Magic Quadrant for Data Integration Tools

3 December 2024 - ID G00806653 - 85 min read

By Thornton Craig, Sharat Menon, and 3 more

The data integration tools market remains buoyant as organizations increasingly seek improved capabilities to support their operational, analytical and AI infrastructure use cases. In this research, we evaluate 20 vendors that compete to meet your organization's data integration needs.

Strategic Planning Assumption

By 2027, AI assistants and AI-enhanced workflows incorporated into data integration tools will reduce manual intervention by 60% and enable self-service data management.

Market Definition/Description

The data integration tools market comprises stand-alone software products that allow organizations to combine data from multiple sources, including performing tasks related to data access, transformation, enrichment and delivery. Data integration tools enable use cases such as data engineering, operational data integration, delivering modern data architectures, and enabling less-technical data integration. Data integration tools are procured by data and analytics (D&A) leaders and their teams for use by data engineers or less-technical users, such as business analysts or data scientists. These products are consumed as SaaS or deployed on-premises, in public or private cloud, or in hybrid configurations.

The common use case scenarios or business problems addressed by data integration tools include:

- **Data engineering —** Data integration by technical user personas to develop, manage and optimize data pipelines in analytical use cases
- **Delivering modern data management architectures —** Data integration to enable modern data management design patterns like lakehouse, data fabric, data mesh and deliverables like data products
- **Enabling less-technical data integration —** Data integration activities by less-technical user personas for various analytical demands of data, such as analytics and business intelligence (ABI), and data science use cases
- **Operational data integration —** Data integration to implement various operational data integration use cases, such as consolidation of master data, delivering and using data hubs, interenterprise and partner data sharing, and application integration

Mandatory Features

Data integration tools must:

- **Support extraction, ingestion and delivery of data using multiple common data integration styles,** including bulk/batch data movement, data replication and synchronization, stream data integration, and data virtualization. This feature includes the availability of out-of-the-box and configurable connectors to seamlessly access data sources and targets.
- **Support data transformation,** including basic (e.g., string manipulation, calculation), intermediate (e.g., data source merging, data aggregation) and advanced (e.g., complex parsing, text mining and multischema data modeling) data transformations. This feature can be delivered using prebuilt, reusable, configurable or custom components and pipelines.

Common Features

Common features for data integration tools include:

- **Data preparation features** to assist nontechnical users belonging to a wide range of business roles in their self-service data integration needs. This includes various techniques such as low-/no-code data ingestion, basic data modeling, data blending and visual exploration.

- **Augmentation features** that leverage generative AI (GenAI) and prepackaged ML algorithms to auto-generate data pipeline code and documentation, optimize data integration operations (e.g., anomaly detection, autorecovery), and use natural language to query as well as transform data.

- **Metadata management features** that support the extensive discovery, access, use and sharing of technical and operational metadata (e.g., usage data, transaction logs, system workloads) and business metadata (e.g., glossary), either through embedded data catalogs or exposing metadata to well-integrated external systems.

- **Data governance features** that assist data governance mandates (e.g., data quality, data lineage, policy enforcement, masking and annotation) while handling data for meeting specific use cases (e.g., master data management).
- **DataOps features** that support data pipeline operations, such as change management of data-related artifacts (e.g., Git integration of data pipelines, data model management), automation (e.g., automated testing) and orchestration of data delivery (e.g., continuous integration/continuous delivery [CI/CD] pipelines), with appropriate levels of security.

- **FinOps features** that enable D&A leaders to iteratively monitor, control and optimize spending, understand product performance and make choices regarding price-to-performance trade-offs for their data integration workloads, resulting in optimal allocation of resources.

Magic Quadrant

Figure 1: Magic Quadrant for Data Integration Tools

Gartner

Vendor Strengths and Cautions

Ab Initio Software

Ab Initio Software is a Leader in this Magic Quadrant; the same as last year. It is headquartered in Massachusetts, U.S. The company offers Ab Initio Data Platform as its data integration tool, which clients mainly use for their end-to-end data management needs such as data integration, data quality, metadata management and data governance execution.

Ab Initio's operations are geographically diverse, with clients across sectors including financial services, telecom and healthcare. Ab Initio is investing in several AI capabilities to simplify complex data management tasks.

Strengths

- **Enterprise focus:** Ab Initio's business model supports complex enterprise-scale data management use cases facilitating a combination of data integration styles (including batch, streaming, replication and virtual), data quality, metadata management and data governance.
- **Offering strategy:** Ab Initio's converged data management platform provides users with the capability to connect business terms (for example, credit score) to physical data elements across the enterprise. This is done through a knowledge graph that is modeled by leveraging large language models (LLMs), machine learning (ML) and natural language processing (NLP)-based technology.
- **Customer experience:** Ab Initio's direct customer relationship approach has resulted in high customer satisfaction with its account management, no-cost proofs of concept (POCs) and technical support, which is consultative in nature. Ab Initio is one of four recipients of the 2024 Customer Choice award for data integration tools, based on Gartner Peer Insights.

Cautions

- **Low penetration among SMB and departmental buyers:** Ab Initio lags behind its competitors' offerings for certain entry-level deployments — such as a stand-alone extract, transform, load (ETL) capability — and for departmental use cases. It addresses some of these challenges via "blueprints" (for example, automating data loading into a data lake, and masking sensitive data) and a low-code data exploration interface to assist less-technical users.
- **Operational complexity:** Ab Initio's platform setup and upgrade can be time-consuming. Coupled with the closed nature of the platform and steep learning curve for technical teams, achieving optimal platform operations may be challenging to customers. Ab Initio is investing in one-click cloud marketplaces, Ansible and Helm Chart deployment capabilities to further ease install and upgrade concerns.
- **High price perception:** In Gartner client inquiries, customers reported high license costs and complexity of managing Ab Initio license agreements. While Ab Initio offers different pricing models (including perpetual, on demand, subscription and unlimited usage), most customers move to an unlimited usage agreement to provide cost containment and predictability while allowing broad access to the software.

Amazon Web Services

Amazon Web Services (AWS) is a Leader in this Magic Quadrant; it was a Challenger last year. AWS is headquartered in Washington, U.S. and offers AWS Glue as its main data integration product. Other products that it provides for data integration tasks include: Amazon AppFlow, Amazon Managed Workflows for Apache Airflow (MWAA), Amazon Kinesis, AWS Database Migration Service (DMS), Amazon Managed Streaming for Apache Kafka (MSK), Amazon Managed Service for Apache Flink, Amazon EMR, and Amazon Athena.

AWS' operations are geographically diverse, with clients across various sectors. AWS is investing in unifying the UX for all of its data integration services, and adding more zero-ETL integrations.

Strengths

- **AWS-centric data ecosystem vision:** AWS provides seamless data access between services, labeled as zero-ETL, across many of its services such as S3, Amazon Redshift and Aurora. AWS Glue integrates with Amazon SageMaker for ML model inference, with services like Amazon Textract and Amazon Kinesis Video Streams for unstructured data extraction, and with AWS Glue Data Catalog and Amazon DataZone for metadata management. It also embeds Amazon Q for AI-powered assistance.
- **Strong vision for supporting multiple personas:** AWS Glue supports a code-based approach through notebook interfaces, a GUI-based approach through Glue Studio, a spreadsheet interface (Glue DataBrew) and natural-language interface (via Amazon Q integration) for data integration. AWS' vision is to provide a unified UX for building data-driven applications.
- **Serverless architecture:** On Gartner Peer Insights, customers report favorably on Glue's serverless architecture as it provides time savings on infrastructure management, efficient autoscaling at the job level, and the ability to limit idle time of compute clusters. This makes deploying jobs much easier than in a self-managed environment.

