Magic Quadrant for Data Integration Tools

Published: 1 August 2019   **ID:** G00369547

Analyst(s): Ehtisham Zaidi, Eric Thoo, Nick Heudecker

The data integration tool market is resurging as new requirements for hybrid/intercloud integration, active metadata and augmented data management force a rethink of existing practices. This assessment of 16 vendors will help data and analytics leaders make the best choice for their organization.

Strategic Planning Assumptions

By 2021, more than 80% of organizations will use more than one data delivery style to execute their data integration use cases.

By 2022, organizations utilizing active metadata to dynamically connect, optimize and automate data integration processes will reduce time to data delivery by 30%.

By 2022, manual data integration tasks (including recognition of performance and optimization issues across multiple environments) will be reduced by 45% through the addition of ML and automated service-level management.

By 2023, improved location-agnostic semantics in data integration tools will reduce design, deployment and administrative costs by 40%.

Market Definition/Description

The discipline of data integration comprises the architectural techniques, practices and tools that ingest, transform, combine and provision data across the spectrum of data types. This integration takes place in the enterprise and beyond — across partners as well as third-party data sources and use cases — to meet the data consumption requirements of all applications and business processes. This is inclusive of any technology that supports data integration requirements regardless of current market nomenclature (e.g., data ingestion, data transformation, data replication, messaging, data synchronization, data virtualization, stream data integration and many more).

The market for data integration tools consists of vendors that offer software products to enable the construction and implementation of data access and delivery infrastructure for a variety of integration use-case scenarios.

Example integration usage scenarios include:

- **Data integration and delivery for optimized analytics —** Accessing, queueing or extracting data from operational systems; transforming and merging that data virtually or physically; and delivering it through an integrated approach for optimized and repeatable analytics (such as those delivered via the data warehouse) and data science purposes.

- **Sourcing and delivery of master data in support of master data management (MDM) —** Enabling the connectivity and integration of data representing critical business entities, such as customers, products and employees. Data integration tools can be used to integrate, consolidate and synchronize master data related to critical business processes.

- **Data consistency between operational applications —** Ensuring database-level consistency across applications, on both an internal and an interenterprise basis. This could involve synchronizing data structures for on-premises applications or cloud-resident data sources in SaaS, and for bidirectional or unidirectional consistency.

- **Interenterprise data acquisition and sharing —** For providing data to, and receiving data from, external trading partners (customers, suppliers, business partners and others). Some interenterprise data sharing requirements involve on-premises or cloud-based environments, or a combination of both. Data integration tools may be used to support data acquisition, sharing and collaborations across applications, which often consist of the common types of data access, transformation and movement components that are also found in other use cases.

- **Data services orchestration —** Deploying all aspects of runtime data integration functionality as data services (for example, deployed functionality can be called via a web services interface).

- **Data migration and consolidation —** Addressing the data movement and transformation needs of data migration and consolidation, such as the replacement of legacy applications, databases or both. Although most are often addressed through custom coding of conversion programs, data integration tools can provide significant support to enterprises undertaking large-scale data migration projects (often due to mergers and acquisitions, modernization or consolidation). However, it should be clear that data integration tools alone do not solve all data migration challenges.

- **Support for data governance and management of data assets —** Increasingly, data integration tools are expected to collect, audit, govern, share and monitor data regarding the deployed data integration service and processes in the organization. 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.

Data integration tools may display characteristics that combine aspects of the individual use-case scenarios listed above. Technologies in this market are required to execute many of the core functions of data integration, which can be applied to any of the above scenarios. (For a detailed list and analysis of all evaluation components and functions, see Note 1.)

Some examples of new and differentiating functionality or characteristics include:

- **Interoperating with application integration technology in a single solution architecture —** This is now go far beyond supporting extraction, transformation and loading (ETL) processes. It can include layered data services such as change data capture (CDC), which can populate data queues, reading message services and accepting streaming data, and extend to the point of provisioning these processes across an enterprise service bus.

- **Supporting data integration across hybrid cloud and intercloud environments —** Hybrid cloud means data is spread across on-premises and cloud ecosystems, while with intercloud it is spread across different cloud infrastructure providers (see "Are You Ready for Multicloud and Intercloud Data Management?"). This is an urgent requirement, as organizations now expect their data integration tools to support this 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.

  This increasingly requires data integration tools to utilize both active and passive metadata (through analysis on this metadata) to recommend and, in some cases, even automate data integration design and infrastructure deployment (see Note 2 for the distinction between active and passive metadata). This assists data and analytics leaders with designing more-flexible data management architectures that account for this hybrid integration ecosystem, and reduces the need for unnecessary data replication or movement to support new data and analytics use cases.

- **Enabling data services for use in broader architecture approaches —** An example is participating in hybrid integration platforms (HIPs). Or, something as simple as enabling a semantic layer, or even historian software queues in IoT and edge devices (historian software is data that collects sensor data as a local cache in IoT environments).

- **Supporting the delivery of data to, and the access of data from, a wide variety of data stores, repositories and data management tiers in application deployments —** This includes but is not limited to: distributed data management solutions, analytic data management repositories, data lakes and platforms typically associated with nonrelational (formerly known as NoSQL platforms) data integration initiatives, such as Hadoop, nonrelational databases and cloud-based data stores.

- **Nonrelational DBMS integration —** This poses data integration challenges but also provides opportunities to assist in the application of schemas at data read time, if needed, and deliver data to business users, processes or applications — or to use data iteratively. Data integration tools must provide connectivity options to integrate different types of nonrelational DBMSs, such as key value stores, graph databases and document stores, among others.

  Most data integration tools are slow to roll out interfacing, integration and orchestration functions with new and other popular nonrelational DBMSs, making this a differentiating capability.

- **IoT/OT data convergence —** Increasingly, the differing structure, latency and throughput requirements of IoT or machine data is introducing new integration requirements. This data is sometimes integrated through stream data integration capabilities, and at other times stored in

time-series databases. Such integration requirements are now expected to be addressed by modern data integration tools for IT/OT data convergence.

- **Self-service data integration —** Finally, there is an increasing expectation of organizations to allow business users or citizen integrators to be able to integrate "their own data" through data preparation tools and techniques. The notion of data management being able to govern and control the flow in a synergistic manner through the data integration tool is a challenge that data integration tool vendors are expected to solve.

In recent years, significant submarkets have emerged in parallel to the main market offerings. These represent a renewed focus on either vision or execution, but do not address *all* data integration and delivery requirements. There are tools that focus on innovative solutions and modern data delivery styles, such as data virtualization, data preparation or stream data integration, among others. These allow organizations to include these new capabilities to support their new data integration requirements. Such requirements include a focus on data virtualization, stream data integration and data preparation, but also specific delivery to support management of data lakes (see "Market Guide for Data Preparation Tools," "Adopt Stream Data Integration to Meet Your Real-Time Data Integration and Analytics Requirements" and "Market Guide for Data Virtualization").

Magic Quadrant

Figure 1. Magic Quadrant for Data Integration Tools

Source: Gartner (August 2019)

Vendor Strengths and Cautions

Actian

Actian is based in Palo Alto, California, U.S. and, including embedded/OEM deployments, has more than 8,000 data integration tool customers. It offers the DataConnect product set, which includes integration tools, technology and services for on-premises deployment through virtual private cloud, multitenant integration platform as a service (iPaaS) and embedded data management.

Note: Actian was jointly acquired by HCL Technologies and Sumeru Equity Partners in July 2018.

Strengths

- **Relevance of targeted capabilities.** Actian continues to leverage its lightweight and small-footprint tool in order to drive long-lasting revenue. As an easily embeddable data integration tool, DataConnect stabilizes quickly and almost bypasses reviews considering its replacement — often for years. Reference customers cited the vendor's ease of use, reliability and ability to handle complex workloads.

- **Opportunity to improve market awareness and mind share.** Actian has an established variety of data integration and data management and analytics tools, often with separate go-to-market approaches. As an acquirer, HCL Technologies should provide substantial reseller, system integrator (SI) and OEM partnership opportunities through leveraging its global ecosystem.

- **Processing optimization.** Actian's data integration tool maintains in-line statistics for data that crosses the integration platform. This continues to be a strength — combining capacity, utilization, data statistics, data profiling and many other components to create a combination of operational alerts for system health and regarding changes in the data, for users and developers alike.

Cautions

- **Acquisition uncertainty.** We consider the acquisition by HCL Technologies to be one that requires a careful balance between maintaining the current embedded solutions business, and a professional services organization subsuming the tool completely. Actian has assured Gartner that it will continue to operate as a separate legal entity after acquisition. Existing customers and prospects can assume with some confidence that, even if brand dilution does occur, the embedded solutions should have a long technology and support life.

- **Lacks role-based delivery.** DataConnect is focused on traditional data integration experts, who deliver integration as part of an application development or in the capacity of supporting data engineering. However, Actian's roadmap involves introducing an integrated design studio to support varying roles (including citizen integrators), through the inclusion of guided development workflows, templates, community knowledge, and issue resolution.

- **Installed base is primarily for bulk/batch integration.** Actian's primary data integration style remains bulk/batch-oriented. While most organizations begin with bulk/batch-based data

integration, this could be a limiting factor if they need to combine bulk/batch with other modern data integration styles (such as data virtualization, for example). The vendor needs to expand the breadth of its data delivery methods to include data preparation, data virtualization and other modern data integration styles, in order to expand into new delivery channels.

