offerings to their advantage even though margins go down, because customer counts go up; being a Challenger in this scenario actually becomes a market advantage if properly pursued.

The proliferation of data types, and applications and devices that capture data, means that digital businesses are facing intensifying data integration challenges — coupled with the rise of citizen integrators that has already added complexity to the mix. There is almost a "personality clash" developing between digital consumers and digital businesses. Periods of rapid change will yield to stable periods, then resume their dynamic nature for only "part" of the business model. The concept of "hybrid" cloud and on-premises will begin to erode and then vanish; data is just data and process is just process.

Importantly, the concept of using smaller, more targeted solutions in data integration practices continues to grow and is permitting tightly focused vendors that do not address the entire

functionality spectrum to achieve good market execution. They are filling the "gap" between broad-spectrum integration for a price and hand-coded or application-architecture-style integration. Since data integration is one constant in the IT universe, implementers do not always seek a more complete solution — because integrating multiple integration tools is the basis of the practice, just like integrating data. This trend is what drives the concept of roles and more-active metadata analysis to be built into the tools.

Finally, while many vendors have reference customers reporting issues with total cost of ownership, there remains a significant level of willful ignorance in the market that the user organizations are causing this problem. Deferring data integration to postapplication deployment is a growing "technical debt." That debt *must* be serviced at the time applications and device apps are integrated — there is no option to leave it as is at that point. This effectively generates a new maintenance point, with a cost that is specifically deferred from the two anchor points on either end. Even in cases where a centralized integration architecture (not necessarily centralized data) is attempted, the design and deployment *must* reach into each data location, which both incurs cost and consumes staff time. Automated discovery and machine learning can provide assistance within tools and platforms to address some portion of this cost, but the software is then a higher value proposition and thus of a higher price. Until application development teams are forced to recognize that some data must always be integrated, some data is integrated some time, and some data does not need to be integrated — and they start planning and designing for all three rings of information governance — they are contributing to their own high TCO for data integration through an irresponsible approach.

Market Overview

The discipline of data integration comprises the practices, architectural techniques and tools that ingest, transform, combine and provision data across the spectrum of information types in the enterprise and beyond in order to meet the data consumption requirements of all applications and business processes.

The biggest change in the market from 2016 is the pervasive yet elusive demand for metadata-driven solutions. Consumers are asking for hybrid deployment not just in the cloud and on-premises (which is metadata-driven combined with services distribution), but also across multiple data tiers throughout broad deployment models, plus the ability to blend data integration with application integration platforms (which is metadata driven in combination with workflow management and process orchestration) and a supplier focus on product and delivery initiatives to support these demands.

The Ability to Execute data integration in a hyperconnected infrastructure (irrespective of structure and origins) through active metadata ingestion, sharing and processing; and the ability to recommend and autoexecute transformations through machine learning, are some of most important requirements that separate traditional tools from modern integration tools.

Technology evolutions in this market to address shifting demand trends, and interim steps to combine cloud and on-premises deployments and broadening hybrid integration approaches are

well on their way to common and broad adoption by 2025 (see "The State and Future of Data Integration: Optimizing Your Portfolio of Tools to Harness Market Shifts").

Adoption of a hybrid approach is increasing in the wake of the cloud-first focus of some digital business strategies, which emphasizes the use of more-lightweight technologies that are user-oriented and adaptable to change. IT professionals like to tout the fast pace of innovation and change; historically, however, this is largely fiction (and remains so in the current era), because computing infrastructure is very large and diverse; for example:

- Data management and integration techniques change in eight- to 10-year cycles (data integration tools took nine years to overcome custom-coding; data warehouses took 11 years to become mainstream.

- Nonrelational DBMS has taken approximately 10 years (from 2005 to 2015) to become mainstream.

- Hadoop and the big data phenomenon emerged around 2010, and had been developed for more than 10 years before 2005 — when it was open-sourced.

This means that "hybrid" is an interim term and will eventually give way to data integration approaches that dynamically reconfigure the optimization and delivery strategy regardless of deployment locations — even migrating themselves to new locations in the expanding, distributed computing infrastructure as we head toward 2025.

In 2017, organizations have begun to take three related pathways to integrating their data:

- First, and foremost, the massive amounts of data being collected by sensor and operational devices has increased the amount of data that can be used in complicated models to determine how single "actors" can participate in multiple events and their analysis. This is best exemplified by the initiative in almost every industry to automate operational processes through the use of models trained by algorithms using various data inputs. These models have become layered decision engines; for example, regarding energy consumption as more than device operation — rather, as a complex model involving environments, cultural influences, alternative fuels, fuel quality, maintenance and more. This is an intensive, deliberate data integration model.