Cautions

- **High cost perception:** On Gartner Peer Insights, customers cite cost as a pain point while using AWS Glue, especially for high-volume data integration jobs. Difficulty in tracking data processing units may pose a challenge, especially for users accustomed to fixed

price models. Customers will need to use AWS' FinOps capabilities — for compute usage visibility and enforcing cost-control guardrails effectively — to keep costs under control.

- **Limited multicloud vision:** Despite AWS Glue's range of source connectors for non-AWS databases (such as SQL Server, Oracle, Salesforce, Snowflake and BigQuery), it provides limited non-AWS target connectors. AWS Glue lags behind the pure-play data integration vendors' offerings for hybrid, multicloud data integration needs because it is not deployed in other public clouds.

- **Challenging to use for more complex pipelines:** AWS Glue may pose a challenge to customers for more advanced data engineering use cases, which require advanced coding. Customers report via Gartner Peer Insights that there is scope for improved integration of Glue with Apache Iceberg, along with improved monitoring and performance management of Spark jobs, which make it challenging to use Glue for complex workloads.

CData

CData is a Niche Player in this Magic Quadrant; a new entrant this year. CData is headquartered in North Carolina, U.S. It offers CData Sync, CData Connect Cloud, CData Virtuality and CData Arc as its data integration tools. These tools are built on top of CData's extensive connector catalog, which it sells directly to enterprises and independent software vendors via OEMs. CData's operations are mostly focused in North America and Europe, with clients in various sectors including financial services, IT and professional services.

Clients predominantly use CData for its connectivity and replication solutions. Through the acquisition of Data Virtuality in April 2024, CData has expanded its data virtualization capability. It is also investing in GenAI and advanced data governance capabilities.

Strengths

- **Strong sales strategy and market responsiveness:** CData has a strong market presence for connectivity and replication solutions, covering over 3,000 source-target-specific use cases, like Workday to Snowflake replication. In addition to direct customers, CData provides embedded connectivity for several software vendors, including Salesforce, Atlassian, UiPath and Collibra.

- **Low total cost of ownership (TCO):** CData provides performant data integration products that allow its customers to scale for high-complexity workloads. Because clients don't

have to build or optimize these connectors as drivers or workbenches, this leads to more cost-effective creation and maintenance of data pipelines.

- **Data Virtuality acquisition:** CData's acquisition of Data Virtuality has extended its product coverage into data virtualization and its market coverage into EMEA. It now offers bimodal integration capabilities — data replication and virtualization.

Cautions

- **Offering strategy:** CData provides a collection of solutions for addressing different data integration use cases. Gartner client interactions reveal that CData's point solution approach to address different data integration use cases may not be a sustainable one for organizations seeking seamless end-to-end business processes.

- **Evolving market positioning:** CData recently transitioned to focus on providing a portfolio of data integration products, whereas traditionally it was considered a vendor focused on selling connectors. Although its data integration tools already make up the majority of the company's enterprise business, it is still in the early stages of evolving and positioning itself as a data integration tool provider.

- **Customer support:** Clients have reported via Gartner inquiry that CData's lack of efficient technical support and inadequate product documentation has led to greater reliance on customer support and longer resolution times. Customers noted the need to subscribe to CData Premium Support to mitigate these issues.

Confluent

Confluent is a Challenger in this Magic Quadrant; it was a Niche Player last year. Confluent is headquartered in California, U.S. It offers Confluent Cloud and Confluent Platform as its data integration tools. Its operations are geographically diverse, with clients primarily in the banking and financial services, healthcare and public sectors.

Through the acquisition of WarpStream in September 2024, Confluent has expanded its deployment options capabilities by adding support for private, single-tenant cloud deployment. It is investing in achieving product feature parity between Confluent Cloud and Confluent Platform, such as by adding Flink to the Confluent Platform. It is also investing in adding more connectors and advanced data governance capabilities.

Strengths

- **Strong stream data integration capabilities across multiple clouds:** Confluent provides strong stream data integration capabilities, with its support for Apache Kafka and Flink for real-time data movement and transformation respectively. Beyond Confluent Cloud for all major public clouds and Confluent Platform for on-premises deployments, Confluent also provides Cluster Linking for hybrid, multicloud stream data integration.

- **Significant data governance enhancements:** Confluent provides Stream Governance, which includes Schema Registry, Stream Catalog and Stream Lineage. Stream Governance offers capabilities such as schema validation and drift handling, business metadata tagging, sensitive data classification, field-level encryption, data quality rules, data lineage graphs, metric tracking at a topic level, and more. These capabilities help customers minimize integrations with stand-alone data governance platforms, thereby reducing latency for real-time use cases.

- **Improved support for modern data management architectures:** Confluent provides data lakehouse support through Tableflow, which converts Kafka topics into Apache Iceberg tables. It also enables users to search and explore Kafka topics as data products through Data Portal, and streamlines user collaboration across domain teams through Stream Sharing to support customers' data product initiatives driven through the data mesh strategy.

Cautions

- **Limited focus on nonstreaming data integration styles:** Confluent's products are rarely used for batch integration or data virtualization. Customers can now use Freight Clusters to offload data into data lakes in batch mode, but they should not otherwise use Confluent for traditional batch workloads. Confluent does not offer in-house connectors for log-based replication capabilities, but packages open-source Debezium connectors instead.

- **Lagging in augmented data integration vision:** Confluent lags behind its competitors in augmented and GenAI-enabled features in both its products and product roadmap. Although Confluent does provide an AI chat assistant and AI model inference within Flink, customers will not find optimized code generation or active metadata-driven automation in the products.

- **High cost perception:** A small but notable set of Gartner clients have indicated that the overall cost of Confluent is high and would prefer more granular billing. To keep costs

under control, Confluent Cloud customers must proactively utilize on-demand cluster sizing via elastic Confluent Units and monitor cloud spend using the Costs API. Additionally, customers have less expensive options from cloud service providers (CSPs) for basic Kafka-based workloads.

Denodo

Denodo is a Leader in this Magic Quadrant; the same as last year. It is based in California, U.S. It offers the Denodo Platform both on-premises and on all public clouds. Denodo's operations are geographically diverse, with customers primarily in the financial services, manufacturing and public sectors. Organizations predominantly use Denodo for data virtualization and data API services.

Denodo is investing in a fully managed SaaS solution called Agora, which simplifies deployments through the separation of control and data planes. The control plane will be hosted by Denodo and the execution plane will remain configured within the customer's domain.

Strengths

- **Strong brand awareness for data virtualization:** Denodo delivers modern data architectures enabled by data virtualization. It offers a strong data virtualization platform, with GenAI simplifying access to structured and unstructured data, improving resource management, and enhancing ease of use for both business and technical users.
- **Partnership growth and sales momentum:** Denodo continues to expand its consulting partners, in addition to its direct sales organization, reporting over 250 partners worldwide covering systems integrators (SIs), resellers, consulting, and more. Over the last year, Denodo has grown its revenue close to three times the average market growth for data integration software.
- **Customer experience:** Denodo is one of four recipients of the 2024 Customer Choice award for data integration tools, based on Gartner Peer Insights data. Customers cite product functionality and customer focus as its key differentiators.

Cautions

- **Physical data movement limitations:** Denodo's products are rarely used for bulk/batch integration or data replication, especially when there are strict performance SLAs. It does not offer log-based change data capture (CDC) and synchronization.

- **Management of distributed deployment:** Distributed Denodo deployments (with multiple instances at different locations) may pose challenges to customers around maintaining development standards, providing centralized support for troubleshooting issues, optimizing performance and securing the source systems. Denodo's fully managed SaaS service, Agora, will further address scaling challenges.

- **Vertical strategy:** Denodo does not offer domain-specific accelerators or industry-starter templates. The company relies on its SI partners to provide support for targeted verticals, as it builds its expertise in vertical presales, sales and implementation.

Fivetran

Fivetran is a Challenger in this Magic Quadrant; the same as last year. Headquartered in California, U.S., it offers Fivetran and Fivetran HVR. Its operations are primarily focused in North America and EMEA, with a growing presence in APAC. The company supports various sectors including IT, finance and insurance, and manufacturing.