Adeptia

Based in Chicago, Illinois, U.S., Adeptia offers Adeptia Connect as its data integration product. The vendor's customer base for the data integration tool market is more than 1,400 organizations.

Strengths

- **Integrated product and flexible delivery.** Adeptia offers its data integration technology alongside other integration capabilities that provide B2B integration needs for data, process and partner community management — all in a single product. Support for distributed runtime processing using SecureBridge Agent provides flexibility for hybrid deployment models.

- **Business role enablement.** Expanded capabilities for large-file data ingestion, role-based security and personalized interfaces extend the applicability of Adeptia for digital business ecosystems. A hub-based architecture, for monitoring data flow, managing configuration and administration, and the use of ML to aid data mapping, seeks to simplify deployments and empower business roles.

- **Flexible pricing and time to value.** Reference customers viewed Adeptia's tools as attractively priced and delivering good value, as they appreciate the tight integration of the underlying components and the ability to support rapid implementation. Adeptia's increasing transition toward subscription-based pricing, based on tiered editions and feature sets, aims to simplify procurement.

Cautions

- **Skills and market coverage.** The availability of implementers and guidance for best practices are concerns cited by customers seeking a wider array of skilled resources, as their implementation complexity and requirements grow. Adeptia continues to focus its products toward ease of use, while building out its partner network and deployment coverage in the market.

- **Degree of metadata support.** Reference customers identified Adeptia's metadata management as an area of relative weakness when enabling reusability across use cases. Customers are increasingly looking for comprehensive functionality and a vision for these requirements, to address the escalating number and variety of datasets as well as distributed data architectures.

- **Technical support and guidance.** Reference customers cited areas of improvement needed for product technical support and documentation. They expressed a desire for a more mature user community for improved access to implementation guidance and practices.

Denodo

Based in Palo Alto, California, U.S., Denodo offers Denodo Platform as its data integration offering. The vendor's customer base for this product is around 700 organizations.

Strengths

- **Market recognition and alignment to evolving needs.** Denodo is frequently considered by buyers evaluating data virtualization technologies. It aligns to diverse use cases such as logical data warehouse/data lake, data service marketplaces and registry-style MDM. Its location-agnostic platform for building capabilities can be seamlessly executed, containerized and reused on many established as well as emerging data and application infrastructures.

- **Targeted offering and robust performance.** Denodo has an established tenure for data virtualization capabilities that work with a diverse range of underlying data sources and data types, both on-premises and in the cloud. Reference customers favored the vendor's technology for its ability to connect to diverse sources and federate data, enable logical abstraction, and support data preparation and data cataloging.

- **Leveraging implementation and technology partners.** Software vendors license or bundle Denodo's functionality as part of their products for analytics, big data and vertical solution use cases. The vendor is also available on AWS Marketplace, Azure Marketplace and Google Cloud Platform. Its partner network encompasses global and regional SIs as well as software vendors, including Infosys, HCL, Deloitte, Wipro, TCS, AWS, Microsoft, Cloudera, Qlik, Snowflake and Tableau.

Cautions

- **Versatility challenges.** While it is well established that data virtualization is a dedicated focus of Denodo, there is a limited awareness in the market of how its data virtualization capability interoperates with other data delivery styles, which sometimes presents competitive challenges. The vendor is addressing this by publishing and promoting numerous case studies and marketing content detailing how organizations can combine data virtualization with other data delivery styles, such as bulk/batch-based ETL.

- **Pricing and negotiation concerns.** Denodo's pricing and contract negotiation flexibility are reported by some existing and prospective customers as areas needing improvement. A small but notable number of prospects and customers who favor the vendor's product strength have nevertheless expressed concerns about high prices in conversations with Gartner.

- **Standardized practices required.** Some Denodo reference customers want better guidance about implementation, including improved technical documentation. Deployments in increasingly complex scenarios are raising customers' expectations for more extensive linkage and enablement of data management infrastructure.

Hitachi Vantara

Based in Santa Clara, California, U.S., Hitachi Vantara offers Pentaho Data Integration, Hitachi Streaming Data Platform (HSDP) and Hitachi Data Instance Director (HDID). The vendor's customer base for this product set is more than 2,000 organizations.

Strengths

- **Expanding portfolio relevance.** Hitachi Vantara continues to evolve its data integration offerings by supporting the data needs arising from IT/OT, edge computing, and integration of stream data and IoT. As part of the Hitachi portfolio, data integration tooling extends toward IT/OT convergence, digitalization, and enhanced analytics experience, metadata collection and runtime/deployment optimization, making the vendor well positioned to tackle IoT integration scenarios.

- **Increased support for data management.** Hitachi Vantara has data sharing support for multienterprise and multicloud integration, and increased synergy of data integration tooling with its suite of data management technologies. Through this it enables a broader scope of integration for operations requirements, and capitalizes on digital business opportunities in conjunction with its content intelligence technology.

- **Growing resources and market reach.** Hitachi Vantara's global reach continues to widen the availability of its data integration products, with financial resources and sales incentives for further product enhancement and market access.

Cautions

- **Mind share and business clarity.** There is a relative lack of mainstream understanding and recognition of Hitachi Vantara's data integration tooling among many organizations specifically evaluating data integration tools. Concerning go-to-market activities, buyers indicated in our reference survey a need for clarity around the vendor's technological/commercial strategy and guidance for benefiting from its enlarged business model and offerings.

- **Integrated use of portfolio.** Reference customers identified a need for improvements to enable integrated usage across different data integration tools — for example, between Pentaho Data Integration and HDID. This would enable smoother adoption for different use cases.

- **Deployment and diagnostic guidance.** Hitachi Vantara reference customers want better guidance about implementation, including improved ease of troubleshooting and diagnostic support to navigate toolset deployment/operation.

IBM

Based in Armonk, New York, U.S., IBM offers the following data integration products: IBM InfoSphere Information Server, IBM InfoSphere Classic Federation Server, IBM InfoSphere Data Replication, IBM App Connect, IBM Streams and IBM Data Refinery. The vendor's customer base for this product set is more than 11,000 organizations.

Strengths

- **Depth of integration offering.** Reference customers highlighted the completeness of IBM's holistic data integration suite, including its rich functionality, variety of prebuilt functions and connectors, and its overall performance.

- **Diverse data integration delivery styles.** Reference customers use IBM's products for traditional data delivery styles (data replication, batch processing), as well as more complex data delivery styles (including data synchronization and stream data integration). They praised IBM's data integration tool portfolio for its ability to deliver complex data integration requirements that demand combinations of traditional and modern data integration styles, such as data replication, data virtualization and stream data integration for real-time analytics.

- **Brand awareness and market presence.** IBM's size and the global coverage of its business systems, infrastructure platforms and analytics solutions enable it to draw on a huge customer base and a wide product distribution model for positioning its data integration tools. Broad usage of IBM technologies within its customer base has driven the wide availability of implementation service providers and approaches to solving complex integration challenges.

Cautions

- **Architecture and upgrade complexity.** Several survey reference customers cited challenges upgrading some or all of their IBM data integration footprint. The need for improved implementation services and ease of deployment was also highlighted. IBM has now introduced a Rapid Migration offering and been expanding its Kubernetes/OpenShift container deployment options for accelerated version migrations, deployments and upgrades.

- **Cost model.** Across all the vendor reference surveying, prospective customers pointed to difficulty in understanding IBM's licensing and pricing methods. Almost half of the reference survey respondents who did not select IBM data integration tools indicated that they avoided it because of its complex pricing model or overall perceived high cost. The vendor continues to rationalize its pricing across products and platforms.

- **Modern deployment and usage patterns.** Some reference customers scored IBM below average for its data virtualization delivery style, as well as for its ability to support multicloud integration patterns.

Informatica

Headquartered in Redwood City, California, U.S., Informatica offers a series of data integration products as part of its Intelligent Data Platform. These products are PowerCenter, PowerExchange, Data Replication, B2B Data Transformation, B2B Data Exchange, Data Integration Hub, Data Services, Intelligent Cloud Services, Big Data Management, Big Data Integration Hub, Big Data Streaming, Enterprise Data Catalog, Enterprise Data Preparation and Edge Data Streaming. The vendor's customer base for this product set is more than 10,000 organizations.

Gartner, Inc. | G00369547

Strengths

- **Augmented data integration through metadata-based AI/ML.** Informatica continues to lead the push to augmented data management, with strategic investments in CLAIRE — its AI-powered engine — for metadata-driven integration. Embedding CLAIRE in the product portfolio provides users with capabilities in integration automation, including automated metadata discovery, cataloging, lineage, impact analysis, AI-based recommendations for integration deployment and performance optimization of integration flows. Reference customers also scored it well above average for ML-enabled augmented data integration and active metadata analysis.

- **Hybrid/multicloud integration.** Informatica has responded well to market demand for application and data integration convergence, serverless big data management, and hybrid/multicloud requirements through continued investment in its microservices-based and API-driven iPaaS, Informatica Intelligent Cloud Service (IICS). It provides services including application, data and B2B integration; API management; mass ingestion; digital integration hub; MDM; data quality; and governance. The entire suite can be managed, monitored and operated with the platform's hybrid unified operational insights features. Reference customers praised the ability of IICS to bridge data silos across on-premises and multicloud ecosystems, and gave it strong scores for delivery of integration capabilities as cloud services, hybrid integration and multicloud integration support.