- Second, human behavior is participating in, or interfering in, a broad range of data collection models. There are opt-in strategies that allow participants to self-select to contribute more data — potentially increasing their influence, but also creating a demand to find other data sources to validate the data from these "volunteers." Opt-in usually means that the participant expects to benefit from the data collection; however, there are also passive systems that may simply use location information and then relate it to other data — as a potential validating dataset to counter-balance the opt-in model. This creates a mesh of human awareness of the new information stream that can be produced in some cases (humans now generate "digital exhaust" everywhere they go, both physically and digitally). Data integration efforts are thus being driven to provide more adaptive models that can be easily embedded; also to leverage machine-learning techniques that are capable of not only recognizing data types and structure, but going beyond into understanding applicable use cases and capacity and into utilizing the

data involved (especially with regard to time and capacity to process the data, communication of metadata about those assets, and more). This embedded, intelligence-driven style of data integration has gone beyond an embryonic stage, but is still immature.

- Third, application and data integration are closely related and, with business processes changing rapidly, either application logic/design or data management/processing must always be the first line of adaptation. Often, the application or the data alone are incapable of absorbing the speed of change in operational process models. As a result, the influence of hybrid platforms for integration are rapidly gaining ground. However, a new challenge has emerged in this space in the form of microservices, which attempt to create reusable data management logic that is encapsulated within a processing object. The definition of integration is thus challenged in a significant way that will alter the landscape in a new world where point-to-point integration between data stores or even applications falls into disuse.

In 2017, however, traditional integration still makes up the bulk of the delivery in the market (easily more than 80% of all organizations make significant use of bulk/batch). While many organizations have the traditional solutions in place, modern demands have increased the utilization of message, virtualization and synchronization so that in a composite somewhere between 35% to 45% of all organizations are using at least two of these other approaches — and an even higher percentage of leading or large organizations are doing so. Distributed mobile devices, consumer apps and applications, multichannel interactions, and even social media interactions are driving these organizations to build highly sophisticated integration architectures that can just as easily be a simple data transfer protocol to a fully contextualized data service that delivers single data points through streams of information in near real time (see "Predicts 2017: Data Distribution and Complexity Drive Information Infrastructure Modernization").

Gartner estimates that the data integration tool market generated more than $2.7 billion in software revenue (in constant currency) at the end of 2016. A projected five-year compound annual growth rate of 6.32% will bring the total market revenue to around $4 billion in 2021 (see "Forecast: Enterprise Software Markets, Worldwide, 2014-2021, 2Q17 Update").

The following trends reflect both the ongoing demand and shifts in demand from buyers in 2017, as well as areas of opportunity for technology providers to deliver thought leadership and innovation in order to extend this market's boundaries:

- **Machine-learning begins to rise to dominance in data/information integration practices.** Systems and applications use data within use-cases — even analytic applications are merely a use case. The flow of data between analytics and back to operational support is moving more toward real time than ever before, but the driver is metadata and metadata-driven machine-learning processes. Too much data that is beyond the capability for processing, needs to be "trimmed" down dynamically and sized to both the use case and the capacity for utilizing that data. While massive processing models can be run on cloud platforms, connectivity, communications, validity, data freshness and most importantly the trust-rating of data sources will allow for dynamic, targeted data delivery to all manner of use cases. Data management is the past, application management is the present, and data integration is the future of information technology.

- **Growing demand for real time and recognition of the required speed of digital business.** In the context of digital business, "business moments" — opportunities of short duration or a point in time that sets in motion a series of events involving people, business and things — are increasingly attracting the attention of enterprises. They want to harness data to seize these moments, which will require data integration support. Data integration functionality provided in a "sandbox" to support analytics is of growing interest; this approach enables data to be delivered and manipulated in a physical or virtual manner, for ingestion regardless of where it resides; it also encourages experimentation with, and the building of, new models with which to use data of interest. As pressures for real-time data integration grow, organizations will need to manage a range of data latencies to make data available for use within acceptable service levels and to match the required speed of business.