Fivetran is investing in scaling its library of connectors by focusing on unstructured data sources and on-demand AI-generated connectors, in addition to simplifying the connector setup experience for its users.

Strengths

- **Sales execution and market momentum:** Fivetran continues to grow its revenue and customer base at a rate faster than the market. While a product-led growth strategy continues to drive SMB segment adoption, the company's focus on enabling large enterprises to standardize on Fivetran with new enterprise-level agreements in 2024 expands its reach beyond the midmarket segment.

- **Solid execution on connectivity and ease of use:** Fivetran currently provides over 580 fully managed connectors, and its "Powered by Fivetran" program enables its customers to embed Fivetran connectors into their own applications. A software development kit (SDK) enables partners to create custom connectors rapidly, and the company provides over 35 predefined, analytics-ready data build tool (dbt) models for data from common source systems like Salesforce. These capabilities facilitate nontechnical users to ingest data from multiple sources without coding.

- **Scalable and flexible pricing options:** Fivetran offers consumption-based pricing (in a tiered model of the "monthly active rows processed"), allowing its SMB customers to start

their projects without securing significant upfront capital expenditure. Customers also enjoy reduced unit costs with increased data volumes. Fivetran offers enterprise license agreements for price predictability. It further supports its customers with free-of-charge, one-time bulk ingestion of data into the cloud.

Cautions

- **Limited data transformation capabilities:** Most of Fivetran's customers rely on dbt Core for data transformation. Fivetran only provides basic in-flight transformations for streaming data, such as normalization and deduplication. It does not offer data virtualization capabilities. Fivetran has introduced connector-specific, built-in dbt data transformations as Quickstart models to help address this gap.
- **Basic metadata and governance support:** Fivetran does not offer metadata analysis/activation support. Instead, the Fivetran Platform Connector exposes metadata to third-party solutions — for example, vendors like Collibra and Alation — to support these capabilities. Fivetran provides basic operational metadata in the Fivetran Platform Connector for logs and auditing, and basic data governance capabilities, such as role-based access control (RBAC), data encryption and column blocking/hashing.
- **Limited support for modern data management architectures:** Fivetran lags behind its competitors in supporting data fabric architecture and data mesh delivery. Its roadmap does not provide for semantic modeling and delivery through data products. Additionally, its GenAI capabilities are limited to integration with Snowflake Cortex AI. Fivetran's support for modern lakehouse architectures is still evolving, with the introduction of Fivetran Managed Data Lake Service earlier this year. This service automatically converts data into open-table formats and facilitates interoperability with natively integrated data catalogs.

Google

Google is a Leader in this Magic Quadrant; last year it was a Challenger. It is headquartered in California, U.S. Google's product suite includes: Cloud Data Fusion for cloud-native data integration; Dataplex for data governance; Dataflow for streaming; Datastream for change data capture and replication; and Cloud Composer for orchestration. Google's operations are spread globally, and it supports customers across all industries.

Over the past year, Google has been working hard to incorporate more AI-assisted features into its products. This is expected to continue through next year, with offerings such as Gemini-powered data preparation and data integration, which supports generating SQL queries from natural language.

Strengths

- **AI-aided workflows:** Google's data integration strategy focuses on AI-powered data integration, with AI integrated into the developer experience. Google's Gemini provides AI-enabled data pipeline generation and data migration to its customers.
- **Data governance and metadata in support of data fabric:** Google offers centralized, unified governance with features for autodiscovery, semantics and end-to-end data lineage. Through Dataplex, it provides customers with the functionality to implement attribute-based governance based on metadata, enabling governance at scale.
- **Developer experience:** Data engineers report that Google's products are easy to use, relative to other vendors in this Magic Quadrant. Google offers comprehensive product documentation, and users report high satisfaction. Users also report that costs are reasonable.

Cautions

- **Products are Google-centric:** Google primarily focuses on delivering data integration tools for use within its own platform, and customers who are not fully invested in its ecosystem may face challenges. Customers not already using Google Cloud Platform will need to leverage its partnerships with other companies to extend those capabilities beyond the Google Cloud ecosystem.
- **Market positioning:** Google operates as a cloud service provider, competing more with other hyperscalers than with specialized data integration vendors. Lower market awareness makes it harder for customers to find employees who are experienced with Google's data integration tools, requiring clients to leverage third-party global and regional SIs.
- **Complex product portfolio:** Google's data integration product portfolio is not unified. Customers wanting to use more than one data integration style may need to use several tools, though Google provides integrations between these tools and has a roadmap to unify some of these capabilities.

IBM

IBM is a Leader in this Magic Quadrant; the same as last year. It is headquartered in New York, U.S. IBM supports data integration through the following products: IBM DataStage, IBM Data Replication, IBM StreamSets, and IBM Cloud Pak for Data. IBM has a diverse customer base, mostly in North America, but also an extensive global presence, primarily in Europe and APAC.

IBM offers a broad data integration portfolio, including AI optimizations using watsonx. In June 2024, IBM expanded its streaming data integration capabilities through the acquisition of StreamSets from Software AG.

Strengths

- **Data integration vision:** IBM provides data integration tooling options that meet most modern data integration requirements; for example, support for multiple integration use cases, inclusion of unstructured data, hybrid deployment models, emerging trends like data fabric, active metadata utilization for integration, and enabling data products. IBM also supports pipeline creation, automation and optimization by leveraging its AI capabilities in watsonx.
- **Global presence and strategy:** IBM has a presence in over 175 countries globally, and maintains strong relationships with over 1,600 partners including consulting firms, SIs and value-added resellers. It supports critical business use cases through scalable, modular and integrated capabilities across infrastructure and hybrid cloud. These capabilities include AI assistants, retrieval-augmented generation (RAG) support for GenAI, and modern AI and data platform designs like data fabric.
- **Enhanced streaming capabilities:** With its recent acquisition of StreamSets, IBM provides comprehensive real-time data ingestion, processing and delivery of data for complex data engineering pipelines in hybrid, multicloud environments. IBM plans to strengthen the streaming connectivity of IBM StreamSets in its near-term roadmap, leveraging its strong CDC technology, along with lineage and cataloging through IBM's modular approach to a data fabric.

Cautions

- **Driving new workloads:** IBM struggles to win new customers at market rates compared with other vendors in this Magic Quadrant. IBM does not come up in many competitive situations and is losing out to more specialized vendors or CSPs. It does, however,

continue to have a strong presence with enterprise clients and invests aggressively in product-led growth and go-to-market strategies, offering native data integration solutions.

- **High cost and solution complexity:** Clients report, via Gartner inquiry and Peer Insights, that IBM's data integration products are priced higher relative to other vendor solutions for comparable use cases. IBM's tool deployments involve complex infrastructure and container platform setups to meet enterprise governance, performance and scalability requirements. This may pose a challenge to customers, but may also lead to greater flexibility in solution deployments.

- **Cloud migration projects:** IBM DataStage customers report a lack of clarity on license portability and associated best practices around migration to IBM Cloud Pak for Data. IBM supports a modernization path for existing solutions and offers built-in, no-cost migration utilities, including partners and programs to help clients considering cloud migration from existing deployments.

Informatica

Informatica is a Leader in this Magic Quadrant; the same as last year. It is headquartered in California, U.S. Its Intelligent Data Management Cloud (IDMC) is a comprehensive data management platform; however, this Magic Quadrant focuses on its data integration components. Informatica's operations are geographically diversified, and its clients are typically large enterprises across various industries.

Informatica focuses on evolving its products to help customers enhance AI-readiness by enabling AI-powered metadata intelligence and automation across its product set.