- **Expansion of market presence and mind share.** Informatica continues to execute strongly in the modern data integration tool market, appearing more frequently than any other vendor in competitive situations and contract negotiations analyzed by Gartner. Reference customers praised its global partner ecosystem (comprising over 500 SI, consulting and strategy organizations), which eases skills sourcing and helps with complex deployments. In addition, Informatica has strategic partnerships with AWS, Microsoft Azure, Google Cloud Platform, Cloudera, Databricks and Tableau.

Cautions

- **Differing maturity within product set.** Informatica markets IICS as an iPaaS for both on-premises and cloud-based data integration needs, having capabilities that are interoperable with its well-established data integration tool, PowerCenter. However, a few respondents commented that, while PowerCenter is mature for complex data integration scenarios (for on-premises data sources), IICS is better suited for multicloud and hybrid data integration scenarios. Customers with on-premises implementations looking to leverage the benefits of the cloud should work with Informatica to develop an architecture based on the HIP framework, to take full advantage of both on-premises and multicloud, hybrid investments.

- **Overall cost and pricing complexity.** Informatica has successfully repositioned itself as a subscription-license-based business and invested in pricing model rationalization. It now offers PAYG, BYOL and serverless metered pricing options, with tiered connectivity options. Despite this, reference customers who selected another vendor cited confusion with Informatica's license models and overall perceived high cost as the main reasons for avoiding it.

- **Some calls for portfolio complexity rationalization.** A small but notable number of reference customers expressed a need for tool portfolio rationalization into modular packages aligned with data integration tool use cases. Informatica has two such packages (Enterprise Data Preparation, Enterprise Streaming and Ingestion) that encourage a start-small approach, but ongoing confusion is contributing to slow sales cycles and redundant purchases.

Information Builders

Based in New York City, New York, U.S., Information Builders offers the Omni-Gen data integration platform. The platform comprises a central platform plus additional tools including iWay Service Manager, iWay DataMigrator and iWay Universal Adapter Suite. Components may be bought separately. The vendor's customer base for this product set is more than 950 organizations.

Strengths

- **Relevance for diverse applicable uses.** Information Builders continues to focus on evolving an integrated environment for a data integration capability that can operate with a combination of data preparation, MDM and prepackaged vertical solutions. This aligns well with demand for data integration activity support so that competency teams can seamlessly support data management and analytics requirements.

- **Customer relationships.** Reference customers reported positive overall experiences with Information Builders, both before buying and after implementation. Selection of this vendor's data integration tools is often influenced by an existing relationship and use of other Information Builders products.

- **Synergy with data management infrastructure.** The breadth of Information Builders' product portfolio and experience in deployments of various data integration styles align it well with the contemporary trends of related markets. Its Omni-Gen platform can be deployed to combine data integration with the adjacent technologies of data quality and MDM.

Cautions

- **Concerns over business evolution.** Recent organizational changes in leadership have generated some concerns for clarity among prospective customers, particularly around Information Builders' strategy and focus toward expanding market presence and use cases in data integration. However, this also presents opportunities for Information Builders to articulate and build on a tenure of providing solutions enabled by data integration.

- **Implementation support and time to value.** Reference customers said they desire improvements to guidance for implementing Information Builders products. They want wider availability of skills and community support to ease expanding use of tools and faster time to value in deployments.

- **Upgrade complexity and metadata support.** Reference customers cited difficulties with version upgrades — specifically with the technical complexity of migrating between major releases. They cited metadata management and modeling functionalities as weaknesses relative to the whole Magic Quadrant group.

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. The vendor also includes data integration as part of its Azure Data Factory (ADF) cloud data integration service. Microsoft SQL Server deployments include SSIS for data integration. Microsoft does not report a specific customer count for its data integration tools.

Strengths

▪ **Usability and TCO.** When asked why they chose Microsoft, reference customers cited its overall low TCO, speed of implementation, ease of use and ability to integrate with other Microsoft SQL Server capabilities/Azure data services.

▪ **Hybrid deployment capabilities.** The tight coupling of SSIS for on-premises data integration with data stored in various Azure services (and data stores) is paying dividends for Microsoft. It had one of the highest numbers of reference customers stating that hybrid deployments are in production or development.

▪ **Service and support.** Reference customers gave Microsoft nearly perfect scores for its overall service and support, and high scores for overall experience with the vendor as well as value for money. Additionally, Microsoft scored well above the average for end-user training as well as administration and management capabilities.

Cautions

▪ **Perception of platform versatility.** There is a perception that Microsoft's integration tooling is useful only in Microsoft-centric environments. While the vendor has begun to address this with SSIS on Linux, its Azure-only development roadmap will become less relevant in a multicloud and intercloud enterprise data estate.

▪ **Pipeline complexity and support for complex data replication.** Reference customers cited challenges implementing complex data integration pipelines with ADF — particularly when they vary or combine data delivery capabilities. Workarounds take additional build time and introduce complexity. A significant number of reference customers reported that they had to partner with competing data integration tools (particularly for data replication) in order to migrate data from their on-premises databases to the Azure ecosystem. Microsoft is addressing these problems through partnerships with other data migration and replication tools, but should continue to enhance ADF to support such complex data synchronization.

▪ **Metadata and governance.** Microsoft received low reference customer scores for metadata management and data modeling, as well as support for active metadata. Survey respondents also scored it near the bottom of this group for its data governance and data quality capabilities. While itself not a challenge for customers using Microsoft's tools only for data integration requirements, this could certainly challenge customers looking for broader data management capabilities (including data quality, data governance and metadata management) for data and analytics use cases.

Oracle

Based in Redwood Shores, California, U.S., Oracle offers a number of data integration products. They are Oracle Data Integration Platform Cloud Service (includes Oracle GoldenGate Cloud Service and Oracle Data Integrator Cloud Service), Oracle GoldenGate (for data replication, synchronization and big data streaming), Oracle Data Integrator (ODI — for bulk/batch integration), Oracle Big Data SQL (for data virtualization scenarios), Oracle Service Bus (for messaging scenarios) and Oracle Integration Cloud Service (iPaaS). The vendor's customer base for this product set is more than 12,000 organizations.

Strengths

- **Extensive stream data integration and replication scenarios.** Oracle GoldenGate and GoldenGate Cloud Service continue to appear very frequently for stream data integration and data replication scenarios on Gartner client tool selection shortlists. Reference customers scored as particularly strong the data synchronization capabilities of GoldenGate, which Oracle has continued to make strategic investments in. These include providing a microservices-based architecture (through full cutover to distributed HTTPS and RESTful APIs for all GoldenGate features), full Spark push-down processing, Kafka support (both as a source and target), and a serverless metered pricing option. It has also added ML capabilities for augmenting certain data integration tasks for nontechnical users.

- **Well positioned for HIP enablement.** Oracle supports a strong portfolio of tools and services that together could provision a HIP. Along with its data integration tools, it provides API management (Oracle API Platform Cloud Service), iPaaS (Oracle Integration Cloud Service), managed file transfer (Oracle Managed File Transfer Cloud Service), data preparation (Oracle Analytics Cloud Service) and IoT integration. In addition, it offers multiple deployment options including cloud and hybrid to its customers for its integration tool portfolio.

- **Global partner network.** Oracle uses its large partner network to assist customers in finding skilled technical consultants and service providers for implementing its data integration technologies. Reference customers reported ease of finding skilled talent as one of the top three reasons for selecting Oracle.

Cautions

- **Concerns with pricing and licensing models.** Oracle has introduced a wide range of alternative pricing options for its data integration tools (including by processor, term, named user, subscription, PAYG and Cloud Credits). However, some reference customers cited "overall pricing" as one of the main reasons for not choosing the vendor in competitive situations. Oracle received below-average scores for its price method in the reference customer survey.

- **Inconsistent support and solution delivery.** A small but notable number of Gartner clients and Oracle reference customers reported that, while Oracle's data integration solutions are architecturally sound, in some cases they are prone to uneven delivery. Some survey respondents reported that it took a long time for the vendor to fix open issues and bugs, and cited a need to make online documentation more accurate for new releases. The vendor was scored below average for service and support.

- **Lack of awareness about modern integration capabilities among target base.** Existing and new customers seem mostly unaware of Oracle's new capabilities supporting modern data integration. The vast majority of reference customers reported that they used Oracle's data integration technology for bulk/batch or data replication scenarios, with very few using it for modern data integration styles (e.g., data virtualization) or modern integration initiatives (e.g., hybrid integration and active metadata support for IoT data integration). This demonstrates that the vendor's capabilities in these areas are currently unknown or unproven among the customer base.

Qlik (Attunity)

Based in Burlington, Massachusetts, U.S., Attunity (a division of Qlik) offers Attunity Replicate, Attunity Compose, Attunity Visibility and Attunity Enterprise Manager. The vendor's customer base for this product set is more than 2,500 organizations globally.

Note: Qlik completed its acquisition of Attunity in May 2019.

Strengths

- **Proven in data replication scenarios.** Attunity continues to be evaluated and selected by clients in the majority of competitive situations for targeted data replication needs. Reference customers gave positive feedback for the vendor's robust CDC-based replication capabilities, and Attunity appeared very frequently in competitive evaluations for data replication across the whole reference survey.

- **Ease of use.** Reference customers repeatedly cited Attunity's ease of use — including ease of installation and configuration, automated code generation, and nonintrusive implementation — as the top reasons for selecting the vendor in competitive situations.

