Magic Quadrant for Data Integration Tools

The data integration tools market is composed of tools for rationalizing, reconciling, semantically interpreting and restructuring data between diverse architectural approaches, specifically to support data and analytics leaders in transforming data access and delivery in the enterprise.

Market Definition/Description

The discipline of data integration comprises the practices, architectural techniques and tools that ingest, transform, combine and provision data across the spectrum of information types. This integration takes place in the enterprise and beyond the enterprise — across partners and third-party data sources and use cases — to meet the data consumption requirements of all applications and business processes.

The market for data integration tools includes vendors that offer software products to enable the construction and implementation of data access and data delivery infrastructure for a variety of data integration scenarios. For vendors, the demand for traditional data integration capabilities alongside the demand for innovative solutions requires robust, consistent delivery of highly developed solutions. Similarly, data integration tools interoperate and integrate with master data tools, data governance tools and data quality tools. Examples of this type of interoperability include:

- Support for governance and management of data assets — Increasingly, data integration tools are expected to collect, audit and monitor information 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. Small devices that roam and attach to data portals will also become prevalent.

- Data acquisition for analytics and business intelligence (BI) and data warehousing — Accessing, queueing or extracting data from operational systems, transforming and merging that data either logically or physically, and delivering it through an integrated approach for analytics 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 build the data access and synchronization processes to support MDM initiatives.

- Data consistency between operational applications — Data integration tools provide the ability to ensure database-level consistency across applications, both on an internal and an interenterprise basis, and in a bidirectional or unidirectional manner. The Internet of Things (IoT) is specifically exerting influence and pressure here.

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

The characteristics of data integration tool usage may not be unique to any one scenario. Technologies in this market are required to execute many of the core functions of data integration, which can apply to any of the above scenarios; examples of the resulting characteristics include:

- Interoperating with application integration technology in a single solution architecture. This is now 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, accepting streaming data, and to the point of provisioning these processes across an enterprise service bus.

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

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

- 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, "stateless" data management tiers, analytic data management repositories, data lakes, and platforms typically associated with big data initiatives — such as Hadoop, NoSQL and cloud-based data stores.

- NoSQL repository, which poses data integration challenges. At the same time, it also provides opportunities to assist in the application of schemas at data read time, if needed, and to deliver data to business users, processes or applications, or to use data iteratively.

- Increasingly, the differing structure of IoT or machine data is introducing new integration requirements.

In recent years, significant submarkets have emerged in parallel to the main market offerings that represent a significant focus on either vision or execution, but do not address *all* data integration delivery requirements. There are tools that focus on the innovative solutions, without the traditional capabilities as well. These include a focus on data virtualization or self-service data preparation, but also specific delivery to support management of data lakes (for further detail, see "Market Guide for Data Preparation" and "Market Guide for Data Virtualization").

Magic Quadrant

Figure 1. Magic Quadrant for Data Integration Tools

Source: Gartner (July 2018)

CHALLENGERS | LEADERS

- Informatica
- IBM
- Talend
- SAP
- SAS
- Oracle
- Microsoft
- Denodo
- Attunity

- Adeptia
- Syncsort
- TIBCO Software
- Information Builders

- Hitachi Vantara

- Actian

NICHE PLAYERS | VISIONARIES

ABILITY TO EXECUTE

COMPLETENESS OF VISION

As of May 2018 © Gartner, Inc

Vendor Strengths and Cautions

Actian

Actian is based in Palo Alto, California, U.S. and, including embedded/OEM, has more than 8,000 customers. Actian offers the revamped DataConnect brand, 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. At the time of writing, HCL Technologies and Sumeru Equity Partners have jointly announced their intent to acquire Actian. They will create a separate board, with no plans to change the leadership of the Actian team/product.

Strengths

- Opportunity to enhance market awareness. Actian has always had a great variety of data management and integration tools, often with separate go-to-market approaches. In 2018, Actian took the opportunity to create the go-to-market branding for DataConnect, which encompasses its overall broad-based software and solution approaches for data integration. As an acquirer, HCL Technologies should provide substantial reseller, system integrator (SI) and OEM partnership opportunities through leveraging its global ecosystem.

- Improving its "niche." We anticipate that HCL Technologies will begin to utilize Actian's data integrator as a "go to" part of its SI professional services, as well as embedding the technology into its next-generation autonomics and orchestration products and platform. Actian will continue to leverage its small footprint, easily embedded solutions in order to drive long-lasting annuity revenue. As a small-footprint tool, it stabilizes quickly and almost bypasses reviews considering its replacement — often for years.

- Processing optimization. DataConnect is a high-throughput, small footprint, versatile ETL data integration platform that supports a broad spectrum of integration patterns including bulk/batch message queue processes, streaming and replication. Actian's data integration tool has always maintained 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 alerts 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 states it will continue to operate as a separate legal entity after acquisition. Existing customers and prospects can proceed with some confidence that even if brand dilution occurs, the embedded solutions will have a long technology life.

- Currently lacks role-based delivery. DataConnect is primarily built for use by traditional data integration experts, who deliver integration as part of an application development or in the capacity of supporting data engineering. Actian plans to introduce role-based interfaces or managed development workflows during the first half of 2019, to guide the development efforts of diverse roles including citizen integrators and integration specialists.

- Customer experience remains primarily bulk/batch. Actian can leverage Apache Spark and other open-source processing models, but primary usage remains bulk/batch. Most organizations begin with bulk/batch, but this can be a limiting factor. A focus on enterprise-class capability for small and midsize businesses (SMBs), combined with its lower pricing, is a current opportunity to extend revenue. Actian needs to expand the breadth of its use cases in order to expand into other delivery channels that serve the SMB market.

Adeptia

Based in Chicago, Illinois, U.S., Adeptia offers Adeptia Connect and Adeptia Integration Suite. Adeptia's customer base for the data integration tools market is estimated to be more than 1,350 organizations.

Strengths

- High productivity for data sharing. Adeptia's data integration tooling's ease of use appeals to IT teams and business roles for interenterprise sharing of data spanning cloud and on-premises data sources and applications. Support for distributed runtime deployment using Secure Bridge enables security/governance policies of data flow, aligned to hybrid deployment models.

- Data and application integration convergence. The assimilation of the enterprise service bus (ESB) technology of Adeptia Integration Suite into Adeptia Connect combines process automation and ETL capability and aligns data and application integration. Added capabilities for large-file data ingestion, and a web-based Process Designer to support the generation of Apache Spark code, extend the applicability of Adeptia for digital business and stream processing requirements.

- Customer experience. Gartner inquiry clients, reference customers and Peer Insights respondents all recognize quality of support services as value points for Adeptia, and reflect a positive perception of value relative to the cost of its subscription licensing. The introduction of metadata inferencing and cataloging enables citizen integrator positons to support self-service deployment and aligns with current market demands.

Cautions

- Adoption coverage. While implementations of Adeptia's data integration tooling resonate well with deployments of a targeted scope — driven predominantly by B2B scenarios — Adeptia's highly focused approach poses a challenge in addressing the breadth of applicability in this market in some competitive situations.

- Availability of skills. While Adeptia emphasizes a codeless paradigm in order to reduce dependence on specialist integrators, finding skilled resources is cited as a challenge — by reference customers and inquiry clients — for organizations needing to implement and maintain deployments as their requirements grow.

- Synergy with data management capabilities. Although Adeptia is working to improve its metadata catalog, customers are looking for more extensive governance support for analytics and data management alongside their use of Adeptia's data integration capability.

Attunity

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

Strengths

- Leadership in replication scenarios. Attunity continues to be referenced and evaluated by clients in the majority of competitive situations for targeted data replication. Reference clients give positive feedback for Attunity's robust, low-cost change data capture (CDC)-based replication capabilities across heterogeneous data sources and types, including its historical strength in addressing mainframe data integration. Continued investment in connectors for data replication across big data and cloud environments, along with a focus on stream data integration (for event data capture), has been well received by clients.