Strengths

- **Strong use of metadata that supports a data fabric vision:** In Gartner client inquiries, customers report favorably on Informatica's AI-powered engine, CLAIRE, which utilizes and analyzes metadata (for example, usage, translation logs, workloads) for intelligence and automation. Informatica offers a comprehensive data fabric vision, utilizing metadata throughout the data life cycle for all data management tasks, including ingestion, transformation and automated generation of data pipelines.

- **Vision for AI-ready data:** Informatica demonstrates a clear vision and understanding of customer requirements for AI-ready data. It automatically catalogs and classifies both structured and unstructured data, and performs unstructured data chunking and

embedding to enable RAG to ground and enrich LLMs through organizational data integration.

- **Mature product portfolio:** Informatica's product portfolio supports a large number of data sources and targets, covering all use cases and styles of data integration. Informatica even provides support to customers with highly specific needs such as FedRAMP certification, heightened privacy/security requirements, and support for uncommon data sources. Customers seeking a robust data integration solution should evaluate Informatica.

Cautions

- **Slower growth than competitors:** Informatica is experiencing slower growth compared with other vendors in this Magic Quadrant and is seeing a relative decline in market share. Client inquiries indicate a general shift toward using native data integration capabilities provided by CSPs, presenting an ongoing challenge to Informatica's growth trajectory.
- **Lingering migration challenges:** Many Informatica customers still use the company's older product offering, PowerCenter, which limits their access to the latest features in IDMC. The migration process to IDMC can be expensive and time-consuming for customers with complex PowerCenter deployments, so it requires careful planning.
- **Consumption-based pricing model is challenging for some customers:** Informatica's consumption-based pricing model — where customers purchase Informatica Processing Units (IPUs) — may not provide substantial advantages to customers who only want to use one part of the product. Customers who use multiple services within IDMC appreciate the flexibility, but still report difficulty forecasting future IPU usage. Prospective clients with limited budgets should rigorously assess Informatica's offerings, and all customers are advised to fully leverage Informatica's FinOps capabilities.

K2view

K2view is a Visionary in this Magic Quadrant; the same as last year. It is headquartered in Yokneam, Israel. It offers the K2view Data Product Platform as its complete data integration tool. K2view's customer base is primarily in North America and Europe. The company focuses on leveraging its Micro-Database technology to enable multiple data integration patterns and use cases.

Strengths

- **Entity-based approach to data integration:** K2view offers a unique approach to integrate and organize data by business entities — such as customer, product and order data — through its Micro-Database architecture. The K2view platform manages millions of Micro-Databases with near-real-time data movement, including the ability to apply in-flight data transformations, all accessed via web service APIs, messaging, CDC, JDBC and streaming integration methods.

- **Strong GenAI support:** K2view provides functionality for preparing and delivering AI-ready data into RAG-based GenAI apps through its GenAI Data Fusion offering. K2view accelerates data pipeline creation by enabling users with GenAI assistants to create and manage pipelines. It also leverages GenAI to generate synthetic data for software testing and ML model training.

- **End-to-end operational pipeline focus:** K2view offers capabilities to design, load and operate Micro-Databases to manage the delivery of data flows such as master data assets. These include multidimensional golden records, hierarchical data, entity matching rules, data classification, governance and exception-handling interfaces.

Cautions

- **Low market visibility and mind share:** K2view's market visibility lags behind incumbent vendors in this Magic Quadrant. Moreover, despite its continued efforts and progress in ramping up global SI partnerships, expertise in the market for its solutions is low.

- **Potential high TCO:** K2view is not a low-cost data integration tool suited for basic ETL use cases. It is more suitable for high-complexity, low-latency data integration use cases. K2view's consumption-based pricing model may enable organizations to start small and grow additional business entities and use cases, leveraging granular monitoring and built-in growth to help manage costs.

- **Pipeline complexity:** Customers report, via Gartner inquiry and Peer Insights, challenges with managing advanced K2view pipelines, which can become complex and are more suitable for data engineers than less-technical users. To address these challenges, K2view released its GenAI assistant to simplify the process of creating and managing complex pipelines.

Matillion

Matillion is a Challenger in this Magic Quadrant; the same as last year. It has dual headquarters in Denver, Colorado, and Manchester, U.K. Its data integration product includes Matillion ETL, Data Loader and Data Productivity Cloud. Matillion's customer base is broad, with organizations primarily in North America and Europe.

Matillion is focusing on expansion beyond its core customers with its recently launched cloud offering, Data Productivity Cloud, which targets more integration use cases and deployment options.

Strengths

- **CDC leadership and selection in competitive situations:** Matillion enables strong support for common CDC-driven use cases and integration patterns. It is evaluated positively in competitive situations for quick deployments and strong integrations with CSPs and cloud DBMS tools.
- **Enhanced GenAI support:** Matillion offers internal GenAI-integrated assistants for enhanced creation and management of data pipelines. It also allows clients to build LLM pipelines using patterns such as RAG, with common connectors to vector databases and AI providers.
- **Market and cloud momentum:** Matillion focuses on expanding its partnerships with popular data management vendors, CSPs and global SIs. Data Productivity Cloud continues to experience growth as a multicloud offering, supporting various deployment options and integration patterns.

Cautions

- **DevOps challenges with initial stages:** Practitioners report struggling with Matillion's DevOps administration on Matillion ETL, especially during the early stages of deployments. DevOps platform integration can also be more complex than expected, leading to pipelines requiring in-house training activities to support users. Matillion's Data Productivity Cloud reportedly resolves this, as it is SaaS with APIs at the core.
- **Custom Connector API challenges:** Multiple customers identified the Custom Connector within Matillion ETL as having scaling issues related to API consumption. The cloud offering reportedly resolves this, but clients should thoroughly validate current deployment requirements and any known capability gaps.

- **Lacks advanced data virtualization:** Matillion's data virtualization capabilities are limited to external table objects and pushdown extract, load and transform (ELT) queries. It lags behind its competitors' offerings in providing support for more advanced integration patterns such as query acceleration, dynamic query optimization and caching.

Microsoft

Microsoft is a Leader in this Magic Quadrant; the same as last year. It is headquartered in Washington, U.S. It offers Azure Data Factory (ADF), SQL Server Integration Services (SSIS) and Power Query as its data integration tools, and Microsoft Fabric as its main data platform, which also includes data integration services. Microsoft's operations are geographically diversified, and its clients are primarily in the banking and financial services industry (FSI), energy and healthcare.

With the goal of making Microsoft Fabric a unified data and analytics (D&A) platform, Microsoft continues to develop its individual data integration functionalities at a sustained pace, expanding its data integration styles and improving its augmentation and governance features.

Strengths

- **Market momentum:** Microsoft released Microsoft Fabric for general availability in November 2023. Since then, it has seen significant adoption from both existing clients that were using Microsoft's other data integration tools, and from new customers. This growing adoption is reflected in the number of times that Microsoft Fabric's data integration services have been considered in competitive scenarios, as an alternative to independent software vendors.

- **Broad data ecosystem vision:** Microsoft Fabric leans toward a converged D&A ecosystem. Its data integration services — for example, Data Factory, Data Engineering and Real-Time Intelligence — are well-integrated with other components such as OneLake and Power BI. Microsoft offers a robust product vision by focusing on constantly expanding its data integration capabilities.

- **AI-powered capabilities:** Microsoft invested in augmentation capabilities through its coding assistant, Microsoft Copilot, which supports both technical and nontechnical users for data engineering and self-service data integration activities respectively. Microsoft provides augmentation capabilities near the UI components in its products, enhancing the user experience.

Cautions

- **Vision less focused on hybrid cloud and multicloud:** Microsoft's vision is primarily centered around its own D&A platform and less on more distributed architectures. This can be limiting for customers who are oriented toward hybrid cloud and multicloud architectures and are not willing to shift their data gravity toward a single platform.
- **More often considered for simpler scenarios:** Gartner client interactions reveal that Microsoft's data integration tools are evaluated in fewer competitive scenarios that involve complex data architectures. While offering all data integration styles, Microsoft's data replication and virtualization capabilities are still immature. To mitigate this limitation, Microsoft is investing in these capabilities through its Mirroring and Shortcuts features.
- **Gaps in supporting capabilities:** Users report low satisfaction with Microsoft's data integration tools support capabilities. For example, they note that FinOps capabilities for managing spending and metadata management capabilities are not completely meeting their needs.