- **Strong OEM support and partner momentum.** Attunity has amassed a strong partner network for cloud data replication and migration (through OEM partnerships with global cloud infrastructure providers such as Amazon Web Services [AWS] and Microsoft Azure). It has likewise done so for data lakes, big data integration and for data warehouse automation and beyond (through partnerships with vendors including IBM, Oracle, SAP and Teradata). Attunity has also developed offerings with upcoming and popular data management vendors, including Snowflake, Databricks and Confluent. It has also expanded its partner network to include global SIs (e.g., Accenture, Cognizant and Infosys) and global resellers (e.g., Hewlett Packard Enterprise and Teradata) to help minimize concerns about talent availability.

Cautions

- **Limited traction beyond core scenarios.** Reference customers cited either current or intended use of Attunity for data replication, data warehouse automation and data lake population, with data integration styles other than data replication/synchronization yet to find traction. The vendor lacks proven deployments in bulk/batch (ETL) and data virtualization scenarios, which

could be an issue for customers looking to use it to combine and interoperate between data delivery styles (e.g., data replication with data virtualization).

- **Licensing costs.** Some reference customers highlighted issues with Attunity's licensing model — specifically highlighting its "per core" licensing, which includes data sources in calculating costs — as expensive. They scored the vendor below average for its pricing method. To alleviate some of these concerns, Attunity has transitioned from a perpetual to a subscription license model based on customer demand.

- **Acquisition uncertainty.** Qlik's acquisition casts some doubt on Attunity's continued existence as an independent offering with an independent development roadmap. However, Qlik and Attunity have both told Gartner that Qlik will continue to sell and support Attunity as an independent and stand-alone data integration solution focused on application- and vendor-neutral data integration use cases.

SAP

SAP is based in Walldorf, Germany. It offers the following data integration products: SAP Data Services, SAP Replication Server, SAP Landscape Transformation Replication Server, SAP Data Hub, SAP HANA and SAP Cloud Platform Integration Suite. The SAP HANA platform includes SAP HANA Smart Data Integration (for bulk/batch and low-latency data movement), SAP HANA Smart Data Access (for data virtualization) and SAP HANA Streaming Analytics (for stream data integration). The vendor's customer base for this product set is more than 60,000 organizations.

Strengths

- **Broad usage and functionality.** SAP's data integration products are regularly deployed for all the data integration use cases defined by Gartner and across the breadth of functionality of SAP's portfolio. This is driven by a vision emphasizing integrated sharing and governance of data, a data hub architecture, and business role empowerment with data preparation and data science solutions.

- **Innovation focused on enabling a digital business.** With SAP's increasing use of ML, a common environment is enabled for discovering and collaborating on reusable artifacts of active-metadata-driven integration flows and orchestration of data delivery capabilities. SAP's roadmap sets out its "Intelligent Enterprise" vision that readily intertwines data integration with other data management technologies, data governance, HIP and various enabling components, including the serverless architectures required for equipping digital business technology platforms.

- **Vast market presence and ecosystem.** Broad use of SAP has produced widely available community support and training as well as methodical implementation practices and problem resolution approaches. Drawing on synergies of data integration with information stewardship, data preparation and systematic delivery via its Integration Solution Advisory Methodology, SAP is making collaboration easier between business leaders and data integration practitioners to concertedly derive value from data.

Cautions

- **License and cost model.** Some reference customers expressed concerns about high prices and a licensing model that they perceive as limiting their ability to expand deployments. Various prospective buyers expressed challenges with SAP's data integration tool pricing relative to their expectations and budgets as adverse factors in competitive evaluations.

- **SAP-focused customer experience.** Although its reference customers generally indicated that they would recommend SAP to others, a substantial number would only do so with qualifications. SAP's data integration tools are largely sold and deployed in SAP-centric environments, which reflects a more targeted focus compared with the major and leading competitors in this market.

- **Deployment and administrative complexity.** Reference customers expressed concerns about the complex use of multiple products together, and overlapping functionality in multiple products. They cited improvements needed in the ease of diagnosing errors as well as in administrating and monitoring data integration operations. SAP continues its efforts to simplify seamless administration, monitoring and orchestration of all integration technologies — especially with SAP Data Hub, which focuses on the integration needs of SAP and non-SAP environments.

SAS

Based in Cary, North Carolina, U.S., SAS offers a number of data integration products. They are SAS Data Management, SAS Data Integration Studio, SAS Federation Server, SAS/ACCESS, SAS Data Loader for Hadoop, SAS Data Preparation and SAS Event Stream Processing. The vendor's customer base for this product is 14,000 organizations.

Strengths

- **Investments aligned to market vision.** SAS has made strategic investments in open-source metadata management projects, including a recent integration with ODPi Egeria. This enhancement will allow SAS customers to inventory their distributed and connected data assets as well as exchange metadata bidirectionally between different data management technologies (SAS or non-SAS). The vendor has also invested heavily in active metadata and ML-based self-healing capabilities, which assist in distributing, optimizing and recomposing integration processing to help developers and citizen integrators automate parts of the data integration process.

- **Technical and after-sales support.** A significant number of reference customers praised SAS for its timely and effective technical product support, as well as its account teams. It also received the highest score of these Magic Quadrant vendors in the service and after-sales support category, and reported strong renewal rates for its data integration tool portfolio.

- **Alignment for operationalizing self-service data preparation.** With self-service data preparation becoming critical to fulfill the ever-growing integration demands of citizen data scientists (a key target group for SAS), the vendor has responded with significant investments in

its data preparation platform. SAS Data Preparation (on SAS Viya) now includes data quality capabilities with options for stream data integration, MPP support, automated data content tagging based on NLP analysis, and cloud data exchange (for a hybrid integration option). SAS Data Integration Studio also has Git integration. Reference customers have embraced these enhancements and gave SAS the highest score for usability support of all the vendors in our reference customer survey.

Cautions

- **Pricing method, price point and contracting.** Reference customers for SAS continued to cite high costs, complex pricing models and issues with contract negotiations as the biggest challenges while adopting SAS's data integration software — a perennial issue. They also identified limited contract flexibility, especially during the renewal process, as an ongoing issue. Vendor lock-in and lack of contract transparency make negotiating with SAS more difficult than with some other vendors. SAS is looking to address these concerns by introducing serverless metered and tiered-based cost models.

- **Perception of analytics-only focus.** While SAS has strengthened its data integration capabilities to address a wide portfolio of operational data integration use cases, its data integration tool portfolio continues to be sought mostly for analytics and data-science-related requirements (per survey respondents and Gartner client inquiries). While not an issue itself, customers should be aware that SAS's data integration functionality is mostly marketed (hence perceived to be strong in its support) for analytics and data science use cases in SAS. Thus, its data integration tool portfolio may not provide specific advantages over other tools beyond these use cases.

- **Ease of upgrade and deployment.** A small number of reference customers reported continuing upgrade difficulties with SAS's data integration tools. Some also cited early deployment issues and a steep learning curve as the two biggest challenges while adopting SAS's data integration technology. The vendor received below-average scores for version upgrades and ease of deployment in the reference survey, reinforcing these issues.

SnapLogic

Headquartered in San Mateo, California, SnapLogic offers SnapLogic Intelligent Integration Platform. The vendor's customer base is around 2,500 organizations.

Strengths

- **Integration convergence and augmented data integration delivery.** SnapLogic's technologies span data and message-oriented integration, as well as optional add-ons for event stream processing, big data processing, data science, B2B integration and API-enabling capabilities. Its integrated offering of data and application integration capabilities on a single platform establishes its relevance for HIP requirements. Developers can use SnapLogic's many connectors (over 500 "Snaps"), build pipelines (integration flows that can be turned into REST APIs that can receive payload and parameters in addition to returning data when invoked), and create, find and modify reusable patterns. The vendor also includes Iris — its metadata-based

AI engine that allows clients to utilize and analyze active/passive metadata to automate some data integration requirements.

- **Accessible to diverse user personas.** Reference customers highlighted SnapLogic's ability to empower less-technical integration roles and business intelligence users, reducing reliance solely on highly skilled specialist integrators. Embedded ML for automated guidance helps accelerate the development of integration flows, which can reduce the reliance on detailed documentation.

- **Pricing model simplicity and trial version.** Reference customers scored SnapLogic well above average for its pricing model transparency and simplicity. Enhanced self-service license options to encourage adoption and onboard prospects faster include a free trial version to ease exploration of product features.

Cautions

- **Code and documentation quality.** Reference customers scored SnapLogic well below average for the quality of its documentation. Some respondents also cited code quality issues during upgrades that sometimes introduced show-stopping bugs.

- **Dated user interface.** SnapLogic provides the ability to address a diverse set of use cases with a single, unified user interface, but several reference customers stated that this interface feels aged when compared to those of competitors. To streamline the efficiency of troubleshooting integration flows and pipeline components, the vendor has introduced some recent user interface improvements aimed at alleviating these concerns.

- **Data management versatility.** Although SnapLogic provides data profiling and data cataloguing, some reference customers expressed a desire for improvements that enable metadata and governance support in data integration activities. They also noted a wish for improved coverage of data management support, including comprehensive data quality and lineage analysis and tracking, as well as metadata management requirements.