- **Intensifying pressure for enterprises to modernize and diversify their data integration strategy.** Organizations are increasingly driven to position data integration as a strategic discipline at the heart of their information infrastructure — to ensure it is equipped for comprehensive data capture and delivery, linked to metadata management and data governance support, and applicable to diverse use cases. In addition, implementations need to support multiple types of user experience via tool interfaces that appeal not only to technical practitioners but also to people in business-facing roles, such as business analysts and end users. Offerings that promote collaboration between business and IT participants are becoming important as organizations seek adaptive approaches to achieving data integration capabilities.

- **Requirements to balance cost-effectiveness, incremental functionality, time to value and growing interest in self-service.** Organizations are now seeking cost and deployment options that balance the traditional demands with more modern infrastructure architectures. Buyers conscious of the shift toward distributed "everything" are taking a targeted approach by acquiring only what they need now, while evaluating future data integration needs and how well vendor product roadmaps line up with those needs. Vendors have responded to this development in various ways, such as by varying their pricing structures and deployment options (open-source, cloud and hybrid models), and extending support for end-user functions so that they work with targeted data of interest, especially when requirements aren't well defined. This approach of utilizing limited functionality or mixing different tools from various vendors based on functionality is referred to as "best-fit engineering."

- **Expectations for high-quality customer support and services.** Buyers are demanding superior customer service and support from technology providers. In addition to highly responsive and high-quality technical support for products, they want direct and frequent interactions with sales teams and executives. Buyers also want wide availability of relevant skills — both within a provider's installed base and among system integrator partners — and forums where they can share experiences, lessons and solutions with their peers.

- **Need to blend traditional deployments with modern infrastructure practices.** The need to support operational data consistency, data migration and cloud-related integration is prompting more data integration initiatives than before. The architectural approach of the logical data warehouse (LDW) optimizes the integrated repositories so they can be combined with new data types using virtualization capabilities. Big-data-related initiatives require the use of opportunistic analytics and the exploration of answers to less-well-formed or unexpected business questions.

The distribution of required computing workloads to parallelized processes in Hadoop and alternative nonrelational repositories will continue to advance the ability of data integration tools to interact with big data sources. Another aspect of this is how vendors/suppliers will choose to respond. A large part of the market will continue to pursue physical consolidation of data and the minority of organizations will seek to augment this with virtualization, message queues and data services buses. Synchronization specialists will be a solution for physically distributed but consistent database deployments.

- **Aligning application and data integration infrastructure.** The expansion of vendors' capabilities into application integration provides opportunities to use tools that exploit common areas of both technologies to deliver shared benefits, such as use of CDC tooling that publishes captured changes into message queues. Organizations have begun to pursue data integration and application integration in a synergistic way in order to exploit the intersection of the two disciplines. Aligned application integration and data integration infrastructure — deployed for the full spectrum of customer-facing interactions and a broad range of operational flows — gradually optimize costs and shared competencies, as compared with the pursuit of disparate approaches to similar or common use cases (see "Converging Data and Application Integration: A Step Toward Pervasive Integration Using a Hybrid Integration Platform"). This combined capability of integration patterns is a key component in enabling a hybrid integration platform-inspired infrastructure.

Acronym Key and Glossary Terms

BI	business intelligence
CDC	change data capture
DBMS	database management system
ESB	enterprise service bus
ETL	extraction, transformation and loading
iPaaS	integration platform as a service
MDM	master data management
ODI	Oracle Data Integrator
SaaS	software as a service
SOA	service-oriented architecture
SSIS	SQL Server Integration Services (Microsoft)
TCO	total cost of ownership

Gartner Recommended Reading

Some documents may not be available as part of your current Gartner subscription.

"Toolkit: RFP Template for Data Integration Tools"

"Research Library: Fundamentals for Data Integration Initiatives"

"Does Custom-Coded Data Integration Stack Up to Tools?"

"Comparing Three Self-Service Integration Architectures"

"Align Data Integration and Data Quality to Strengthen IoT Solutions"

"Use iPaaS to Unify Data and Application Integration"

"Innovation Insight for Hybrid Integration Platforms"

"Citizen Integrators Bring Application and Data Integration Into a Common Focus"

"Maverick* Research: Revolutionizing Data Management and Integration With Data Mesh Networks"

"How Markets and Vendors Are Evaluated in Gartner Magic Quadrants"

Evidence

The analysis in this Magic Quadrant is based on information from a number of sources, including:

- Extensive data on functional capabilities, customer base demographics, financial status, pricing and other quantitative attributes gained via an RFI process engaging vendors in this market.

- Interactive briefings in which the vendors provided Gartner with updates on their strategy, market positioning, recent key developments and product roadmaps.