Oracle

Oracle is a Leader in this Magic Quadrant; the same as last year. It is headquartered in Texas, U.S. It offers Oracle Cloud Infrastructure (OCI) as its main data platform, which includes data integration services, and offers Oracle GoldenGate and Oracle Data Integrator (ODI) as its data integration tools. Oracle's operations are geographically diversified, with a greater presence in the Americas. Its clients are primarily in banking and FSI, retail and telecommunications.

Oracle is investing in a multi- and hybrid-cloud-oriented data ecosystem focused on both operational and analytical workloads. This includes all data integration styles and provides an augmented experience by leveraging AI capabilities.

Strengths

- **Market vision for multicloud deployments:** Oracle leans toward a cloud-agnostic data integration approach for creating a multicloud data ecosystem. It positions OCI as a central hub between different CSPs by maintaining partnerships with CSPs and investing in metadata-sharing cross-clouds and FinOps capabilities.

- **Support for complex and distributed data architectures:** Oracle's data integration tools offer robust support for complex integration scenarios, including hybrid architectures that require integration of cloud and on-premises systems. Oracle provides support to more than 50 regions around the world and offers multiple technologies to integrate with private data centers and edge configurations securely.

- **Robust operational data integration:** Oracle GoldenGate is a leading tool in real-time data replication, capable of performing low-latency data synchronization. Users often leverage it to implement high availability and disaster recovery in their organizations. Oracle also provides robust stream-processing capabilities, and its focus on ensuring transaction integrity is ideal for mission-critical and operational scenarios.

Cautions

- **Diminished sales and marketing execution:** Customers are less frequently comparing Oracle's data integration tools with others in the market, denoting improvement needed in sales and marketing execution. Gartner client interactions also reveal a decline in Oracle's competitive evaluations for data integration tools.

- **Limited selection outside Oracle ecosystem:** In our client interactions, Oracle data integration solutions are more often considered when integrating data from and into Oracle Database, despite the high number of heterogeneous connectors and the position of OCI as a more cloud-agnostic option.

- **High-cost perception continues:** Perceived high cost continues to be a frequent reason for clients not shortlisting Oracle in their evaluations, despite them being aware of the technical capabilities offered. To mitigate this trend, Oracle is offering free tiers for its products and by adopting a more scalable and agile consumption-based pricing model.

Palantir

Palantir is a Visionary in this Magic Quadrant; the same as last year. It is headquartered in Colorado, U.S. It offers Palantir Foundry Data Integration Suite and AI Platform (AIP) as its data integration tools. Palantir's clients are mainly in North America and Europe, in various sectors including government, healthcare and aerospace. Organizations predominantly use Palantir for their operational data integration to support real-time decision making.

Palantir is investing in AIP boot camp programs to help prospective clients evaluate their pilot use cases such as supply chain optimization and preventive asset maintenance.

Strengths

- **Operational use-case focus:** Palantir provides operational data integration by modeling the enterprise's data landscape, enabling the business to make real-time decisions. It delivers this use case through its embedded consulting team of forward-deployed engineers (FDEs). Palantir is one of four recipients of the 2024 Customer Choice award for data integration tools, based on Gartner Peer Insights data.
- **Offering strategy:** Palantir's converged data management platform enables developers to build ontology-driven data services and real-time decisioning applications. It provides capabilities such as GenAI-enabled integrations in Foundry, enhanced security, lineage and ontology APIs, and LLM-driven augmentation to further simplify business operations on the ontology graph.
- **Growing customer base beyond the defense sector:** Palantir reported 83% year-over-year growth in the commercial sector by the end of 2Q24, taking its commercial customer count to about 300. Gartner client feedback indicates that Palantir's AIP boot camp program is not only well-received, but also attracts high demand among customers.

Cautions

- **Overreliance on consulting:** Palantir's business model relies heavily on its FDEs' playbooks via industry/domain application recipes and ontologies. Further, Palantir has limited consulting partners and global sales channels.
- **Complex pricing model and high costs:** Palantir charges its customers "per use case, per year," but provides limited detail into what encompasses the success criteria for a use case. There are many variables that affect the customers' TCO, including storage, compute, ontology and professional services consulting costs. Customers report that their engagement with Palantir becomes complex and challenging to scale beyond their initial use cases, mainly because of Palantir's pricing model and high costs.
- **Closed ecosystem:** Palantir's technology partnerships are very limited, and some of its customers report lock-in concerns associated with moving their data into Foundry. Customers express that Palantir does not satisfactorily serve their request to use general-purpose analytical tools in addition to Foundry.

Precisely

Precisely is a Niche Player in this Magic Quadrant; last year it was a Challenger. It is headquartered in Massachusetts, U.S. Its data integration products include: the Precisely Data Integrity Suite, which is used for various data management tasks (although this Magic Quadrant focuses on the data integration features); Precisely Connect, which is used for traditional batch ETL and CDC database replication, typically from on-premises databases; and Ironstream, which is used for data extraction from IBM Z mainframe and IBM i systems. Precisely's operations are diversified, with a presence in North America, Europe and APAC.

Precisely is continuing to develop the end-to-end, comprehensive capabilities of the Precisely Data Integrity Suite.

Strengths

- **Support for legacy/on-premises architecture:** Precisely supports on-premises products, offering both batch and real-time replication for sources such as IBM Db2 for z/OS, with robust performance.
- **Customer centricity and satisfaction:** Customers report general satisfaction with Precisely products, services and user experience. Precisely Strategic Services help customers plan their product implementation, and the wizard-driven pipeline configuration process allows business users to integrate data without significant technical knowledge.
- **Suitability for operational data integration:** Precisely offers real-time replication of large volumes of data. Customers performing operational data integration tasks that require, for example, data replication should evaluate Precisely.

Cautions

- **Customer perception that the product is limited to mainframe replication:** Users report that Precisely is mainly evaluated for expertise with mainframe sources and not other types of data integration scenarios, which reflects on its market positioning.
- **Limited consumption-based pricing:** Precisely offers a consumption-based pricing model for Precisely Connect, if purchased through the AWS marketplace, but not for the Data Integrity Suite. This current lack of consumption-based pricing reduces flexibility for clients.
- **Limited customer adoption:** The Precisely Data Integrity Suite currently has a low number of data integration customers. While the suite has gained traction with customers using

data quality and data governance capabilities, adoption for data integration remains limited.

Qlik

Qlik is a Leader in this Magic Quadrant; the same as last year. It is headquartered in Pennsylvania, U.S. Its data integration products include Qlik Talend Cloud, Talend Data Catalog, Qlik Data Integration, Stitch and Qlik Talend Data Fabric, with most product development efforts focused on its newest product, Qlik Talend Cloud. Qlik's operations are geographically diversified, with a substantial presence in Europe, and its clients are primarily large enterprises across various industries.

Since acquiring Talend in 2023, Qlik has spent the past year consolidating and updating its product portfolio, including introducing new capabilities such as data products. Its product roadmap focuses on continued improvements in data products, AI and data quality.

Strengths

- **Scope of product portfolio:** Qlik's product portfolio supports a vast range of data sources and targets. It offers robust data preparation and governance capabilities, and also provides ample support for bulk/batch data movement and data replication.
- **Support for data replication/synchronization:** Customers report favorably on Qlik's data replication capabilities. Their products support customers with tasks such as real-time reporting and analytics, real-time data integration, legacy modernization and cloud data movement.
- **Strong data governance:** With the Talend acquisition, Qlik provides strong governance capabilities, and supports customers with tasks such as metadata-driven data quality, lineage, policy enforcement, masking and annotation. It also offers a data product catalog and data marketplace, enabling a data-driven culture in the organization.