Syncsort

Based in Pearl River, New York, U.S., Syncsort offers the Syncsort Connect family of products. This includes Connect CDC, Connect for Big Data, Connect ETL, Connect AppMod and Connect Sort. It also includes Ironstream, which integrates security and operational data from z/OS mainframe as well as IBM i systems for use in technologies such as Splunk and ServiceNow, and Elevate MFSort for mainframe data integration. The vendor's customer base for this product set is around 2,250 organizations.

Strengths

- **High-performance functionality.** Strong performance for ETL, lower TCO compared with the market leaders, and time to value were all cited as key attractions for reference customers who chose Syncsort. A focus on offloading mainframe data processing workload to Hadoop or cloud data warehouses — along with added support for containerization, cloud deployment and

blockchain in both its data integration and data quality tools — extends this vendor's relevance in digital business.

- **Aligned support for data management and governance needs.** By connecting legacy data to emerging technologies including multicloud and blockchain as part of its Syncsort Invent initiative, Syncsort assimilates existing and acquired data-focused technologies into a unified portfolio. This enhances alignment across its data integration, quality/governance and data management solutions. For tightening product interoperations, the vendor sets out to enhance metadata sharing, integrated design and administration, and a common user interface across components of its portfolio.

- **Customer relationship and track record.** Reference customers identified their overall relationship with the vendor as a positive, acknowledging its partnering posture toward clients. With a track record established in ETL optimization, an enlarged product portfolio, a loyal customer base and a go-to-market approach via strategic partnerships, Syncsort has a solid foundation on which to grow its market presence.

Cautions

- **Integrated use of vendor portfolio.** To facilitate a seamless ability to expand deployments across use cases, reference customers cited their desire for improved metadata management support and simpler interoperability of data integration tooling with Syncsort's other products.

- **Skills availability.** Finding skilled resources was reported as a challenge for reference customers when trying to deploy, interoperate and maintain multiple products of Syncsort in complex integration environments.

- **Appeal for diversifying integration roles.** Syncsort has its roots in technical offerings that are well aligned to communities in the IT arena. With current market momentum that trends toward the needs of multiple personas, a lack of business-facing or self-service data integration options presents constraints to less- and nontechnical integrators. The vendor recently launched an initiative to create a consistent user experience for its data integration and data quality products, to help less-technical integrators configure data connections and monitor project statuses, and to simplify common usage patterns (e.g., populating data lakes).

Talend

Based in Redwood City, California, U.S., Talend offers Talend Open Studio, Talend Data Fabric, Talend Data Management Platform, Talend Big Data Platform, Talend Data Services Platform, Talend Integration Cloud and Talend Stitch Data Loader. The vendor's paying customer base for this product portfolio is more than 3,200 organizations.

Strengths

- **Market momentum and mind share.** Talend executed strongly in the data integration tool market in 2018/19, growing license revenue approximately 30% and doubling its licensed data integration customers, boosted by its acquisition of Stitch. The vendor has also captured mind

share, appearing in over 50% of Gartner contract review calls for data integration tools and featuring very frequently in competitive situations during our inquiry calls.

- **Strong portfolio and strategic investments in data integration innovation.** Talend fields a complete data management and integration portfolio as part of Data Fabric, which includes data integration, data preparation, data catalog, data stewardship and API management capabilities. It has expanded its data integration portfolio through the recent acquisitions of Stitch (for cloud data loading and migration capabilities) and Restlet (for API integration), along with organic development of its data catalog tool (for inventorying connected data assets). These strategic investments along with existing iPaaS capabilities position the vendor well for hybrid and multicloud integration support.

- **Open-source commitment.** Talend has an active open-source user community committed to product support. Like other Leaders, it continues to leverage a broad variety of open-source innovations, including Apache Beam, Spark and Kafka, which are now used extensively in Talend Data Fabric. Reference customers said they value the openness (via bidirectional metadata sharing) and configurability of the vendor's data integration tools, along with its ability to deliver reusable open-source integration artifacts built by its practitioner community.

Cautions

- **Pricing model scalability.** Although Talend's pricing remains lower than that of some Leaders, both Gartner clients and Talend reference customers have noted a steady price increase in its subscription named user license costs year over year. The vendor has long appealed to customers due to its "per seat" pricing, which provides a predictable cost model. But this advantage is slowly eroding for existing customers wishing to scale yet feeling that the vendor should provide more license options (including enterprise license models) beyond the named user license model, which could prove expensive for large customers. Existing customers should investigate its alternative pricing models to achieve additional savings through the adoption of innovations such as serverless technology.

- **Online documentation, migration and upgrading.** Some reference customers identified difficulties with migration between versions. In some cases, a refresh of integration flows was required to take advantage of new capabilities, which lengthened the transition process. Some respondents requested better provision of online documentation, relevant and customized to their software stack. Talend's overall survey score for version upgrade or migration experience was below the average for this Magic Quadrant.

- **Professional services scarcity.** Some reference customers and Gartner clients reported challenges with finding and engaging skilled professional services partners specializing in strategy and deployment of Talend's new data integration and management products. The vendor received below-average survey scores for professional services satisfaction. To rectify these concerns, Talend has initiated an expanded partner training program and added new delivery partners in Europe and Asia/Pacific, as well as introduced "Customer Success Architects" to assist larger customers and complex integration efforts.

TIBCO Software

Based in Palo Alto, California, U.S., TIBCO Software offers a number of data products. They are TIBCO Data Virtualization (TDV), TIBCO Cloud Integration, TIBCO EBX, TIBCO StreamBase, TIBCO Messaging and TIBCO Spotfire. TIBCO Messaging includes TIBCO FTL, TIBCO eFTL, TIBCO Flogo Enterprise and TIBCO Enterprise Message Service. TIBCO Spotfire is included for its embedded data preparation and data catalog capabilities. The vendor's customer base for this product set is more than 5,000 organizations.

Strengths

- **Strategic acquisitions and investments.** TIBCO has made a few recent strategic acquisitions to strengthen its data management portfolio. These include Scribe Software (June 2018), which added capabilities including iPaaS, data synchronization and support for data migration, and MDM provider Orchestra Networks (December 2018) to integrate MDM, data governance and metadata management capabilities in its Connected Intelligence strategy. TIBCO also acquired SnappyData (March 2019), an in-memory DBMS that is expected to improve the performance of its data integration tools portfolio.

- **Established stream data integration and messaging capabilities.** TIBCO has a comprehensive portfolio of capabilities for supporting popular and upcoming data delivery styles, including stream data integration and message-oriented data movement. It recently announced support for TIBCO Eclipse Mosquitto Distribution (for connecting IoT devices using MQTT to TIBCO messaging components) and for Apache Kafka (for real-time streaming data pipelines). Reference customers scored the vendor highest of all this group for its messaging and streaming data integration capabilities.

- **Comprehensive data virtualization capabilities.** TDV has been recognized by reference customers as a reliable data virtualization tool that allows TIBCO to execute on its Connected Intelligence strategy by connecting to data "in place" for analytics and operational use cases. Respondents also praised the vendor's commitment to TDV's product development, which includes an MPP engine for dynamically distributing big data workloads across multiple processes (via Apache Drill embedding), and support for Elasticsearch.

Cautions

- **Talent sourcing and ease of use.** Some reference customers identified challenges with finding, sourcing and retaining talented data integration architects and developers adept with TIBCO's comprehensive data integration tools portfolio. Some also called out the vendor's tools for having a steep learning curve. TIBCO is addressing these concerns actively by expanding its SI and partner network, and ensuring more certifications within it.

- **Confusion over integration across product set and product messaging.** With multiple acquisitions and new tool developments, TIBCO's product portfolio for data integration has become complex. With multiple offerings for data and application integration, reference customers cited "disconnected capabilities" spread across multiple offerings as a challenge, and requested a more simplified and unified portfolio. Some respondents cited confusion and lack of clarity concerning TIBCO's roadmap for its newly acquired tools.

- **Limited support for complex bulk/batch.** While TIBCO has been praised for its leadership in messaging and data virtualization capabilities, reference customers cited complex ETL/ELT capabilities as needing improvement. While TDV can complement ETL/ELT by providing robust caching and querying — and even support a data service that an ETL system can use as a source — TIBCO's data integration tools alone cannot address the full scope of bulk/batch-related workloads.

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

- Qlik — through its acquisition of Attunity, finalized in May 2019

- SnapLogic

Dropped

None

Inclusion and Exclusion Criteria

The inclusion criteria represent the specific attributes that Gartner 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 the overall range of capabilities expected of data integration tools. Specifically, vendors must deliver the following functional requirements:

- Provide a reference customer base representing any mix of the vendor's products for use in at least three of these seven key technical data delivery styles:

  - **Bulk/batch data movement —** Includes single pass or multipass/step processing that incorporates 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.

  - **Data services orchestration —** The ability to deploy any of the other data integration styles, but with the specific capability to interoperate with application services (logic flows,

interfaces, end-user interfaces, and so on). Also, the ability to 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 virtualization —** 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. Data virtualization may or may not include redefinition of the sourced data.

- **Data replication —** 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.

- **Data synchronization —** Can utilize any other form of data integration. However, it 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. This 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.

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

- **Stream data integration —** Data consists of datasets that follow consistent content and structure over long periods of time, and large numbers of records 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. These outcomes may or may not be aggregated/parsed into separate event streams from the same continuous stream. The logic for this approach is embedded in the data stream processing code.