- 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 and technology platform providers, and Attunity being a data integration partner for Amazon Web Services [AWS] and Microsoft Azure). Likewise, for data lakes and big data integration (through partnerships with Hadoop distributions), and for data warehouse automation and beyond (through partnership with vendors such as Teradata, Oracle, SAP and IBM). It has also expanded its partner network to include global SIs such as Cognizant, Infosys and Accenture, and global resellers including Teradata and Hewlett Packard Enterprise (HPE) to help minimize concerns about talent availability.

- Strong alignment to growth and modern data integration needs. Attunity has expanded its capabilities to cater to high-demand areas in data integration, including end-to-end data lake enablement (to automate data lake pipelines), cloud data migration and steaming data integration. An ever-expanding list of connectors for replication/CDC for nonrelational, cloud, mainframe and data warehouse sources and targets — including Apache Kafka message queues, Azure Event Hub, Spark, Hadoop, Amazon Kinesis — has been well received by its customers.

Cautions

- Uneven reports regarding upgrades and technical help. A small but significant number of reference customers cited issues with version upgrades, the technical complexity of migrating between major releases, and the quality of documentation. Some customers had to spend a significant amount of time working with Attunity's development team in order to identify issues and workarounds until hot fixes were released.

- Reference customers want more "operations" information. Some reference customers cited issues with Attunity's inability to provide reports when a data integration job fails, while others requested that the platform display key statistics such as integration job reload time. Attunity has developed metadata-based operations management capabilities to address these concerns, but customer inputs regarding the new functionality have not been received as yet.

- Skills are hard to find. Gartner inquiry clients ask about Attunity skills in the broader market, because expert or even moderately skilled users are difficult to find. Often, Gartner clients and those attending our events report it is quite common to identify an experienced user of any data integration platform, then hire and convert them to using Attunity.

Denodo

Based in Palo Alto, California, U.S., Denodo offers the Denodo Platform as its data integration offering. Denodo's customer base for this product set is estimated to be around 500 organizations.

Strengths

- Strong mind share, momentum and customer support. Denodo continues to expand its leadership and mind share in data virtualization, reaching almost 95% of Gartner client inquiries on the subject. Reference customers cite Denodo for its timely and effective technical and sales support, as well as its account teams.

- Investment in innovation trends. With its latest version 7.0 release, Denodo has introduced new features. The Denodo Platform's support for integrated massively parallel processing (MPP) capabilities along with dynamic query optimization, provides performance optimization, incremental caching of large datasets, reuse of complex transformations and persistence of federated data stores. Denodo has introduced metadata functionality that allows users to inventory distributed data assets connected to Denodo and to collect, access and use metadata to inform data integration activities and data directly.

- Broad connectivity, streaming and cloud support. All available connectors are included within the Denodo Platform's cost. Denodo has improved its connectivity support for streaming data (with support for Apache Kafka, Storm and Spark Streaming, and OSIsoft's PI System). It also supports cloud services on the AWS, Azure, Google and IBM Bluemix marketplaces, and has partnerships with database platform as a service (dbPaaS) vendors such as Amazon Redshift and Snowflake, and support for popular SaaS applications. Denodo can also interoperate with Docker technology for containerization.

Cautions

- Complementary technology needed for diverse integration. Denodo executes well on data virtualization-based use cases. Clients that need to combine data virtualization with additional relevant data delivery styles (such as messaging, ESB, bulk/batch, streaming) or data replication, sometimes need to complement Denodo with competing data integration tools. This challenges Denodo's role as an organizational standard. Denodo continues to address this issue by expanding its functionality and educating customers on how to combine different data delivery styles.

- Interoperability with newly emerging infrastructure. Gartner clients and Denodo's reference customers have begun making demands for their data integration tools to have greater flexibility to support integration with edge devices, iPaaS and other modern infrastructure designs. Users are now looking for a lightweight, iPaaS version of the Denodo platform that is similar to the offerings of other providers in the market — now available on AWS and Azure. (No active references indicating total cost of ownership [TCO] or agility advantages were available at the time of writing.)

- Low availability of skilled practitioners with best practices. Denodo reference customers have expressed a desire for greater ease in obtaining skilled implementers in the market. Denodo has initiated on-demand training capacity — included in its subscription pricing — and seen increased interest for its SI partner training (Denodo reported increased participation of

nearly 300% in 2017) — including training for customers based in Singapore, Sydney and Tokyo.

Hitachi Vantara

Hitachi Vantara's global headquarters is in Santa Clara, California, U.S. Hitachi Vantara does not report customer counts for specific products, but has more than 1,500 commercial customers. Hitachi Vantara offers Pentaho Data Integration and Lumada.

Strengths

- Aggressive leadership. Hitachi Vantara was formed by merging three Hitachi subsidiaries to focus on broad data management and integration markets: Pentaho, Hitachi Data Systems and Hitachi Insight Group. A new C-level team is in place, with a mandate to increase synergy with adjacent operational areas. This new team has a mandate to pursue a new go-to-market strategy and has already implemented sales force training worldwide and established quotas and compensation for sales to focus on data-driven solutions and service.

- Platform fit to new market demands. Hitachi Vantara's data integration strategy targets "edge to cloud," but is noted for supporting analytics and IoT/machine data integration. It has targeted the IoT in the past and has combined the existing Hitachi presence with the market demand for sensors and devices in order to create a big data analytics opportunity. In addition, its existing customer reference base includes high-speed, high-volume transaction customers with cloud-to-cloud integration. Reference customers also report good orchestration in Hadoop markets.

- Simple "first experiences." Hitachi Vantara's community edition presents a "try then buy" model, which is particularly appealing to organizations during initial pilots or when experimenting with advanced data integration solutions. Additionally, the Kettle engine is available as both community and open-source implementations. Reference customers summarize the initial experiences as being easy to implement with good scaling capabilities.

Cautions

- Reorganization could slow data integration innovation. It is the aspiration of every business unit in a larger corporation to become an indispensable contributor. However, this has proven difficult for virtually *all* data integration product teams within larger corporations worldwide. While the new leadership plans to be aggressive, the newly combined Hitachi Vantara mixes different types of data management products and go-to-market strategies, which may prove incompatible. The inclusion of Hitachi Data Systems in the reorganization may present a challenge as well as an opportunity.

- Needs modernization. Gartner inquiry clients and Hitachi Vantara reference customers report that upgrade issues persist in 2017/2018. At the same time, a sometimes basic user interface makes Pentaho Data Integration difficult for users coming from other tools. In general, Hitachi Vantara responds with functionality as demand emerges. Hitachi Vantara added AWS and Azure support as those customer bases grew, and recently added support for Google BigQuery in Pentaho 8.1 (May 2018). However, customers report unexpected

operations management issues when utilizing third-party scheduling and administrative tools.

- Linkage to deployments of data management infrastructure. Some implementations cited a need for more-extensive experience from the vendor when manipulating and delivering data beyond analytics scenarios, and for more guidance to enable data quality and governance in relation to the breadth of data management activities.

IBM

Based in Armonk, New York, U.S., IBM offers the following data integration products: IBM InfoSphere Information Server Enterprise Edition, IBM InfoSphere Information Server Enterprise Hypervisor Edition, IBM InfoSphere Federation Server (now part of IBM Db2 Advanced Enterprise Server Edition), IBM InfoSphere Data Replication, IBM Data Integration for Enterprise and IBM Data Integration for Hadoop, IBM BigIntegrate, IBM Streams and IBM Data Refinery (previously IBM Bluemix Data Connect). IBM's customer base for this product set is estimated to be more than 10,700 organizations.