- A web-based survey of the reference customers provided by each vendor, which captured data on usage patterns, levels of satisfaction with major product functionality categories, various nontechnical vendor attributes (such as pricing, product support and overall service delivery), and more. In total, 398 organizations across all major regions provided input on their experiences with vendors and tools in this manner.

- Feedback about tools and vendors captured during conversations with users of Gartner's client inquiry service.

- Market share estimates developed by Gartner's Technology and Service Provider research unit.

Note 1 Detailed Components of the Evaluation Conditions

Gartner has defined several classes of functional capability that vendors of data integration tools provide in order to deliver optimal value to organizations in support of a full range of data integration scenarios:

- **Connectivity/adapter capabilities (data source and target support).** The ability to interact with a range of different types of data structure, including:

  - Relational databases

  - Legacy and nonrelational databases

  - Various file formats

  - XML

  - Packaged applications, such as those for customer relationship management (CRM) and supply chain management

  - SaaS and cloud-based applications and sources

  - Industry-standard message formats, such as electronic data interchange (EDI), Health Level Seven International (HL7) and Society for Worldwide Interbank Financial Telecommunication (SWIFT)

  - Parallel distributed processing environments such as Hadoop Distributed File System (HDFS) and other nonrelational-type repositories such as graph, table-style, document store and key-value DBMSs

  - Message queues, including those provided by application integration middleware products and standards-based products (such as Java Message Service)

  - Data types of a less-structured nature, such as those associated with social media, web clickstreams, email, websites, office productivity tools and content

  - Emergent sources, such as data on in-memory repositories, mobile platforms and spatial applications

  - Screen-scraping and/or user interaction simulations (for example, scripts to interact with the web, 3270 or VT100 terminals, and others)

  Data integration tools must support different modes of interaction with this range of data structure types, including:

  - Bulk/batch acquisition and delivery

  - Granular trickle-feed acquisition and delivery

  - Change data capture (CDC) — the ability to identify and extract modified data

  - Event-based acquisition (time-based, data-value-based or links to application integration tools to interact with message request/reply, publish/subscribe and routing)

- **Data delivery capabilities.** The ability to provide data to consuming applications, processes and databases in a variety of modes, including:

  - Physical bulk/batch data movement between data repositories, such as processes for ETL or extraction, loading and transformation (ELT)

- Data virtualization

- Message-oriented encapsulation and movement of data (via linkage with application integration tool capability)

- Data synchronization when distributed datasets must resolve data collisions resulting from distinct changes in disparate copies of data to retain data consistency

- Replication of data between homogeneous or heterogeneous DBMSs and schemas

- Migration of data across versions of data repositories (such as databases, file systems, and so on) and applications (resolving logical differences to achieve physical migration)

In addition, support for the delivery of data across the range of latency requirements is important, including:

- Scheduled batch delivery

- Streaming/near-real-time delivery

- Event-driven delivery of data based on identification of a relevant event

- **Data transformation capabilities.** Built-in capabilities for achieving data transformation operations of varying complexity, including:

  - Basic transformations, such as data-type conversions, string manipulations and simple calculations

  - Transformations of intermediate complexity, such as look-up and replace operations, aggregations, summarizations, integrated time series, deterministic matching and the management of slowly changing dimensions

  - Complex transformations, such as sophisticated parsing operations on free-form text, rich media and patterns/events in big data

In addition, the tools must provide facilities for developing custom transformations and extending packaged transformations.

- **Metadata and data modeling support.** As the increasingly important heart of data integration capabilities, metadata management and data modeling requirements include:

  - Automated discovery and acquisition of metadata from data sources, applications and other tools

  - Discernment of relationships between data models and business process models

  - Data model creation and maintenance

  - Physical-to-logical model mapping and rationalization

  - Ability to define model-to-model relationships via graphical attribute-level mapping

  - Lineage and impact analysis reporting, in graphical and tabular formats

- An open metadata repository, with the ability to share metadata bidirectionally with other tools

- Automated synchronization of metadata across multiple instances of the tools

- Ability to extend the metadata repository with customer-defined metadata attributes and relationships

- Documentation of project/program delivery definitions and design principles in support of requirements definition activities

- A business analyst/end-user interface to view and work with metadata

- **Design and development environment capabilities.** Facilities for enabling the specification and construction of data integration processes, including:

  - Graphical representation of repository objects, data models and data flows

  - Management of the development process workflow, addressing requirements such as approvals and promotions