- **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, cloud-based data asset types (including data of SaaS applications and cloud data stores), and streaming data.

- **Mode of connectivity/adapter support for interaction (against a range of sources and targets) —** Support for bulk/batch, change data capture and real-time/near-real-time connectivity/interaction (in a continuous or streaming data environment, for example).

- **Data transformation support —** Packaged capabilities for basic transformations (such as data type conversions, string manipulations and calculations) and moderately complex transformation (such as integration with data quality and MDM tools to access changes in data validation directly from the metadata within those solutions).

- **Metadata and data modeling support —** Automated metadata discovery (such as profiling new data sources for consistency with existing sources), lineage and impact analysis reporting, and the ability to synchronize metadata across multiple instances of the tool. Also includes an open metadata repository with mechanisms for bidirectional sharing of metadata with other tools. As a desirable advantage, some data integration tooling may provide ML-enhanced metadata discovery as well as internal analytics to enhance human data management and integration requirements, using both passive and active metadata collection, sharing and analysis.

- **User- or role-specific variations in the development interface —** These are capable of various workflow enhancement mechanisms. These mechanisms may include supporting templates, version modification (via internal library management or other mechanism) and quality assurance capabilities — either via audit/monitor metadata (manual) or through 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 service deployments in production environments, as well as alternative or collaborating development environments.

- **Data governance support —** The ability to import, export and directly access metadata with data profiling, data quality tools and/or other information-governance-enabling technologies (such as MDM, information stewardship, metadata management and data catalog tooling). Accepting business and data management rule updates from data stewardship workflows and sharing data profiling information with such tools are highly desired.

- **Runtime platform support —** Windows, UNIX or Linux operating systems. Demonstrated capability to operate on more than one commercially available cloud environment is desired.

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

In addition, each vendor has to satisfy the following quantitative requirements regarding their market penetration and customer base:

- **Revenue —** Generate at least $30 million of its annual software revenue from data integration tools (perpetual license with maintenance or subscription with support). Or, it must maintain at least 300 production maintenance/subscription-paying customers for its data integration tools. (Note that the number of downloads without license or maintenance revenue is informative, but not a qualifying piece of information.) Gartner will use as many independent resources for validating this information as possible, specifically to validate provided information.

- **Geography —** Support data integration tool customers in at least two of the following geographic regions or specific national markets: North America, South America, EMEA and Asia/Pacific. Specific exceptions may be possible relative to also meeting the revenue or verifiable supported customer counts — understanding that Gartner analyst decisions are final.

- **Presence —** Demonstrate sufficient market presence, reviewed and 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 that focus on narrow use cases that are too specific for broader market application could be excluded. In the past, some vendor/supplier tools were excluded because they:

- Focused on only one horizontal data subject area — for example, the integration of customer-identifying data.

- Focused on only a single vertical industry.

- Served only their own, internally managed data models and/or architectures (this includes tools that only ingest data to a single proprietary data repository).

- Were used by a single visualization or analytics processing platform.

- Had a DBMS/data management solution for analytics (DMSA)/data lake management vendors that use their data integration tools only to ingest/integrate data into their own repository.

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, including 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/Description section or described below). Product strategy is expected to be in support of traditional integration needs filling current gaps, weaknesses and opportunities to capitalize on less-advanced demand trends in this market.

Given the requirement for data integration tools to support diverse environments for data, delivery models and platform-mix perspective, we also assess vendors on the degree of openness of their technology and product strategy. 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 MDM) becomes an integral support for all use cases that can share high-quality data as well as lineage metadata, with runtime management and monitoring support.

For broader-spectrum solutions, the market has deemphasized the product capability and emphasized the ability to break out pricing and components. Various capabilities are crucial to the success of data integration tool deployments. These include:

- How well the vendor supports the range of distinguishing data integration functionalities required by the market

- How this functionality is delivered

- Support for established and emerging deployment models

- Overall usability and consumption of the tools

Overall Viability

An assessment of the vendor's overall financial health as well as the financial and practical success of the business unit.

We view the likelihood of the organization continuing 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 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. Organizations increasingly seek "modularity" or the capability to isolate specific required functions in data integration that are then reflected in their implementation approach and cost allocation.

The focus on pricing by vertical — which allows for pricing by use case, role, and volumetric and performance metrics (all considered applicable for different market needs) — has increased in 2019. In addition, pricing by feature, deployment model, user persona 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, as well as the effectiveness of its direct and indirect sales channels, was scored as part of the evaluation.

Market Responsiveness/Track Record

The vendor's 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 virtulization, message-oriented data movement, replication and synchronization, and stream data integration support. However, traditional bulk/batch processing is still the predominant demand. Not only do most solutions overlap; the market is now 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, as well as how well the vendor acted on the vision of prior years, is also evaluated.

Marketing Execution

The clarity, quality, creativity and efficacy of programs designed to deliver the vendor'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, sales, events and other promotional 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. Providers must develop a means of converting community "chatter" and excitement to support delivery and go-to-market campaigns.

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 interactions with buyers, technical support or account support. This may also include ancillary tools, customer support programs, the availability of user groups, service-level agreements (SLAs) and so on.

Data integration has evolved to include a broad range of expectations when it comes to customer experience. We evaluated the level of satisfaction expressed by customers with the vendor's product support and professional services as well as customers' overall relationship with the vendor, and their perceptions of the value of the vendor's data integration tools relative to cost and expectations.

The distinction between advanced use cases and "pedestrian" applications is becoming more pronounced. The evaluation this year is focused on separating success in "traditional" market delivery from "innovative" in reviewing the customer experience. The evaluation of vendors against this criterion will continue to be driven directly by the results of our customer reference survey (see the Evidence section).

Operations

The ability of the organization to meet goals and commitments.

Factors for this criterion 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 deployment with greater focus on data services, metadata management and semantic tiers. Also, they must demonstrate that they are well positioned to 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	High

Source: Gartner (August 2019)

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

The vendor's ability to understand customer needs and translate them into products and services.

Vendors that show a clear vision of their market will listen to and 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. However, metadata-driven warnings will be needed, 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 and 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 into 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 the 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. Organizations now expect differentiated pricing based upon their use cases, as well as rapid conversion to scalable pricing models when new

demands are introduced. Base pricing must include development and test environments that are minimal- or no-cost — or represent a convertible investment when going into production. Additionally, cost models that address mass production environments versus "citizen" or data-science-driven efforts require flexible pricing.

The market seeks efforts from vendors to break free from pricing models oriented on hardware metrics alone, to provision different pricing models and license packaging that are targeted and customized toward user persona type and use case addressed. Suppliers must consider if their internal compensation models incentivize delivery that matches customer demand and implementation profiles.

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 and weaknesses as well as opportunities to capitalize on advanced demand trends. There is now significant increased expectation in "active" metadata understanding, conversion, utilization and analysis of this metadata (see Note 2). This active metadata is used in profiling, ML learning, evaluation of assets and comparison to existing integration upon connection. Self-correcting optimization in processes is now important and expected. Utilizing metadata to assist in user "push" recommendations for new data assets, as well as create semantic knowledge graphs to assist with data fabric design that enables a more consistent (and application-neutral) semantic model for integration, is considered a differentiator.

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 both 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 on its strategy for the data integration tool market — including diversity of delivery models, packaging and pricing options, and partnerships — 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, defensive or preemptive purposes. The current innovation demands in the market are centered on managing location-agnostic capability in data integration (i.e., the ability to not have to move or replicate data necessarily but to connect to data in-place when feasible and take the processing to the data [rather than vice versa] to execute integration).

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 recommend/determine when data needs to be moved for optimal processing. As data 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 now expected.

The degree to which the vendor demonstrates creative energy in the form of enhancing its practices and product capabilities is important here. Also important is introducing thought-leading and differentiating ideas/product plans with the potential to significantly extend or reshape the market in a way that adds real value for customers. Finally, the importance of ML-based automation using internal analytics on all kinds of collected metadata to support integration activities is another area of improvement that the market currently demands. The growing diversity of users indicates a much higher demand for administrative, auditing, monitoring and even governance controls that utilize job audit statistics. Graph analysis to determine user classification, tendency toward types of data and processing design, and optimization "hints" is also increasingly demanded.

Geographic Strategy

The vendor's strategy for directing 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 is appropriate for that geography and market.

User organizations are seeking local support with differing levels of confidence in the various approaches possible (i.e., VARs, resellers, channel partners, OEM offerings and distributors). They are seeking continuity of support across regions as well. 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 into markets beyond its home region/country and its approach to achieving global presence (e.g., direct local presence and use of resellers/distributors) are 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	Medium
Sales Strategy	High
Offering (Product) Strategy	High
Business Model	Medium
Vertical/Industry Strategy	Low
Innovation	High
Geographic Strategy	Medium

Source: Gartner (August 2019)

Quadrant Descriptions

Leaders

Leaders in the data integration tool market are front-runners in their capability to not only support a full range of data delivery styles but also allow for the combination of these different data delivery styles (e.g., data virtualization and ETL). In 2019, they have been advancing their metadata capabilities, with some introducing highly dynamic optimization and advanced design assistance functions. They have been extending their capabilities to allow for ML over active metadata, to assist developers with various degrees of support and automation in integration design and implementation. Leaders have recognized the growing affinity between data and application integration, and introduced tools that bridge the gap between application and data integration through iPaaS capabilities. Leaders are also approaching location-agnostic deployments in a synergistic manner.