Strengths

- Broad appeal for modernized integration. IBM has introduced machine-learning-enabled data integration for recognizing schemas, relationships and data integrity, data consistency and quantitative performance and volume metrics for processing automation. Managing remote processing is supported on multicloud deployments — enabling data integration orchestration from cloud to cloud. When this capability is combined with cloudlike deployments on-premises, this cloud-to-cloud scenario becomes a way to address hybrid cloud and on-premises demands. IBM's integration tools can serve as an "axle" between "spinning" sets of the same data in multiple locations. Rudimentary machine learning has been added to recognize data schemas and also analyze data profiles, system design and enhanced optimization.

- Reconstituted pricing. Through mid-2018, IBM has reconstituted its client engagement strategy and tuned it more toward subscription and cloud-based offers and pricing. By introducing "flex points," customers can now purchase what amounts to a platform credit that allows organizations to apply their points to different components based upon current needs. This addresses flexibility issues relating to IBM pricing, although it does not specifically address the total cost.

- 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's technologies within its customer base has driven the wide availability of implementation service providers and approaches to solving complex integration challenges.

Cautions

- Difficulties when competing with "point" solutions. IBM is often perceived as being too big. Formerly more about price, this discussion has changed recently. It is no longer about IBM being too big, it is about developing a rapport with its customers and the broader market, which makes it clear that IBM's offerings are available as point solutions in competitive scenarios. This trend predates the introduction of its flex points pricing.

- Cost model. Prospective customers point to their difficulty in understanding IBM's licensing and pricing methods. Existing customers often express concerns about high costs relative to the alternatives in this market.

- Integrated usage and support across portfolio. To facilitate a seamless expansion of deployments across use cases, reference customers for IBM cite the need for improvement in areas of technical help. They also see the need for simpler integrated use of IBM's data integration tooling alongside its broad set of data and analytics technologies.

Informatica

Based in Redwood City, California, U.S., Informatica offers a series of data integration products as part of its Informatica Intelligent Data Platform. These products are: PowerCenter, PowerExchange, Data Replication, B2B Data Transformation, B2B Data Exchange, Data Integration Hub, Data Services, Integration Cloud Services, Cloud Integration Hub, Big Data Management, Big Data Integration Hub, Big Data Streaming, Enterprise Data Lake, Edge Data Streaming, Enterprise Data Catalog and Operational Insights. Informatica's customer base for this product set is estimated to be more than 9,000 organizations.

Strengths

- Innovation and product strategy. Informatica continues to deliver on a strong product strategy that resonates well with market demands for data management, iPaaS and "big data." Informatica continues to mature its Intelligent Data Platform (IDP) strategy with investments in CLAIRE — its artificial intelligence-powered engine — and the Enterprise Data Catalog (its data catalog offering) for metadata-driven integration. These provide clients with advanced capabilities in integration — including automated metadata discovery, lineage, impact analysis, bidirectional metadata sharing and artificial intelligence/machine-learning-based process automation and recommendations for self-integration and repair of integration flows.

- Active, bidirectional metadata. Informatica's ability to share metadata across its data integration tools and its other data management solutions ensures that organizations can adopt it as an enterprise standard. Reference customers indicated that the unified integrated platform, comprising all major data delivery styles, allows them to use Informatica across a broad range of use cases.

- Expansion of market presence and mind share. Informatica continues to expand its global partner network of more than 500 partners and multiple independent software vendor ecosystems, ensuring continuous availability of talent. Informatica continues to be the vendor most referenced by Gartner clients. It appears in more than 75% of the contract

review calls for data integration tools and features most frequently in competitive situations on our inquiry calls.

Cautions

- Customers have difficulty aligning use cases with product offerings. Prospective customers should take care to align their needs with the proposed solutions offered. Gartner inquiries indicate that in competitive situations there is confusion around Informatica regarding overlapping features. This complicates competitive evaluations even further as reference customers struggle to clarify the components they use in their enterprise. Clients report that name changes can result in redundant or unused products, without careful review of their portfolio. Informatica is simplifying its product messaging to address these concerns, which has initially resulted in another round of product name changes.

- Alternatives for accessing "data in place" as opposed to data virtualization. Gartner clients and reference customers for Informatica voiced concern over the technical capabilities of Informatica Data Services (IDS) as a robust data virtualization tool. IDS is used frequently by customers for broad data federation needs. However, reference customers request a more robust tool for advanced data virtualization capabilities (for example, including push-down processing, abilities to improve caching performance via massively parallel processing support, and a more advanced query optimization engine). Informatica is promoting Data Integration Hub, APIs, microservices, data pipeline development and other alternatives for data-in-place for this growing demand.

- Perception of high cost remains. Informatica made the bold move of introducing a subscription-first pricing strategy for all its data integration tools in May 2016. However, reference customers still report their reluctance over perceived high TCO and price points, often citing them as the reason for evaluating other competitors. Informatica is intent on fixing this perception through a continued focus on pay-as-you-go options, bring your own licensing (BYOL) and hourly pricing models on the AWS and Azure marketplaces. However, concerns about higher subscription price points, additional costs for connectors and "add-on" maintenance and support costs remain.

Information Builders

Based in New York City, New York, U.S., Information Builders offers the Omni-Gen data integration platform, which is composed of a central platform plus additional tools, including iWay Service Manager and iWay DataMigrator, and the iWay Universal Adapter Suite. Components may be bought separately. Information Builders' customer base for this product set is estimated to be more than 900 organizations.

Strengths

- Accelerating adoption with targeted use cases. Information Builders has recently completed a reorganization and restructuring of its go-to-market strategy, to effectively combine solution templates, use-case frameworks and an overall approach for accelerating delivery. While this may, at first, seem to be a more complex offering and strategy, it creates a

cohesive approach for customers. It also pits Information Builders directly against larger-solution competitors with a lower price point and a less demanding engagement model.

- Broad platform capabilities. With a broad range of connectors (iWay) that are also used by a wide range of its competitors, Information Builders can access a very large variety of sources and load just as many targets. In addition, API developer solutions are coupled with administration/management tools that are specifically positioned to manage both application integration and data integration needs. From a B2B perspective, the same administrative and management functions establish performance and completeness (quality) thresholds, and enable automated resilience along with automated notifications.

- Customer experience. Ease of use, breadth of data integration capabilities and customer support continue to be cited by reference customers as primary reasons for selecting and continuing with Information Builders. Additionally, reference customers and Gartner inquiry clients report that the value proposition is highly positive with a low cost and high levels of benefits for utilizing the platform.

Cautions

- Inconsistent presence in the market. Over the years, Information Builders has made money and remains viable. However, its continuous efforts to reposition itself make it difficult to expand its presence in the market. Its most effective lead generation continues to be by word of mouth and reference-based. In itself, this would be an enviable position, except that Information Builders' customer base is not large enough to leverage this approach into a greater market presence.

- Tool maturity and delivery challenges. Error handling, a difficult learning curve and lack of skilled personnel in the broader marketplace hamper Information Builders' reference customers in their ability to expand the use cases for implementation. Some organizations attribute both learning curve and skills issues to their continued use of the Focus proprietary language base with iWay DataMigrator. Reference customers reported difficulties during installation, without any apparent guidance for issues.

- Inconsistent support and solution delivery. Gartner's clients report that Information Builders' solutions are often architecturally sound, but are prone to uneven delivery. This has an impact on support when the design used to deliver a solution includes unsupported "work-around" practices. While specifically reported as an issue by its reference customers, there is an incongruity here because the primary reason that organizations give for selecting Information Builders is its willingness to adapt quickly to individual client delivery needs.

Microsoft

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

Strengths

- Relevant capabilities and TCO. Reference customers for Microsoft cite the overall low TCO, speed of implementation, ease of use and ability to integrate with other Microsoft SQL Server capabilities and Azure data services as the main reasons for choosing Microsoft over the alternatives.