Cautions

- **Pace of R&D:** Qlik has not demonstrated strong R&D efforts for AI-related features over the past year. Gartner's analysis of the vendor's product roadmap leads us to believe this is a temporary issue due to Qlik's consolidation of its product portfolio following the acquisition of Talend.

- **Price increases and lack of communication:** Qlik is transitioning to a consumption-based pricing model, and the frequency and severity of complaints about pricing (as expressed during Gartner inquiries) has been high. Qlik's customers, particularly customers of the prior Talend product, report price increases in the past year and a lack of communication around these. Additionally, Talend Open Studio (a free, open-source product) has been discontinued.

- **Limited data virtualization:** Qlik does not offer mature data virtualization capabilities. Customers seeking data virtualization tools should thoughtfully evaluate Qlik's capabilities in this space.

Safe Software

Safe Software is a Niche Player in this Magic Quadrant; the same as last year. It is headquartered in British Columbia, Canada. It offers the FME enterprise integration platform, which is composed of FME Form, FME Flow and FME Flow Hosted. Safe Software's operations are primarily focused in EMEA and North America, and it supports clients mostly in the government, energy and utilities, and architecture, engineering and construction (AEC) sectors.

Safe Software is investing in improving its data integration styles, in particular data virtualization. It is also going to introduce new security features and augmentation capabilities such as integrating AI into new parts of its offerings.

Strengths

- **Core specialization in geospatial data:** Safe Software provides robust connectors for popular GIS applications and IoT sources to its customers. It supports geospatial semantic frameworks such as ArcGIS Knowledge. The company also offers data governance capabilities to prepare and enforce specific standards, such as the EU's INSPIRE Directive and the Indoor Mapping Data Format (IMDF).

- **Quick time to value and postsales customer experience:** Customers cite Safe Software as a time-to-value alternative to more comprehensive solutions. Users report high satisfaction with FME's ease of setup and use, and particularly appreciate the postsales experience both from a technical and business point of view. Safe Software is one of four recipients of the 2024 Customer Choice award for data integration tools, based on Gartner Peer Insights data.

- **Continuous product consolidation and improvement:** Safe Software keeps up with the market trends by introducing new capabilities in its product. For example, its Remote Engine Services allow the separation of the control and processing planes, favoring multi- and hybrid-cloud architectures. Safe Software provides support for emerging systems, and has introduced AI capabilities to expand the user base to ensure coverage for assisting less-technical users.

Cautions

- **Lagging data integration styles:** With FME, Safe Software lags behind its competitors in various data integration styles. While it covers all the required styles for this market, FME provides limited data virtualization and replication capabilities, with only basic CDC features.
- **A vision less focused on supporting features:** FME is a less comprehensive solution than others in this market, with Safe Software continuing to focus on developing core data integration capabilities. FME meets basic expectations for metadata management, data governance and DataOps capabilities, but needs to be complemented with third-party solutions in more complex, enterprise scenarios.
- **Limited brand awareness:** Safe Software does not execute marketing at the same level as its support for product and customer experience. Therefore, customers still evaluate its offerings primarily for niche scenarios, and rarely compare it with other vendors in this Magic Quadrant.

SAP

SAP is a Leader in this Magic Quadrant; the same as last year. It is headquartered in Walldorf, Germany. It offers SAP Datasphere as its main cloud data platform, which includes data integration services, and other data integration tools — SAP Data Services, SAP Data Intelligence Cloud and SAP Landscape Transformation Replication Server (SLT). SAP's operations are geographically diverse, with a majority of its customers in the EMEA region.

SAP is investing in creating an integrated cloud offering that supports the operational and analytical needs of different users through a semantically enriched platform.

Strengths

- **Prebuilt models and business logic:** SAP's offering includes several out-of-the-box transformers that are built by leveraging its own business expertise and ecosystem

partner knowledge to integrate SAP data and third-party sources. This allows SAP clients to overcome the challenges of replicating business logic from the source systems.

- **Data ecosystem vision:** Through SAP Datasphere, SAP intends to provide its customers with a single D&A ecosystem composed of well-integrated D&A tools. This combines the functionalities of SAP Data Intelligence Cloud and SAP Data Warehouse Cloud to provide a unified end-to-end solution, covering both business and IT needs.

- **Extensive metadata management and governance capabilities:** SAP offers a single entry point for data, and metadata discovery and management, including features such as data lineage, impact analysis, data quality checks and data access logging for auditability. Its semantically enriched, multidimensional data catalog enables metadata activation and supports recommendation engines and the creation of knowledge graphs.

Cautions

- **SAP-centric solutions:** SAP's data integration tools are not typically considered in the competitive evaluations of non-SAP clients due to SAP's platform-centric focus. Although SAP is investing in third-party connectors, SAP customers report the need to supplement SAP with other data integration tools to cover their integration requirements of non-SAP sources.

- **Uncertainty in product roadmap:** Gartner clients have expressed confusion about SAP's data integration tools portfolio, highlighting overlapping capabilities in product offerings and a lack of clear company direction. Customers are also uncertain about the life cycles of SAP Data Services and SAP Data Intelligence Cloud, including product support periods and any plans for retirement.

- **Lack of augmentation capabilities:** SAP's augmentation capabilities are lagging behind those of other Leaders in this Magic Quadrant. Although SAP offers predictive analytics capabilities and anomaly detection features, and has several AI features in its 2025 roadmap, augmentation and GenAI features are still embryonic and not adequately distributed in its product portfolio.

SnapLogic

SnapLogic is a Visionary in this Magic Quadrant; the same as last year. It is headquartered in California, U.S. It offers The SnapLogic Platform for Generative Integration as its data

integration offering. SnapLogic's operations are primarily focused in North America and EMEA. It has clients from diverse sectors, including technology, retail and healthcare.

SnapLogic is investing in its vision of generative integration, which proposes a single platform for generating all types of integrations using GenAI.

Strengths

- **Clear vision for GenAI:** SnapLogic's GenAI integration solution, SnapGPT, offers users the capability to generate, analyze and describe data pipelines in natural language. SnapLogic AgentCreator helps users to build GenAI applications, assistants and agents by collaborating with leading LLMs and also provides prompt generation. SnapLogic's native JSON-streaming architecture allows users to process large numbers of documents as part of RAG implementations.

- **Aligned with the needs of business users:** SnapLogic provides its customers with no-code data ingestion capabilities through AutoSync, and no-code data wrangling within the platform through AutoPrep. AutoSuggest offers the functionality to leverage AI to provide the next best Snap (prebuilt connector) when designing a data pipeline. SnapLogic has revamped its product packaging, and now offers separate editions for the business and enterprise.

- **Enhanced multicloud support for data and application integration:** SnapLogic provides an integration platform as a service (iPaaS), where the data plane can reside on-premises or in any of the major public cloud providers, enabling customers to centralize integration on a unified platform. It also combines data and application/API integration patterns in the same platform via Snaps. On Gartner Peer Insights, customers report high satisfaction with SnapLogic's cloud-native architecture.

Cautions

- **Challenges with job monitoring and debugging:** Some SnapLogic users report the need for clearer error messages upon job failures, and an improved UX for handling exceptions and restarting failed jobs. A job failure becomes challenging while processing complex data transformations, which sometimes requires rechecking too many Snaps to debug the root cause.

- **Limited data governance and metadata support:** SnapLogic's platform lags behind many leading vendor offerings in this Magic Quadrant in terms of data governance and metadata management tool integration, either from third-party vendors or its own.

Customers today need to expend effort to integrate SnapLogic with governance and metadata products, as SnapLogic's product integration with vendors like Collibra and Monte Carlo is still on the roadmap. The SnapLogic Integration Catalog is limited to capturing metadata generated within the platform, such as Snap statistics and logs.