Providers or solutions not limited to cloud- or on-premises-only that can be deployed beyond a specific location are gaining traction (i.e., on devices, as agents/daemons on sensors, in the IoT and more). Leaders are adept at providing tools that can support both hybrid integration and multicloud integration options, bridging the data silos that now exist across on-premises and multicloud ecosystems. Leaders are strong in establishing their data integration infrastructure as an enterprise standard in at least one primary use case, with the capabilities to deliver in multiple use cases. They

become 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 market mind share, and resources skilled in their tools are readily available. These vendors recognize the need for new and emerging market demands — often providing new functional capabilities in their products ahead of demand — 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 — either directly or through a parent company.

Challengers

In 2019, the Challengers have been making significant strides in understanding that self-service data preparation, data replication, cloud integration and data virtualization are no longer differentiating features but must-have capabilities within a broader metadata-driven data integration solution. And even as Challengers, they are expected to have good executable technology covering these trends (a facet previously associated with Leaders only).

Challengers include vendors that have converted their messaging, roadmap and vision into market delivery — or, having failed to do so, have become Niche Players. In general, Challengers are well positioned in light of the key existing practices in the 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 established presences. They exhibit 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. Some Challengers have developed best practices for leveraging their strongest product in new delivery models, such as data virtualization that can manage message queues or incorporate databases as cache overflow repositories for easy retrieval and/or as an optimization technique.

Overall, the market is pushing Challengers to utilize metadata and perform ML over this metadata to deliver solutions that can automate various data integration tasks (e.g., automated profiling, transformations, data preparation, performance optimization, query optimization, scaling, tuning, movement of workloads to data stores and engines best suited for processing). Overall, this will be a key area in 2019 that will determine which Challengers can move into the Leaders quadrant next year.

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 expected to grow in demand. In 2019, the Visionaries have taken an early focus on alternative go-to-market strategies or specific capabilities that capitalize on their capacity to leverage either:

- ML-assisted integration

- Serverless integration tooling that supports a multicloud and hybrid cloud integration architecture, *or*

- The growth in demand for connectors and API/microservices

In addition, a significant driver of vision for the market this year has been the ability of tools to connect to and analyze all forms of metadata — not just passive metadata but increasingly active as well (see Note 2 for definitions). With this, tools can provide key statistics to developers and citizen integrators that aid with integration design and, increasingly, integration automation. The Visionaries are doing all of these things. Additionally, the development of specific solution designs that mix repeatable delivery templates with appropriate custom-deployed options represents the potential to address specific new markets demanding data integration in highly specific vertical offerings.

Visionaries sometimes lack market awareness or credibility beyond their customer base or 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. They may be large, established players in related markets that have only recently placed an emphasis on data integration tools.

Niche Players

With the market now matured, Niche Players here generally don't exhibit gaps in primary market functionality or features. Instead, they are simply either challenged in increasing their execution or have not identified a specific market approach that expands use cases for their technology. This means that almost every Niche Player will be able to deliver against standard market expectations both in functionality and cost-price options.

Niche Players do not appear very frequently in competitive situations for comprehensive data integration tools for enterprise-class deployments. Many have very strong offerings for a specific range of data integration problems (e.g., a particular set of technical environments, application domains or use-case scenarios), and deliver substantial value for their customers in the associated segment. Niche Players now exhibit particular advantages in pricing, in their small footprint and even in vertical or horizontal solutions. This 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.

Clients should note that more than 80% of all end-user organizations still seek bulk/batch processing (even in hybrid and intercloud scenarios — see the Evidence section for details). This means that a highly efficient but batch-oriented data integration tool vendor could exhibit high-level execution capabilities without ever crossing to the right-hand side of the Magic Quadrant (i.e., to the

Visionaries or Leaders quadrants). Niche Players all exhibit batch capabilities, from back-office and operations data through to massive volumes of sensor or IoT data.

Context

The market for data integration tools continues to evolve and is supported by strong levels of adoption. More data and analytics leaders are realizing that data integration is a critical component of their data management infrastructure. They understand that they need to employ data integration functions to share data across all organizational and systemic boundaries. Organizations are, therefore, increasingly seeking a comprehensive range of improved data delivery capabilities to modernize their data, analytics and application infrastructures.

Data and analytics leaders must navigate a market brimming with products that claim to solve a range of data integration problem types. However, not all have experience in — nor evenly provide — all of the relevant capabilities needed across our key use cases. Some vendors focus heavily on moving data in bulk or batch, but may place less emphasis on capabilities such as data virtualization, data replication, message-oriented integration, stream data integration or data synchronization.

In 2019, traditional integration has begun to shift from the bulk/batch dominance for delivery in the market — but only marginally. More than 80% of all organizations in our survey make significant use of bulk/batch, but as much as 30% also utilize data virtualization or data synchronization with layers of data processing afterward.

Some organizations have determined that basic functions are adequate for them, and are, therefore, seeking tools with focused and targeted capabilities. As a result, they are interested in evaluating and procuring tools that are specialists in one data delivery style (e.g., data replication or data virtualization). Others have decided to initiate an enterprisewide effort to optimize the breadth of approaches. They seek tools that are generalists or best of breed in data integration, have varied data delivery styles and support a combination of these different data delivery styles (see "Modernize Your Data Integration Architecture for Digital Business by Combining Data Delivery Styles").

While many organizations have traditional solutions in place, modern demands have increased the utilization of stream data integration (to capture data as it is generated in near real time), data virtualization and data replication/synchronization. Somewhere between 35% and 45% of all organizations surveyed are using at least two of these alternative approaches — and an even higher percentage of leading or large organizations are doing so (see "Market Guide for Data Virtualization").

In the context of digital business, business moments are increasingly attracting the attention of enterprises. ("Business moments" are opportunities of short duration or a point in time that sets in motion a series of events involving people, businesses and things.) They want to harness data to seize these moments, which will require data integration support that includes a focus on stream/ event-oriented capabilities. This is where stream data integration is now becoming a strong

capability and a relevant data delivery style for modern data integration requirements (see "Adopt Stream Data Integration to Meet Your Real-Time Data Integration and Analytics Requirements").

The era of metadata-powered integration tools is well underway. Metadata as a byproduct of the design and operations management of a data integration platform is a minimum requirement of data integration tools in 2019. Platforms and solutions are now expected to provide continuous feedback regarding the profiles, quality, location, performance optimization, lineage, use cases, access points, context, frequency of access and content analysis of integrated data assets. As far as architects and solution designers are concerned, this feedback is long overdue. It is expected that graph analytics powered by every conceivable type of metadata (both passive and active) will provide the necessary dynamic data fabric designs for introducing ML capabilities into data integration platforms (see "'Unlearn' Existing Practices for Success in Multicloud and Hybrid Data Integration").

Gartner sees that the need to acquire and integrate data from cloud environments, typically for hybrid cloud and multicloud integration, is becoming critical to many data integration use cases. The expansion of vendors' capabilities into application integration provides opportunities to use tools that exploit common areas of both technologies to deliver shared benefits. Organizations have begun to pursue data integration and application integration in a synergistic way in order to exploit the intersection of the two disciplines. This combined capability of integration patterns is a key component in enabling an HIP-inspired infrastructure (see "Use iPaaS to Extend Your Data Integration Strategy to the Cloud in Hybrid Ways").

An interesting data point from the reference customer survey was that close to 58% of surveyed organizations utilize data integration tools for their data replication demands, up from 30% in 2018. This drastic increase is because organizations are looking to utilize the CDC capabilities of their data integration tools to replicate the data from their operational DBMSs to cloud data warehouses supported by dbPaaS. This has been a significant driver of growth for many data integration providers (such as Attunity). Such providers have formed significant partnerships with cloud service providers (AWS, Microsoft Azure, etc.) to deliver integrated data from on-premises data stores and applications to cloud data warehouses and lakes for analytics. They are doing this through forward engineering, often as ready for consumption as integrated data with schema assigned for analytics and data science use cases.

In a continuation from 2018, organizations are seeking solutions that facilitate role-based data integration. This includes the capability to promote or manage the workflow of converting individually developed processes into enterprise-capable ones (see "Market Guide for Data Preparation Tools").

A mix of data integration approaches thus becomes crucial, spanning physical delivery to virtualized delivery, and bulk/batch movements to event-driven granular data propagation. In particular, when data is being constantly produced in massive quantities and is always in motion and constantly changing (e.g., IoT platforms and data lakes), attempts to collect all of this data are potentially neither practical nor viable. This is driving an increase in demand for connection to data, not just the collection of it (see "Modern Data Management Requires a Balance Between Collecting Data and Connecting to Data").

The distribution of required computing workloads to parallelized processes in Hadoop, alternative nonrelational repositories and data lakes will continue to advance data integration tools' ability to interact with, deliver data to and execute integration tasks in emerging data-and-analytics-associated platforms.

Market Overview

Data integration is central to enterprises' data management infrastructure. Enterprises pursuing the frictionless sharing of data are increasingly favoring data integration tools that are flexible in regard to time-to-value demands, integration patterns, optimization for cost and delivery models, and synergies with hybrid integration platform (HIP) and digital business transformation.

Gartner estimates that the data integration tool market revenue was slightly over $3.2 billion in constant currency in 2018, an increase of 4.2% from 2017 (see "Forecast: Enterprise Infrastructure Software, Worldwide, 2017-2023, 2Q19 Update").