- Productivity in data delivery-, process- and business-oriented use. The ability to redeploy SSIS artifacts to Azure Data Factory, plus an enhanced deployment option for SSIS on Linux, extends its hybrid deployment capabilities and supports the combined use of ETL, enterprise business workflows and data preparation. Plans for cohesive access of metadata and multicatalogs are aimed at supporting the active data profiling requirements of ad hoc and citizen integrators.

- Widespread tool presence and usage experience. Broad familiarity with the implementation of Microsoft technologies spurs usage of SSIS (available also on Linux) and, increasingly, that of Azure Data Factory. A broad range of choices in terms of community collaboration, training, and third-party documentation and guidance for deployment practices are reported as key points of value.

Cautions

- Maturity of integrated portfolio. While Microsoft has advanced in aligning the offerings of SSIS and Azure Data Factory for integrated use, these capabilities are still relatively unproven or unknown to buyers in this market.

- Perception of platform versatility. A large customer base to leverage and a strong focus on the enablement of Microsoft developers, have generated market perceptions that this vendor's data integration tooling has limited relevance in a non-Microsoft-centric environment. The recently enhanced SSIS, which is deployable on Linux, begins to address these challenges.

- Hybrid experience. Although users of Microsoft are, increasingly, offered a mix of cloud and on-premises deployment models, usage experiences reflect a need for easing the implementation and migration of data integration artifacts (processes) between deployment choices in order to enable HIP scenarios. Microsoft has shifted the customer delivery experience all the way from "boxed" software to PaaS, for other products, and it is time for customers to consider the same implications for data integration.

Oracle

Based in Redwood Shores, California, U.S., Oracle offers the following data integration products: Oracle Data Integration Platform Cloud, Oracle GoldenGate (OGG), Oracle GoldenGate Cloud, Oracle Data Integrator (ODI), Oracle Big Data SQL and Oracle Service Bus. Oracle's customer base for this product set is estimated by Gartner to be more than 11,000 organizations.

Strengths

- New appeal to business roles. In 2018, Oracle client purchasing decisions appear to have shifted to business users seeking to integrate governance and data quality requirements during data integration delivery. When this shift is coupled with the recent advances in Oracle Big Data Preparation, and links to both ODI and GoldenGate (utilizing Kafka), Oracle's data integration platform is more broadly appealing and is solving technical issues in the stack. Oracle's knowledge modules continue to be an effective way to decouple business logic from ETL code.

- Restructured pricing aligned with the market. Oracle has introduced a free trial component. Its customers have options for pricing by volume or compute capacity. Customers pay for "credits" that are utilized in an account management approach, which can be applied to the various Oracle products. Through a range of term-based, subscription-based, name-based and metered-based pricing options (even hourly on OGG and ODI), Oracle is improving its pricing flexibility.

- Synergies and enhanced coordination across broad technologies. Recognition of Oracle's diverse portfolio for addressing data integration and other data and application-oriented requirements (spanning data quality tools, MDM solutions, ESB, analytic appliances and enterprise applications) continues to fuel its appeal in deployment scenarios. Oracle also supports big data SQL, plus streaming in the cloud and on-premises, and both are supported with in-built data virtualization.

Cautions

- Cost model raised by users. Calls with Gartner customers on our contract review and inquiry services, together with ratings on Gartner Peer Insights, indicate concerns with the cost model for Oracle customers desiring more flexibility in license models and pricing options. This represents one of the biggest reasons why prospects choose the competition over Oracle during our contract reviews.

- Difficult to learn so you need a plan. Gartner clients report that Oracle's ODI, specifically, requires a closely aligned development environment for testing and that, overall, usage is hampered until significant experience has been gained with the solution. Many users report the documentation is not as complete as it could be. Developing an overall plan for your implementation early on, as opposed to fleshing it out as progress is made, is almost mandatory. Also, when coordinating ODI and OGG, it is necessary to develop robust testing plans at the integration points.

- Ease of use can be improved. Some users that have moved from other tools or platforms report the Oracle interface itself could be improved relative to the response times for monitoring jobs. Customers also report that the initial setup is difficult and the maintenance functions are complicated by a lack of suitably skilled personnel in the market.

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 Remote Data Sync, SAP Data Hub, SAP HANA, SAP Cloud Platform, SAP Cloud Platform Integration and SAP

Streaming Analytics. SAP HANA has SAP Data Integration and SAP Smart Data Access for data federation. SAP's customer base for this product set is estimated at more than 50,000 organizations.

Strengths

- Extensive functionality and usage. The breadth of capabilities available across SAP's data integration tooling supports a diverse mix of data delivery styles, and use cases of increasing complexity. The SAP Data Hub offering, spanning data, process, deployment platforms, personas and diverse application challenges, resonates with SAP's customers.

- Portfolio innovation. Customers value the linkage of SAP's data integration tooling and adjacent products — for application and data integration, data preparation, information governance and master data management — in combination, for solving complex problems in the SAP landscape. A focus on multicloud, and IoT- and AI-enablement enhances the relevance of its solutions for applications, data and analytics scenarios.

- Alignment to trends of digital business technology platform. SAP's investment in business transformation capabilities through "intelligent technologies," aims at a digital platform that combines data management technologies, active metadata, HIP and various enabling components — including serverless architectures and machine-learning-based self-healing — that distribute, optimize and recompose integration processing.

Cautions

- Integrated use of products. SAP continues to simplify seamless use between data integration tools and integrated deployment across its broader portfolio. However, making multiple tools work together across SAP's integration ecosystem is reported by its reference customers as being a challenge. SAP continues to address integrated deployment — through SAP Data Hub — to ease and accelerate the interoperability of its integration offerings.

- Procurement and cost sensitivity. Concerns about TCO are increasingly surfacing in organizations that are broadening their deployments of SAP's data integration tooling. Some of these concerns are caused by confusion about overlapping product functionality and features, which leads to the purchase of similar products and results in redundancy.

- Customer experience. Reference customers for SAP indicate concerns about the overall customer experience. Customers want better guidance in following best practices. They need solutions that ease product/feature navigation, their approaches to implementation, consistency of support services and broader availability of skills in the market.

SAS

Based in Cary, North Carolina, U.S., SAS offers the following data integration products: SAS Data Management, SAS Data Integration Server, SAS Federation Server, SAS/ACCESS, SAS Data Loader for Hadoop and SAS Event Stream Processing. SAS's customer base for this product set is estimated to be 14,000 organizations.

Strengths

- Customer loyalty and persistent usage. The flexibility of SAS tools in data integration projects of various scope and sizes lies in their extensibility in implementations and proven strengths for amassing, integrating and sharing data. Many SAS customers report high-quality engagements for professional services, support and maintenance with the vendor. The result positions SAS well to compete with the larger, more-established vendors in the data integration tools market.

- Alignment to innovation trends. Customers identify SAS's very good usability and its resonance with data and analytics trends as key strengths. For example, enhanced capabilities spanning highly interactive development interfaces and metadata-driven rapid integration of sources. Machine learning functions — to recognize data schemas and content patterns — have also been introduced to assist developers and support the reuse of processing components.

- Integration with broader data and analytics offerings. SAS's data integration tools benefit from advances in its related data and analytics portfolio. Prebuilt integration exists with master data, data governance, the Viya discovery platform, and data quality tools associated with the complex requirements of large enterprises. Deals for SAS data integration tools support analytics goals and data management as a strategic discipline, but organizations moving to the Viya integrated approach need a broader reference base to determine its success.

Cautions

- Use-case versatility. While achieving traction in an analytics-oriented scenario, SAS supports a narrower range of data integration use cases relative to the major providers in this market.

- Pricing model and price points. As implementation complexity grows and causes concerns about the high cost of SAS solutions, some customers have expressed a desire for more procurement and cost options, which SAS is addressing.