- **Low market visibility:** Gartner client inquiries reveal that SnapLogic garners lower market interest when compared with its competitors in this Magic Quadrant. Fewer customers with data integration requirements evaluate SnapLogic than those with broad integration and API management requirements.

TIBCO

TIBCO is a Niche Player in this Magic Quadrant; last year it was a Challenger. It is based in Palo Alto, California. It offers the TIBCO Platform (which includes the Integration, Event Processing, Messaging Process Management, and Data Grid modules) and TIBCO Data Virtualization as its data integration tools. The TIBCO Platform focuses on offering customers a single control plane for data management and connecting their data integration needs with others, such as application integration and software development.

TIBCO did not respond to requests for supplemental information. Gartner's analysis is therefore based on other credible sources.

Strengths

- **Data virtualization and messaging capabilities:** TIBCO provides strong data virtualization capabilities and its TIBCO Data Virtualization product has extensive client deployments. Also, TIBCO Platform-Event Processing offers comprehensive event-processing capabilities.

- **Vision for unified data management:** TIBCO has a shared metadata layer across all of its data management products, and offers a unified view within the TIBCO Platform. Clients can better observe and manage data integration pipelines from a central interface.

- **Diversified data delivery styles:** TIBCO supports data and application integration patterns through a combination of its own data virtualization and event-based tools. The TIBCO BusinessWorks tool enables less-technical users to create data workflows.

Cautions

- **Market execution:** Gartner client inquiries indicate a decline in market awareness and interest in TIBCO. Gartner clients are not familiar with the TIBCO Platform, and only know

of TIBCO stand-alone tools such as TIBCO Data Virtualization, TIBCO EBX Master Data Management Software or TIBCO Messaging.

- **Limited vision for augmented data integration:** TIBCO lags behind its competitors in offering a holistic vision for leveraging AI (including GenAI) capabilities in data integration pipelines.
- **Challenges in cost forecasting:** TIBCO clients report challenges with migrating to new subscription licensing. TIBCO is not providing guidance on the impact of pricing changes on pipeline costs, or on product improvements related to the new pricing model.

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

- CData

Dropped

- SoftwareAG (StreamSets) was dropped due to StreamSets' acquisition by IBM in July 2024.
- Talend was dropped due to its acquisition by Qlik in May 2023.

Inclusion and Exclusion Criteria

Inclusion Criteria

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

To qualify for inclusion, the vendor's data integration tool (or tools) must be a "stand-alone" product directly usable by the buyer. The vendor must offer a generally available software product that meets Gartner's definition of a data integration tool. The tool must not be part of other solutions, such as analytics and business intelligence (ABI), database management system (DBMS) or other packaged SaaS applications, or be a general-purpose development platform or programming interface that requires extensive customization for data integration.

To use the product, customers should be able to procure only the data integration tool as an independent offering and not as a part of some other offerings — such as another form of tool suite, an application or other technology solution — of which the data integration capabilities are an "embedded" subset.

Vendor should adhere to the Market Definition/Description.

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

- **Revenue or customer count:**
  - Either generate at least **$40 million in software revenue from data integration tools** in calendar year 2023 — that is, from perpetual license with maintenance, or subscription with support (which would include payment only for data integration software), or through a consumption-based licensing model where the consumption metrics are being used only for the data integration software (on an annual basis).
  - Or maintain at least **500 subscription or maintenance paying customers** (where "customers" does not mean individual "user" license seats) for its data integration tools in **production**. (The number of downloads without license or maintenance revenue is informative, but not a qualifying piece of information.
- **Geography:** Have market presence in **at least three of the following regions** (regional market presence is defined as a minimum of 5% of the revenue of the verified production customer base, as well as the existence of dedicated sales offices or distribution partnerships in a specific region):
  - North America (Canada, Mexico and the U.S.)
  - Central and South America

  - Europe (including Western Europe and Eastern Europe)
  - Middle East and Africa (including North Africa)
  - Asia/Pacific region (including Japan)
- **Market presence:** 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.
- Have a data integration tool service **generally available** as of midnight, U.S. Eastern Daylight Time, on 30 June 2024. This includes any new functionality added to the service by the specified date. We do not consider beta, "early access," "technology preview" or other not generally available functionality or services. Additionally:
  - Any acquired product or service must have been acquired and offered by the acquiring vendor as of 30 June 2024. Acquisitions after this date were considered under their preacquisition identities, if appropriate, and are represented separately until the publication of the following year's Magic Quadrant.

Exclusion Criteria

Vendors that focus on narrow use cases that are too specific for the broader data integration market were 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 provided data integration as a capability embedded within their broader data management/analytics/data science platform, but did not provide a stand-alone/independent or commercially off-the shelf generally available data integration tool product.
- The vendor only provided support for open-source platforms/frameworks or development platforms, which need to be heavily engineered/customized for specific data integration tasks/use cases and/or are specific to a single data integration/data delivery style (such as stream data integration only).

- The vendor provided adapters or drivers to various D&A sources and targets, thereby indirectly supporting data integration, but did not market a stand-alone data integration tool.
- The vendor only provided self-service data preparation tools for power users, analysts and line-of-business users. These tools, however, do not have the ability to physically move data or operationalize self-service data flows and models into production through data movement, governance and sharing, if and when needed.

Honorable Mentions

- **Airbyte** is headquartered in California, U.S., and offers Airbyte Open Source, Airbyte Cloud, Airbyte Teams and Airbyte Enterprise. Customers can benefit from Airbyte's 400+ low/no-code connectors available in the Airbyte Marketplace and create more than 10,000 custom ones using Connector Builder, Connector Development Kit and AI Assistant. These connectors, embeddable within other systems, enhance productivity for data-intensive, multiplatform applications. The 2024 Airbyte 1.0 launch has strengthened DataOps support (Terraform provider) and data governance capabilities, including data masking, encryption and access controls. Airbyte supports GenAI use cases with native vector DB compatibility and integrations with RAG frameworks to accelerate AI application development. Airbyte did not meet our inclusion criteria for revenue and market presence, and was therefore not included in this year's Magic Quadrant.

- **Dataddo** is headquartered in Czech Republic, and offers three products: Data to Dashboards (to move data into the caching layer of analytics tools); Data Anywhere (to move data into cloud data warehouses); and Headless Data Integration (to move data into customers' data products via an API). Customers use Dataddo for its self-service capabilities that empower nontechnical and technical users to bidirectionally move data across systems, with basic data governance/privacy/compliance guardrails built in. Dataddo did not meet our inclusion criteria for revenue and market presence, and was therefore not included in this year's Magic Quadrant.

- **dbt Labs** is headquartered in Pennsylvania, U.S., and offers dbt Cloud to help organizations adopt and operate dbt to enable data transformation at scale, with a focus on the transformation part of data integration. dbt is an open-source tool that allows data teams to collaboratively build and deploy templated SQL-oriented data transformation workflows so that teams can build and operationalize data pipelines using software development best practices through directed acyclic graphs (DAGs). dbt Cloud is a

hosted service that simplifies deployment processes and operates dbt at enterprise level. It includes an online GUI, orchestration, observability, documentation and discovery capabilities, model governance, and the possibility to build a semantic layer across multiple dbt instances. Although very popular for data transformation, dbt Labs did not meet the inclusion criteria related to support for mandatory features, such as supporting extraction, ingestion and delivery of data using multiple common data integration styles. It was therefore not included in this year's Magic Quadrant.