  - Granular, role-based and developer-based security

  - Team-based development capabilities, such as version control and collaboration

  - Functionality to support reuse across developers and projects, and to facilitate the identification of redundancies

  - A common or shared user interface for design and development (of diverse data delivery styles, data integration and data quality operations, cloud and on-premises environments, and so on)

  - A business analyst/end-user interface to specify and manage mapping and transformation logic through the use of end-user functionality for data integration/preparation

  - Support for testing and debugging

- **Information governance support capabilities (via interoperation with data quality, profiling and mining capabilities with the vendor's or a third party's tools).** Mechanisms to work with related capabilities to help with the understanding and assurance of data quality over time, including interoperability with:

  - Data profiling tools (profiling and monitoring the conditions of data quality)

  - Data mining tools (relationship discovery)

  - Data quality tools (supporting data quality improvements)

  - In-line scoring and evaluation of data moving through the processes

- **Deployment options and runtime platform capabilities.** Breadth of support for the hardware and operating systems on which data integration processes may be deployed, and the choices of delivery model — specifically:

- Mainframe environments, such as IBM z/OS and z/Linux

- Midrange environments, such as IBM i or Hewlett Packard Enterprise (HPE) NonStop

- Unix-based environments

- Windows environments

- Linux environments

- On-premises (at the customer site) installation and deployment of software

- Hosted off-premises software deployment (dedicated, single-tenant implementation)

- Integration platform as a service (iPaaS), consumed by the customer completely "as a service" — the vendor provides cloud infrastructure; the customer does not install or administer the software

- Cloud deployment support (requires organizations to deploy software in a cloud infrastructure); importantly, the ability to design once but deploy across multiple or even hybrid/mixed environments, on-premises, in the cloud, or both

- In-memory computing environment

- Server virtualization (support for shared, virtualized implementations)

- Parallel distributed processing, such as Apache Hadoop, MapReduce, or leveraging Apache Spark or Hadoop YARN (Yet Another Resource Negotiator)

- **Operations and administration capabilities.** Facilities for enabling adequate ongoing support, management, monitoring and control of the data integration processes implemented by the tools, such as:

  - Error-handling functionality, both predefined and customizable

  - Monitoring and control of runtime processes, both via functionality in the tools and through interoperability with other IT operations technologies

  - Collection of runtime statistics to determine use and efficiency, as well as an application-style interface for visualization and evaluation

  - Security controls, for both data in-flight and administrator processes

  - A runtime architecture that ensures performance and scalability

- **Architecture and integration capabilities.** The degree of commonality, consistency and interoperability between the various components of the data integration toolset, including:

  - A minimal number of products (ideally one) supporting all data delivery modes

  - Common metadata (a single repository) and/or the ability to share metadata across all components and data delivery modes

  - A common design environment to support all data delivery modes

- The ability to switch seamlessly and transparently between delivery modes (bulk/batch versus granular real-time versus federation) with minimal rework

- Interoperability with other integration tools and applications, via certified interfaces, robust APIs and links to messaging support

- Efficient support for all data delivery modes, regardless of runtime architecture type (centralized server engine versus distributed runtime)

- The ability to execute data integration in cloud and on-premises environments, as appropriate, where developed artifacts can be interchanged, reused and deployed across both environments with minimal rework

- **Service enablement capabilities.** As acceptance of data service concepts continues to grow, so data integration tools must exhibit service-oriented characteristics and provide support for SOA, such as:

  - The ability to deploy all aspects of runtime functionality as data services (for example, deployed functionality can be called via a web services interface)

  - Management of publication and testing of data services

  - Interaction with service repositories and registries

  - Service enablement of development and administration environments, so that external tools and applications can dynamically modify and control the runtime behavior of the tools

---

Evaluation Criteria Definitions

Ability to Execute

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

Overall Viability: Viability includes an assessment of the overall organization's financial health, the financial and practical success of the business unit, and the likelihood that the individual business unit will continue investing in the product, will continue offering the product and will 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/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 initiatives, 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, service-level agreements and so on.

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 to 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 website, advertising, customer programs and positioning statements.

Sales Strategy: The strategy for selling products 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 sets 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 vertical markets.

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.

GARTNER HEADQUARTERS

Corporate Headquarters
56 Top Gallant Road
Stamford, CT 06902-7700
USA
+1 203 964 0096

Regional Headquarters
AUSTRALIA
BRAZIL
JAPAN
UNITED KINGDOM

For a complete list of worldwide locations,
visit http://www.gartner.com/technology/about.jsp