- Participation in a broadening market ecosystem. Market perceptions of SAS being driven to enhance capabilities targeted at its own developer base present competitive disadvantages, particularly for buyers whose data integration requirements involve open-source or non-SAS software. It is noteworthy that the alignment of SAS's strategy and offerings based on SAS Viya is moving toward making the SAS platform more inclusive, HIP-enabling and embeddable, when working in diverse ecosystems.

Syncsort

Based in Pearl River, New York, U.S., Syncsort offers DMX, DMX-h, DMX CDC and Ironstream for relational, nonrelational and legacy mainframe data integration needs. Syncsort's customer base for this product set is estimated to be around 2,200 organizations.

Strengths

- Robust tools targeted at basic needs. High-performance ETL capabilities applied to mainframe and Hadoop ecosystems represent the established position of Syncsort in this

market. Deployments of Syncsort emphasize data movement, ingestion of streaming machine and application data, and data transformation offloading from data warehouses through to Hadoop.

- Customer relationship. Syncsort offers high-quality customer engagement and many reference customers identify its technical support and their overall relationship with Syncsort as positives. With an established and loyal customer base, and an expanding range of strategic partners, Syncsort has a solid foundation on which to grow its market presence. Internally, Syncsort's key personnel are stable and bring continuity to the overall team.

- Broadening applicability for data management scenarios. Syncsort's acquisition of Trillium Software in 2016 indicated a roadmap for bringing data integration and quality/governance into a common focus. An improved ability for publishing resilient CDC, data lineage, and the use of on-premises and cloud deployment models extends Syncsort's versatility for data and analytics use cases. Publishing resilient CDC is supported from mainframe, IBM i and relational sources to the data lake, to cloud data stores and to streaming platforms such as Kafka.

Cautions

- Deployment skills availability. Prospects and reference customers expressed concern that successful Syncsort deployments heavily depend on a high degree of proficiency from both developers and data architects. Also, the lack of a Syncsort skills pool in the market presents challenges.

- Mind share and market positioning. Adoption of Syncsort's data integration tooling continues, though with a relative lack of mainstream recognition. Syncsort plans to extend the applicability of the toolset by optimizing integration and assurance of data in a common environment for the metadata models, data delivery and governance targeting enterprisewide requirements.

- Acquisitions could dilute market presence. The combined benefits of Syncsort together with its acquisition of other data management and integration tools (such as Trillium), are often not immediately understood by prospects during competitive evaluations. As a result, some customers are uncertain of the broader benefits when engaging with Syncsort. This vendor has set out to enhance its market messaging with new, more simple branding and a focus on partners.

Talend

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

Strengths

- Cloud, containers and leveraging open source. Talend has developed specific capabilities leveraging Docker deployment and Apache Beam, which is based on Google's Cloud Dataflow. Customers are leveraging DataStreams to deliver Lambda architectures for dissonant data speeds. At the same time, all of these technical capabilities are being used for cloud-to-cloud, cloud-to-premises and multicloud integration solutions. Talend also leverages technology partnerships with Cloudera Navigator and Hortonworks, for their integration with Apache Atlas.

- Adaptable capabilities in an aligned product set. Talend's portfolio, including data quality, MDM, and the recently added API management tooling, sets out to deepen synergies across data management- and application infrastructure-related use cases. Customers value the configurability of Talend's tools, which makes them flexible enough to adapt to the business requirements of data integration processes, and the availability of artifacts built by Talend's practitioner community.

- Maintaining a low-cost pricing "edge." Talend's pricing remains lower than that of other Leaders in the market; more than a marginal saving. As the price differential begins to erode, Talend is now competing on the completeness of its solution — and better ease of use and time to delivery — among the Leaders. Talend has long appealed to customers as a "lower price point" for enterprise-class tools. However, we have begun to receive information from customers that pricing is increasing, so we recommend performing a thorough review of proposals.

Cautions

- Continuing stability and new release issues. Gartner clients continue to report issues with version upgrades and implementation for Talend. There is some anticipation that many of these issues will be addressed by the broader availability of skills in the market. However, customers have significantly punished other Leaders in this market when the answer to product or platform issues (even minor ones) is paying someone to fix it.

- Still thin partner and market skills. Gartner clients still report difficulties in finding skilled developers. Talend initiated an expanded partner training program across 2017/2018 and added new delivery partners in Europe and Asia/Pacific. "Customer Success Architects" have been introduced to assist larger customers and complex integration efforts. Gartner clients report that significant progress in this area has resulted in improved availability. We believe partner and skills efforts will continue, but they still need to be a focus for Talend.

- Portfolio messaging challenged by customer experience. Some prospective customers, particularly organizations considering putting Talend's portfolio to enterprisewide use, indicate it is difficult to integrate the overall solution and that Talend does not adequately articulate the more evolved capabilities of its products or their synergistic use.

TIBCO Software

Based in Palo Alto, California, U.S., TIBCO Software offers the following data integration products: TIBCO Data Virtualization (TDV), TIBCO StreamBase, TIBCO BusinessWorks, TIBCO Flogo Enterprise, TIBCO FTL and eFTL, TIBCO Enterprise Message Service, TIBCO Cloud Integration and TIBCO Spotfire.

TIBCO Spotfire is included for its embedded data preparation and data catalog capabilities. TIBCO's customer base for this product set is estimated to be more than 5,000 organizations.

Strengths

- Brand recognition, global presence and large non-data-integration installed base. TIBCO has a large customer installed base in the related data and analytics markets, particularly in application integration, business process management and MDM. It has proven its ability to capitalize on this when marketing its data integration tools, although this marketing is usually aimed at IT-based end users rather than the business sponsors of integration. TIBCO's global presence — through its significant ecosystem of partners, value-added resellers, consultants and external service providers — ensures that availability of talent is generally not an issue for its clients.

- Acquired data virtualization for expanded delivery. TIBCO's recent acquisition of Cisco Information Server — (CIS), which was renamed TIBCO Data Virtualization (TDV) — has given it an opportunity to embed data virtualization capabilities into its data integration product portfolio. Data virtualization presents TIBCO with the opportunity to expand into new use cases. Coupled with TDV, TIBCO can execute on its "Connected Intelligence" strategy by connecting to data sources "in place" for data sharing.

- Targeted offering components leverage existing strengths. TIBCO continues to draw on its leadership in stream data integration, through continued investment in StreamBase (an ultra-high-performance stream-processing engine) and BusinessEvents (a rule-centric platform capable of detecting complex patterns). Strength in stream data integration, plus investment in Mashery (API management), TIBCO Cloud (iPaaS), TIBCO Enterprise Message Bus (message-oriented data encapsulation and movement) and the newly introduced Project Flogo, means TIBCO can position itself as an HIP provider. (Project Flogo offers machine and IoT data integration.)

Cautions

- Complex pricing and high maintenance. TDV continues to be called out by TIBCO's reference customers for complex pricing models that are difficult to scale. Reference customers for TIBCO also continue to cite higher-than-average maintenance and support costs for its integration products. TIBCO has introduced subscription-based cost models to its integration product portfolio, and allows customers to transfer their licenses across on-premises and cloud-based environments for hybrid integration scenarios. TIBCO has also started simplifying the packaging of TDV, to include adapters to popular data sources and applications along with production and nonproduction licenses, to address some of these challenges.

- Limited depth of coverage for bulk/batch data delivery. TDV can complement ETL/ELT by providing robust caching and querying, and posting datasets into staging tables (as a cached dataset). It can also provide a data service that an ETL system then uses as a source. However, TIBCO's data integration tools alone cannot address the full scope of bulk/batch data movement-related workloads.

- Lack of mind share in non-analytics-oriented data integration use cases. With the acquisition of TDV, TIBCO can now address popular use cases such as data migration, prototyping for batch data integration, interenterprise data sharing and support for big data (including data lake constructs). However, Gartner client inquiries about TDV are skewed toward analytics/BI-oriented use cases.