The competitive landscape of this market reflects vendors' pursuit of a more comprehensive offering strategy in support of a broad range of use cases, and to capitalize on new demand. Evolving their relevance and competitive positioning requires vendors to extend their vision and deepen their capability to harness market inertia and broaden applicability of data integration offerings. This is in line with buyer expectations for optimal functions, performance and scalability in data integration tools, so that those tools operate well with the same vendor's technology stack and, increasingly, interoperate across related data management and application infrastructures.

Organizations will need to monitor and exploit major themes that are affecting enterprise requirements and vendor offerings in the data integration tool market:

- **Data integration must become proactive to seize business moments.**

    Demand for data to be delivered or processed at the scale of events must match the speed of a business. Enterprises want to harness data so that they can seize "business moments" (i.e., moments that present opportunities of short duration or a point in time, which set in motion a series of events, involving people, business and things).

    There have been extensive efforts in enterprises to provide an architecture that can deliver data in various modes — from batch to real time, bulk to granular, physical to virtualized and central to distributed. Vendors are extending their tool capabilities' focus on flexible latency with a mix of data delivery optimization to meet event-driven data availability requirements. Tools are exhibiting enhanced characteristics in data integration architecture, with deepened integration between bulk/batch delivery, data virtualization, granular and low-latency data capture and propagation, and stream data integration.

- **Extensive metadata support for any data, anywhere.**

    Metadata as a byproduct of design and operations management from a data integration platform is regarded as a bare minimum in this market (this is called "passive metadata"). Data integration tools are increasingly expected to provide continuous feedback regarding the

profiles, quality, use cases, access points, context and content analysis of integrated data assets. Graph analytics on every conceivable type of metadata will provide the necessary information for introducing ML capabilities into data integration activities (this is called "active metadata analysis," which is now required and differentiating).

At the same time, this metadata analysis can and should be coupled with application processing optimization strategies, and even influence how application integration can be supported by data integration tools. Often, the application or data alone is incapable of absorbing the speed of change in operational process models. As a result, the influence of HIP approaches in data integration is rapidly gaining ground.

- **Use of ML in data integration.**

Although at an early stage in the market, ML techniques are increasingly being used within data integration tools (see "Innovation Insight for Digital Integrator Technologies"). Approaches such as natural language processing can facilitate the resolution of integration problems and potentially increase time to value, reduce complexity and empower less-technical roles for data integration tasks. A focus on AI-powered data integration technologies is surfacing in areas such as engagement via chatbots or voice, automation of flow creation via next-best action and intelligent data mapping, and insight for processing optimization and self-healing platform operations.

- **Use of ML for augmented data integration.**

In addition to the above, data integration platforms amass vast experience regarding utilized, its flow and patterns, and even applicable scenarios. As such, these platforms represent what is potentially the richest environment for amassing ML training data, and an early opportunity for the data tagging that is necessary for ML.

- **Autonomous optimization for blending traditional deployments with modern infrastructure practices.**

Operational data consistency, data migration and cloud-related integration, the logical data warehouse, big-data-related initiatives, distributed processing workloads such as Hadoop, and alternative nonrelational repositories will all continue to advance. Another aspect of this is how vendors/suppliers will choose to respond.

- **Data integration is at the heart of a dynamic data fabric.**

A data fabric is generally a custom-made design that provides reusable data services, pipelines, semantic tiers or APIs via combination of data integration approaches (bulk/batch, message queue, virtualized, streams, events, replication or synchronization), in an orchestrated fashion.

The steady growth of the data management and analytics markets has demonstrated the value of consistent, semantically harmonized and governed information assets. Increasingly, dynamic recognition of assets of critical importance to business outcomes exists for the organization's ecosystem. This is being driven by the need for consumption, modeling and effective visualization of a growing and varied source of information assets. All of this needs to be carried

out in a consistent and semantically harmonized way, and should be enabled by layers of active metadata utilization.

Data integration tools are now tasked with creating agile and trusted data that enables frictionless information exchange without necessarily having to replicate data or assign schema to it upfront. Tool providers are now expected to deliver these integration approaches (bulk/batch, stream, messaging, etc.) and ensure that users can combine them and deliver the resultant integrated data as reusable data services (physically and virtually) for multiple use-case scenarios across ecosystems.

Data fabrics can be improved by adding dynamic schema recognition or even cost-based optimization approaches (and other augmented data management capabilities). As a data fabric becomes increasingly dynamic or even introduces ML capabilities, it evolves from a data fabric into a data mesh network.

(See "Top 10 Data and Analytics Technology Trends That Will Change Your Business.")

- **Data integration becomes everyone's responsibility.**

Role-based development and deployment of data integration is now incumbent on the data integration tools to deliver. But not all integrators are equally skillful. These new integrators must be either guided or controlled by the platform, to ensure that solutions developed using a data integration platform can and do meet their requirements.

- **Links to data preparation.**

Organizations need their data integration tools to also provide capabilities for self-service data preparation — that is, nontechnical users such as business analytics or citizen integrators need to be able to integrate data with minimal IT support for their specific business requirements (see "Market Guide for Data Preparation Tools").

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.

- **Diverse interfaces for different data integration user persona.**

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.

- **Customer support and services tuned to different users.**

Buyers are demanding highly responsive and high-quality technical support for products for advanced users. At the same time, they want direct and frequent interactions with sales teams and executives. Buyers also want broad availability of relevant skills — both within a provider's installed base and among its SI partners — and forums where they can share experiences, lessons and solutions with their peers.

- **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. Organizations have begun to pursue data integration and application integration in a synergistic way in order to exploit their intersection. This combined capability of integration patterns is a key component in enabling an HIP-inspired infrastructure. At the same time, data as a service (DaaS) provides platform- and context-independent data in a similar fashion, but a completely platform-agnostic design option.

Gartner Recommended Reading

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

"Critical Capabilities for Data Integration Tools"

"Adopt Stream Data Integration to Meet Your Real-Time Data Integration and Analytics Requirements"

"'Unlearn' Existing Practices for Success in Multicloud and Hybrid Data Integration"

"Market Guide for Data Preparation Tools"

"Market Guide for Data Virtualization"

"Toolkit: RFP Template for Data Integration Tools"

"Research Library: Fundamentals for Data Integration Initiatives"

"Are You Ready for Multicloud and Intercloud Data Management?"

"Modernize Your Data Integration Architecture for Digital Business by Combining Data Delivery Styles"

"Rebalance Your Integration Effort With a Mix of Human and Artificial Intelligence"

"How to Implement a Hybrid Integration Platform to Tackle Pervasive Integration"

"Top 10 Data and Analytics Technology Trends That Will Change Your Business"

"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 research is based on information from sources including, but not limited to:

- 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 product capabilities.

- Gartner's 2019 Data Integration Customer Reference Survey — An online survey of the reference customers provided by the vendors in this Magic Quadrant. This 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, 355 organizations across all major regions provided input on their experiences with vendors and tools in this manner. The survey was conducted during April 2019; the results were collated and analysis completed in May 2019.

- Feedback about tools and vendors captured during Gartner's data integration tools customer reference surveys of 2016, 2017, 2018 and 2019 in reviewing data-integration-related Gartner client inquiry topics and discussion notes from June 2016 through June 2019.

- Market share estimates developed by Gartner's technology and service provider research unit.

- 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.

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 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 for 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 the following 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 installation and deployment of software

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

  - (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

Gartner, Inc. | G00369547

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

Note 2 Passive vs. Active Metadata Support

Passive Metadata

Passive metadata includes automated acquisition of metadata, data model creation, documentation and maintenance. It also includes lineage and impact analysis reporting, an open metadata repository, and synchronization of metadata with an end-user interface to view and work with metadata.

Passive metadata is static metadata that is either updated manually or captured periodically for design versions. The primary distinction from active metadata is that passive metadata consists primarily of documentation, ranging from fixed schema of sources and/or targets all the way through to business definitions acquired in a glossary and maintained as a formal data dictionary.

Active Metadata

The ability of the data integration tool to deliver machine learning (ML)-enhanced metadata discovery, and internal analytics to support, optimize and even automate human data management and integration tasks.

Due to the explosion of data in today's highly connected and digital business environments, the growth in data volume and diversity is fast exceeding the ability to process and integrate this data. Organizations, therefore, expect their data integration tools to provide abilities to autoexecute transformations through ML capabilities. In order to support any ML-based automation, enterprises pursuing frictionless sharing of data must have metadata capabilities that far exceed passive metadata practices. Passive metadata is metadata that is static in nature, usually emerges at design time and often requires human or manual updates. Passive metadata most often consists of simple documentation or design-time technical metadata.

Organizations now need their data integration tools to provide continuous access, analysis and feedback on metadata parameters such as frequency of access, data lineage, performance optimization, context and data quality (based on feedback from supporting data quality/data governance/information stewardship solutions). As far as architects and solution designers are concerned, this feedback is long overdue.

It is expected that graph analytics powered by every conceivable type of metadata will provide the necessary information for introducing ML capabilities into data integration platforms. The result will be systems that utilize both cost-based and priority-based optimization in a policy-driven solution that will eventually consider combinations of data across on-premises and multicloud deployments. These systems will be able to dynamically relocate data, provide data processing services and coordinate with governance policies that consider issues such as rights to privacy and legal jurisdictions and provenance.

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

Gartner, Inc. | G00369547