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

- TIBCO Software — Cisco Integration Server's intellectual property, customer base and business was acquired by TIBCO Software in 4Q17.

Dropped

- Cisco — Its Cisco Integration Server (intellectual property, customer base and business) was acquired by TIBCO Software in 4Q17.

Inclusion and Exclusion Criteria

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

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

- Vendors whose customer reference base cannot represent any mix of their products or administrate and orchestrate compatible technologies for use in at least three of seven key technical deployment styles will be excluded. These styles are: bulk/batch, data services bus (service-oriented architecture [SOA])-style deployments, message-oriented, replication, streaming/event, synchronization and data virtualization. We describe them (alphabetically with no reference to order of importance) as follows:

  - Bulk/batch — 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 bus (SOA) — 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). And, with 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.

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

- o 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.

- o Streaming/event — Data consists of datasets that follow a 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, which 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.

- o 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.

- o 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.

- 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 and data asset types (such as JavaScript Object Notation [JSON]).

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

- Data transformation support — It is no longer acceptable to support packaged capabilities for basic transformations only (such as data-type conversions, string manipulations and calculations). Data integration solutions now must support moderately complex needs 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, the ability to synchronize metadata across multiple instances of the tool. Also, an open metadata repository, including mechanisms for bidirectional sharing of metadata with other tools.

- Capability to deliver machine-learning-enhanced metadata discovery, and internal analytics to enhance human data management and integration requirements.

- User- or role-specific variations in the development interface that are capable of various workflow enhancement mechanisms. These mechanisms may include supporting templates, version modification (via internal library management or other mechanisms), 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 services deployments in production environments, as well as alternative or collaborating development environments.

- Data governance support — Ability to import, export and directly access metadata with data profiling and/or data quality tools, MDM tools and data discovery tools. Accepting business and data management rule updates from data stewardship workflows and sharing data profiling information with such tools is highly desired.

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

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

- Generate at least $30 million of their annual software revenue from data integration tools (perpetual license subscription or maintenance/support). Or, they must maintain at least 300 maintenance-paying/subscription-paying customers for their 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.

- 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.

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

Vendors 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.

- They focused on only a single vertical industry.

- They served only their own, internally managed data models and/or architectures (this includes tools that only ingest data to a single proprietary data repository). Or, they were used by a single visualization or analytics processing platform, or 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 tools market by using the following criteria:

- Product/Service. Core goods and services that compete in and/or serve the defined market. This includes current product and service capabilities, quality, feature sets, skills and so on. This can be offered natively or through OEM agreements/partnerships (as defined in the Market Definition section, or described below). Some consumers are prepared to accept less-capable products from many different suppliers and assemble them together on their own. Connecting data integration activities to data quality and governance-related capabilities (such as MDM) becomes an integral support for all use cases that can share high-quality data as well as lineage and nonlineage metadata, with runtime management and monitoring support. For broader spectrum solutions, the market has de-emphasized the product capability and emphasized the ability to break out pricing and components. 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, and the overall usability and consumption of the tools.

- Overall Viability. Viability includes an assessment of the vendor's overall financial health as well as the financial and practical success of the business unit. It views the likelihood of the

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

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

- Market Responsiveness/Track Record. Ability to respond, change direction, be flexible and achieve competitive success as opportunities develop, competitors act, customer needs evolve and market dynamics change. This criterion also considers the vendor's history of responsiveness to changing market demands. Market track record is itself one measure of market responsiveness, and in this case data integration tools are much like other infrastructure-focused solutions. Often, organizations demand data virtualization, message-oriented data movement, replication and synchronization, and streaming/event processing. Traditional bulk/batch processing is still the predominant demand. Not only do most solutions overlap, but the market is demanding a capability to deliver all forms of integration to differently skilled implementers with everything from simple data preparation through self-service data integration to enterprise-class systems. The degree to which the vendor has demonstrated the ability to respond successfully to market demand for data integration capabilities over an extended period, and how well the vendor acted on the vision of prior years, are 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 and sales activities. Marketing execution was traditionally considered to be the positioning and declarations of a supplier, but now end-user organizations use it frequently as a gauge of how in-tune supplier roadmaps are with overall market demand. Suppliers need to be aware of emerging best practices for data management infrastructure and if they and their customers can specifically benefit from specialized horizontal or vertical capabilities, geographically targeted approaches or partner-supported implementation

practices. The overall effectiveness of the vendor's marketing efforts — which impact its mind share, market share and account penetration — is important. The ability of the vendor to adapt to changing demands in the market by aligning its product message with new trends and end-user interests was scored as part of the evaluation.

- Customer Experience. Products and services and/or programs that enable customers to achieve anticipated results with the products evaluated. Specifically, this includes quality supplier 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 customers' experience. The level of satisfaction expressed by customers with the vendor's product support and professional services is evaluated. As are 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 2018 Data Integration Customer Reference Survey.[1]

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

Table 1: Ability to Execute Evaluation Criteria

AWS	Amazon Web Services
BI	business intelligence
CDC	change data capture
DBMS	database management system
ESB	enterprise service bus
ETL	extraction, transformation and loading
HIP	hybrid integration platform
iPaaS	integration platform as a service
IoT	Internet of Things
MDM	master data management
ODI	Oracle Data Integrator
PaaS	platform as a service
SaaS	software as a service
SI	system integrator
SMB	small and midsize business
SOA	service-oriented architecture
SSIS	SQL Server Integration Services (Microsoft)

Source: Gartner (July 2018)

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 tools market by using the following criteria:

- Market Understanding. Ability to understand customer needs and translate them into products and services. Vendors that show a clear vision of their market will listen 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 their customer base. This criterion covers the alignment of the vendor's sales model with the ways in which customers' preferred buying approaches will evolve over time. Scaled pricing models are becoming particularly interesting. Suppliers must consider if their internal compensation models incentivize delivery that matches customer demand and implementation profiles. Customers and prospects are less concerned with "positioning" than they are with "try then buy" models. This increases the demand in the market for limited freeware versions that can be easily converted to robust solutions once proven.

- Offering (Product) Strategy. An approach to product development and delivery that emphasizes market differentiation, functionality, methodology and features as they map to current and future requirements. Existing markets and use cases have begun to weaken in favor of more distributed data integration needs — which increases the demand for self-healing and wizards/tutors for recognizing new sources and information asset types. Product strategy vision includes the roadmap for continued support of traditional integration needs — filling current gaps, weaknesses and opportunities to capitalize on less-advanced demand trends in this market. In addition, given the requirement for data integration tools to support diverse environments for data, delivery models and platform-mix perspective, we assess vendors on the degree of openness of their technology and product strategy.

- Business Model. The design, logic and execution of the organization's business proposition to achieve continued success. A visionary business model will balance the emerging (and increasingly stringent) demand for managing internal and external compliance and risk while providing support for existing customers. While broad, all-inclusive models represent one solution approach, it is also 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 tools market — including diversity of delivery models, packaging and pricing options, and partnership — is important.

- Vertical/Industry Strategy. The strategy to direct resources (sales, product, development), skills and products to meet the specific needs of individual market segments, including

verticals. This is the degree of emphasis the vendor places on vertical solutions, and the vendor's depth of vertical market expertise.

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

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

Table 2: Completeness of Vision Evaluation Criteria

Enlarge Table

-

Evaluation Criteria

Market Understanding

Marketing Strategy

Evaluation Criteria

Sales Strategy

Offering (Product) Strategy

Business Model

Vertical/Industry Strategy

Innovation

Geographic Strategy

Source: Gartner (July 2018)

Quadrant Descriptions

Leaders

Leaders in the data integration tools market are front-runners in the capability to support a full range of data delivery styles. In 2018, they have advanced their metadata capabilities, with some introducing highly dynamic optimization and advanced design assistance functions. Leaders have recognized the growing affinity between data and application integration. They are approaching location-agnostic deployments in a halting manner. Providers or solutions not limited to only cloud or only on-premises that can be deployed beyond a specific location are gaining traction (that is, on devices, as agents/daemons on sensors, in the IoT and more). 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 mind share in the market, 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, and by identifying new types of business problem to which data integration tools can bring significant value. Examples of deployments that span multiple projects and types of use case are common among Leaders' customers. Leaders have an established market presence, significant size and a multinational presence — either directly or through a parent company.

Challengers

In 2018, the Challengers have made significant strides in understanding that cloud and data virtualization are no longer differentiating features, and even as Challengers they are expected to have good executable technology. Challengers include previous Visionaries 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 an established presence. 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 new 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.

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 2018, our Visionaries have taken an early focus on alternative go-to-market strategies or on specific capabilities that capitalize on their capacity to leverage open-source, cloud or the growth in demand for connectors and API/microservices. Our lone Visionary is doing all of these things. Additionally, the development of specific solution designs that mix repeatable delivery templates with appropriate custom-deployed options represents a 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 a single application domain. Visionaries may also fail to provide a comprehensive set of product capabilities — including those that focus on a single data integration style and simply import, export or leverage that primary functionality to create alternative outputs for transformed or integrated data. They may be new entrants lacking the installed base and global presence of larger vendors, although they could also be large, established players in related markets that have only recently placed an emphasis on data integration tools.

Niche Players

As a mature market in 2018, Niche Players here generally don't exhibit gaps in primary market functionality and features. Instead, Niche Players simply either are 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 Niche Players have very strong offerings for a specific range of data integration problems (for example, a particular set of technical environments or application domains), and deliver substantial value for their customers in the associated segment. Niche Players now exhibit particular advantages in pricing, 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.

Importantly, more than 80% of all end-user organizations still seek bulk/batch processing (even to and within the cloud).[2] This means that a highly efficient but batch-oriented vendor can exhibit high-level execution capabilities without ever crossing to the right-hand side of the Magic Quadrant. Niche Players all exhibit batch capabilities — from back-office and operations data through to massive volumes of sensor or IoT data. Traditional Visionaries and Leaders may regress in their Completeness of Vision during the coming years — as they focus on existing demand and maximizing revenue relative to the current market, and maximize delivery in their strongest areas of the market.

Context

Before 2018, we started to see citizen integrators creating more data silos, more competing data stores and a devolved data integration tools market. We hope this will be a temporary situation and that more machine-enabled development and deployment options will fill the skills gap. Despite this regressive behavior, self-service tools are appearing in abundance in response to a belief that "anyone can integrate," when the consideration ought to be whether everyone *should* integrate. Data is always used in adjacent cases and context, and this is becoming the theme in 2018.

Data has now permanently overwhelmed the physical environment that is primarily composed of processing and storage capacity. In 2017, the pace of information capture was already beyond the market's capacity to process the resulting data volumes without some form of machine learning. "Pseudo" artificial intelligence solutions are taking advantage of consumer ignorance by marketing machine-learning-enabled decision engines as "artificial intelligence." The market has embraced this inaccurate, but appealing, message. Previously, there was a pendulum swing about every four to five years — between abundant capacity and a constrained infrastructure that required better integration processing designs. This pendulum swing, spanning the past three decades, has now come to an end. Cloud providers are encountering the same connectivity and management issues that on-premises providers encountered two decades ago. We have to get smarter, and that means our tools have to help us.

The era of supercomputers in information technology has long since faded and the era of hybrid environments is in full swing, but the era of metadata is just beginning for all information technology professionals. There will be analysis of metadata, machine learning against metadata, and graph analytics over metadata; and not just for technical metadata. Use-case metadata, context metadata, multiple contextual models within different versions of metadata, and even inferred metadata from analyzing other surrounding and adjacent metadata will permeate solution approaches. Enterprises pursuing frictionless sharing of data *must* have metadata capabilities that far exceed simple documentation. Metadata analytics will now issue dynamic instructions to processing engines. Microservices that specifically collect metadata will begin to emerge. This metadata "boom" will power tools that can be designed once for delivery across multiple platforms, mixed architectural approaches and broad deployment, without significant rework.

Throughout the market, the pattern of deferring data integration to post-application deployment will grow the "technical debt." That debt *must* be serviced at the time when applications and device apps are integrated — there is no option to leave it "as is" at that point. This effectively generates a new maintenance point, with a cost that is specifically deferred from the two anchor points (or applications) on either end. Even in cases where a centralized integration architecture (not necessarily centralized data) is attempted, the design and deployment *must* reach into each data location, which both incurs cost and consumes staff time. All technical debt must be serviced, without option; it simply becomes a question of when and how. Data integration tools are one means of servicing that data management debt; better design upfront is the other end of the continuum. Since there *are* existing applications, no one organization will ever succeed at the "design upfront" of the continuum. That means data integration is never going away.

The demand for easily deployed basic functionality is accelerating with data preparation demands; some suppliers will be tempted to, and will, specifically pursue "down market" sales that have smaller margins. From a broader data integration perspective, data preparation is considered a submarket that crosses between data integration and BI and advanced analytics tools (as well as some aspects of data quality and data profiling). We do not evaluate data preparation solutions in this Magic Quadrant (see "Market Guide for Data Preparation"). When market demand changes and vendors continue to only enhance established functional requirements, they may position targeted offerings to their advantage even though the margins go down, because their customer counts go up. Being a Challenger in this scenario actually becomes a market advantage, if properly pursued.

Despite all these advances, there is still room in the market for specialists with small footprint solutions and a specific focus to become the best at doing one thing. It has become increasingly difficult for those vendors hoping to develop new approaches as Visionaries to emerge in the market.

Many reference customers still report issues with TCO, and there remains a significant level of willful ignorance in the market — which blames the user organizations for this problem. However, the cloud capabilities are easing this cost issue, because data integration providers are now becoming SaaS vendors as well as PaaS providers. Automated discovery and machine learning can provide assistance — within tools and platforms — to address some portion of this cost, but the software is then a higher value proposition and therefore of a higher price. Application development teams need to recognize that some data must always be integrated; some data is integrated some of the time; and some data does not need to be integrated at all. They must start planning and designing for all three of these rings of information governance, or they are contributing to their own high TCO for data integration through an irresponsible approach.

Market Overview

Data integration is a persistent practice that is not specifically focused on the market as it is today. The presence of legacy, resilient systems and innovation all in the same market, requires robust, consistent delivery of highly tradtional practices as well as emerging or innovative approaches. Gartner's view of the market likewise includes transformational technologies or approaches delivering on the future needs of data and analytics leaders.

The biggest change in the market in 2018 is the shift from an anticipated future demand for metadata-driven solutions to a current market expectation that these solutions will be delivered as

part of the data integration platform. There is currently no forgiveness for inadequate metadata capabilities. Another major shift is customers asking for hybrid deployment (cloud and on-premises), as in 2017, but now with the expectation of multicloud and cloud-to-cloud integration.

Continuing in a longer trend, organizations expect to have the capability to blend data integration with application integration platforms, and are significantly increasing their own evaluation of whether their supplier's focus on product and delivery initiatives supports these demands. In continuation from 2017, organizations are seeking solutions that facilitate role-based data integration, which includes the capability to promote or manage the workflow of converting individually developed processes into enterprise-capable ones.

- The era of metadata has begun. As stated previously, metadata as a byproduct of design and operations management from a data integration platform is at a bare minimum in 2018. Platforms and solutions are now expected to provide continuous feedback regarding the profiles, quality, use cases, access points, context 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 will provide the necessary information for introducing machine learning capabilities into data integration platforms. 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 platforms. The requirements for more complete and better access to metadata underpin combining application and data integration practices. Often, the application or the data alone are incapable of absorbing the speed of change in operational process models. As a result, the influence of hybrid platforms for integration is rapidly gaining ground. That said, a new challenge has emerged in this space — in the form of microservices and their subsequent use in developing intelligent APIs. From a data management perspective, microservices attempt to create reusable data management logic that is encapsulated within a processing object. These objects can then be layered, or used as modular components, to build APIs. The current best practices for data integration are thus challenged in significant ways that will alter the landscape in a new world where point-to-point integration between data stores or even applications falls into disuse.

- Cloud is a concept — not a place — and data is expected to traverse the gap. This means that data is not in a place, but is a communication medium and a network of linked information assets. Data integration is expected to solve data issues by separating context from data and delivering both, side-by-side, to be recombined in use cases on a discretionary basis. Data is becoming cloudlike in a way that is more aligned with the first and best intent of the cloud. Before implementers/designers began referring to "cloud providers," they referred to cloud as the concept that different parts of your IT infrastructure could be designed to run anywhere and resources could be connected and disconnected from each other on demand. This is the "real cloud." Automobiles move around as mobile data centers with inherent processing, communication and data storage capacity — so does your phone and your tablet. Massive capacity on demand is available from so-called "cloud providers" (who really only provide certain solutions "in" the cloud). These components give a false sense of location for data that is not actually inherent in data architecture, but is part of the engineering approach

used when solutions are designed. Data integration platforms must now be able to design once, deploy many and optimize everywhere. Organizations expect data integration solutions to assist with three different data integration demands: cloud-to-cloud, cloud-to-ground and ground-to-cloud. When data gravity centers develop and create an imbalance between where the data is deployed and its management, and the need or demand for where the data needs to be used, data integration is expected to restore the balance. There are some important requirements separating traditional tolls from modern integration tools. For example, the Ability to Execute on data integration in a hyperconnected infrastructure (irrespective of structure and origins) through active metadata ingestion, sharing and processing; and the ability to recommend and autoexecute transformations through machine learning. This means that "hybrid" is an interim term. It will eventually give way to data integration approaches that dynamically reconfigure the optimization and delivery strategy regardless of deployment locations — even migrating themselves to new locations in the expanding, distributed computing infrastructure as we head toward 2025.

- Data integration is everywhere and is everyone's responsibility. Role-based development and deployment management is now incumbent on the data integration platform 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 the requirements for:

  o Machine learning that helps resolve when too much data that is beyond the capability for processing needs to be "trimmed" down dynamically and sized to both the use case and the capacity for utilizing that data. While massive processing models can be run on cloud platforms, connectivity, communications, validity, data freshness and, most importantly, the trust-rating of data sources will allow for dynamic, targeted data delivery to all manner of use cases.

  o Real-time recognition of the required speed of digital business. In the context of digital business, "business moments" — opportunities of short duration or a point in time that sets in motion a series of events involving people, business and things — are increasingly attracting the attention of enterprises. 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.

  o Diverse interfaces with the same design result. 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.

  o 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.

- o 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 continue to advance. Another aspect of this is how vendors/suppliers will choose to respond.

- o 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 the intersection of the two disciplines. 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.

Trends in 2018 continue from 2017, and organizations have begun to take three related pathways to integrating their data:

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

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

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

- Third, application and data integration are closely related and, with business processes changing rapidly, either application logic/design or data management/processing must always be the first line of adaptation.

In 2018, traditional integration is beginning to shift from the bulk/batch dominance for delivery in the market, but only slightly. More than 80% of all organizations in our survey make significant use of bulk/batch, but as much as 40% also utilize data virtualization, message queues or simple replication with layers of data processing afterward.[1] While many organizations have traditional solutions in place, modern demands have increased the utilization of message virtualization and synchronization. Somewhere between 35% to 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.[1] Distributed mobile devices, consumer apps and applications, multichannel interactions and even social media interactions are driving these organizations to build highly sophisticated integration architectures. Yet, these can just as easily be a simple data transfer protocol to a fully contextualized data service that delivers single data points through streams of information in near real time (see "Predicts 2017: Data Distribution and Complexity Drive Information Infrastructure Modernization").

Gartner estimates that the data integration tools market generated more than $2.9 billion in software revenue (constant currency) at the end of 2017. This is an increase of 6.9% from the end of 2016, and the market is expected to approach $4.5 billion in 2022 (see "Forecast: Enterprise Infrastructure Software, Worldwide, 2016-2022, 1Q18 Update").

Technology evolutions in this market to address shifting demand trends, and interim steps to combine cloud and on-premises deployments and broadening hybrid integration approaches, are well on their way to common and broad adoption by 2025 (see "The State and Future of Data Integration: Optimizing Your Portfolio of Tools to Harness Market Shifts").

Acronym Key and Glossary Terms

AWS	Amazon Web Services
BI	business intelligence
CDC	change data capture
DBMS	database management system
ESB	enterprise service bus

ETL	extraction, transformation and loading
HIP	hybrid integration platform
iPaaS	integration platform as a service
IoT	Internet of Things
MDM	master data management
ODI	Oracle Data Integrator
PaaS	platform as a service
SaaS	software as a service
SI	system integrator
SMB	small and midsize business
SOA	service-oriented architecture
SSIS	SQL Server Integration Services (Microsoft)
TCO	total cost of ownership

Evidence

[1] Gartner's 2018 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, 238 organizations across all major regions provided input on their experiences with vendors and tools in this manner. The survey was conducted during May and June 2018; the results were collated and analysis completed in June 2018.

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

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

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

- A web-based survey — Gartner 2018 Data Integration Customer Reference Survey — of the reference customers provided by each vendor (see above for details).

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

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

Note 1Detailed Components of the Evaluation Conditions

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

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

  - Relational databases

  - Legacy and nonrelational databases

  - Various file formats

  - XML

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

  - SaaS and cloud-based applications and sources

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

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

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

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

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

- o 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:

  - o Bulk/batch acquisition and delivery

  - o Granular trickle-feed acquisition and delivery

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

  - o 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:

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

  - o Data virtualization

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

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

  - o Replication of data between homogeneous or heterogeneous DBMSs and schemas

  - o 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:

  - o Scheduled batch delivery

  - o Streaming/near-real-time delivery

- o 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:

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

  - o 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

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

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

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

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

  - o Discernment of relationships between data models and business process models

  - o Data model creation and maintenance

  - o Physical-to-logical model mapping and rationalization

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

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

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

  - o Automated synchronization of metadata across multiple instances of the tools

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

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

  - o 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:

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

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

- Granular, role-based and developer-based security

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

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

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

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

- Support for testing and debugging

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

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

  - Data mining tools (relationship discovery)

  - Data quality tools (supporting data quality improvements)

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

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

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

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

  - UNIX-based environments

  - Windows environments

  - Linux environments

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

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

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

- o 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

- o In-memory computing environment

- o Server virtualization (support for shared, virtualized implementations)

- o 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:

  - o Error-handling functionality, both predefined and customizable

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

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

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

  - o 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:

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

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

  - o A common design environment to support all data delivery modes

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

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

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

- o 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:

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

  - o Management of publication and testing of data services

  - o Interaction with service repositories and registries

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

Evaluation Criteria Definitions

Ability to Execute

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

Overall Viability: Viability includes an assessment of the overall organization's financial health, the financial and practical success of the business unit, and the likelihood that the individual business unit will continue investing in the product, will continue offering the product and will advance the state of the art within the organization's portfolio of products.

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

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

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

Customer Experience: Relationships, products and services/programs that enable clients to be successful with the products evaluated. Specifically, this includes the ways customers receive

technical support or account support. This can also include ancillary tools, customer support programs (and the quality thereof), availability of user groups, service-level agreements and so on.

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

Completeness of Vision

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

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

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

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

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

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

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

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

By Mark Beyer, Eric Thoo, Ehtisham Zaidi