- **eQ Technologic** is headquartered in California, U.S., and offers the eQube Data as a Service (eQube-DaaS) platform. This includes six different tools to enable end-to-end analytics use cases, and supports various data and application integration requirements. eQ's comprehensive suite of tools, including eQube-MI (Migration and Integration) and eQube-TM (Transformation Modeler), enable metadata-driven, low/no-code integration and transformation across multiple data and application sources and targets. These tools use a data virtualization layer with a security-first approach. The tool suite supports all common integration patterns and requirements, such as data and application integration, API management and product life cycle management. This makes it popular with clients in asset-intensive industries. Although eQube-DaaS offers a complete range of data integration capabilities for this market, eQ Technologic did not meet the inclusion criteria related to market presence for this year's Magic Quadrant.

- **Nexla** is headquartered in California, U.S., and offers a unified platform for data integration, data products and GenAI RAG pipelines, with enterprise-grade collaboration, monitoring and governance. It supports multiple data integration styles, including ETL, ELT, CDC, streaming and API integration with no-code, low-code and code-centric interfaces. The flexibility of Nexla's platform comes from "Nexsets" — virtual schema-centric data products that create an abstraction over data sources and destinations. Nexsets encapsulate data, metadata, transformations, validation, access controls, schema, and automated versioning without creating a new copy of data, simplifying catalog integration and discovery. Despite its strong offering, Nexla was not included in this year's Magic Quadrant because it did not meet our market presence criteria.

- **Prophecy** is headquartered in California, U.S., and offers Data Transformation Copilot, a tool that helps users build, deploy and observe data pipelines. It supports a number of different use cases, such as making all data users more productive on Databricks, modernizing ETL and data preparation processes (including migrating from legacy tools), and transforming data to support building AI applications, such as processing

unstructured data for RAG. Prophecy features a visual drag-and-drop UI, a coding integrated development environment (IDE), and AI-assisted suggestions for data transformation operations. Once pipelines are built visually, they are turned into standard, cloud-native code (PySpark, Scala or SQL). Despite its strong features, Prophecy was not included in this year's Magic Quadrant because it did not meet the criteria for market presence, revenue and customer count.

- **The Modern Data Company** is headquartered in California, U.S., and offers the DataOS (data operating system) platform, with LLM capabilities to manage the data product life cycle. It aims to accelerate value delivery through data product creation and delivery capabilities. From a technical point of view, the platform closes the gap between conceptual and physical data modeling. Users can define conceptually what they want, and the platform's software traverses and integrates data. As a result, organizations can avoid long project cycles on technical implementations and benefit from faster iterations of conceptual models. The Modern Data Company did not meet our inclusion criteria for revenue and market presence, and was therefore not included in this year's Magic Quadrant.

- **Rivery** is headquartered in New York, U.S., and offers Rivery Platform, a SaaS ETL/ELT data integration platform with data ingestion, data transformation and pipeline orchestration capabilities. Rivery offers 200+ built-in connectors, as well as the ability to develop custom ones via low code. Transformation can be implemented through SQL, Python/Pandas code or dbt. The company allows its customers to orchestrate workflows with automation and custom logic, as well as using other orchestrators like Apache Airflow to execute Rivery's data pipelines. Users particularly appreciate Rivery's quick time to value and its ease of use, by both technical and less-technical personas. They also praise the quality of its sales and support services. Rivery did not meet the inclusion criteria for revenue and market presence, and was therefore not included in this year's Magic Quadrant.

- **Stratio Big Data** is based in Spain. It offers Stratio Generative AI Data Fabric, which aims to simplify data understanding and accessibility for business users by utilizing GenAI to automatically create a "Business Data Layer" (BDL). Through it, Stratio drives governed data consumption and improves business user productivity. The BDL exposes industry standard ontologies in a knowledge graph representation with several data integration, automatic cataloging and governance capabilities. As new technical assets and new relationships are discovered, the corresponding graph, metadata and semantic schemas

are automatically inferred and established within the BDL. A chat interface allows users to do data transformation tasks and ask business questions in natural language, receiving answers based on the BDL. Stratio did not meet our inclusion criteria for revenue and market presence, and was therefore not included in this year's Magic Quadrant.

- **Workato** is headquartered in California, U.S., and offers the Workato Data Orchestration cloud platform, which enables a variety of data integration use cases and common integration patterns, including "recipes" allowing for integration flow reuse. Workato Data Orchestration was released as a stand-alone product in late 2023 and is built on Workato's iPaaS platform, which is included in Gartner's Magic Quadrant for Integration Platform as a Service. Workato Data Orchestration enables stand-alone data integration capabilities while leveraging the technical foundations of Workato's iPaaS, including connectors, metadata management and data pipeline observability. Workato did not meet our inclusion criteria for market presence, and was therefore not included in this year's Magic Quadrant.

Evaluation Criteria

Ability to Execute

Gartner analysts evaluate providers on the quality and efficacy of the processes, systems, methods or procedures that enable IT provider performance to be competitive, efficient and effective, and to positively impact revenue, retention and reputation within Gartner's view of the market.

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 etc. This can be offered natively or through OEM agreements/partnerships as defined in the market definition and detailed in the subcriteria.

We evaluate each vendor in the following areas:

- The capabilities it has to address current market requirements. These include, but are not limited to, bulk/batch data movement, CDC-based data replication and synchronization, data services orchestration, stream data integration for real-time use cases, data

migration support, support for data engineering for analytics and data science, and other integration efforts for operational use cases (like master data management [MDM]).

- The degree of openness of the vendor's technology and product strategy — that is, the ability to exchange metadata with third-party offerings.
- The ability of its product to allow interoperability to open-source solutions and third-party offerings, which is appreciated by end users. Some consumers are prepared to accept products from many different suppliers and assemble them together on their own.
- Connectivity options as both sources and targets to not only nonrelational databases, cloud applications and cloud data stores (such as cloud object stores and cloud data warehouses), but also traditional stores (including relational databases and enterprise applications).
- Connecting data integration activities to data quality and governance — these are integral in supporting operational data integration use cases that require sharing high-quality data along with its lineage, such as MDM and B2B data sharing.
- The ability to offer both serverless metered pricing options (for net new use cases) and traditional pricing models such as node-/core-based models (when the use cases do not require flexibility of compute).
- The ability to support DataOps and orchestration capabilities to enable agile pipeline delivery, management, operationalization and maintenance. CI/CD integration, support for Git and Jenkins, support for regression test/validation, support for data observability, and support for schema drift handling are all expected.

Overall Viability (Business Unit, Financial, Strategy, Organization): Financials

Viability includes an assessment of the organization's overall financial health, and the financial and practical success of the business unit. It considers the likelihood of the organization continuing to offer and invest in the product, as well as the product position in the current portfolio.

We evaluate each vendor in the following areas:

- The appropriateness of its financial resources, the continuity of its people and its technological consistency, and how that affects the practical success of the business unit or organization in generating business results.

- The growth of its product lines, its annual recurring revenue (ARR), profitability and growth in new geographies/use cases.
- Revenue growth in products/services to determine vendor growth in the data integration software market.
- Growth in high-momentum use cases such as cloud data integration.
- Revenue growth through cloud integration tools (iPaaS, SaaS, etc.)
- Other metrics to determine financial viability and spend on R&D efforts to continue differentiation and growth in product lines.
- Investment in skills, personnel and roles for product development, delivery and support. Retention and growth metrics — both are necessary.
- Industry-specific certifications — FedRamp, HIPAA, GDPR, etc.

Sales Execution/Pricing

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

We evaluate each vendor in the following areas:

- The ability to offer modular solutions. Organizations increasingly seek "modularity" or the capability to isolate specific required functions in data integration, which are then reflected in their implementation approach and cost allocation.
- The ability to provide tools and capabilities through different pricing models appropriate by use cases, persona and environment.
- The ability to support buyers that are looking for new pricing metrics that abstracts them from the underlying metrics of cloud pricing. Buyers are looking for vendor options that support serverless metered pricing metrics that are a true reflection of the actual work done and that can separate compute from storage/infrastructure.
    - Having said that, organizations are also wary that serverless metered pricing options that don't consider good financial governance can soon get out of control. Thus, we

also closely evaluate vendors on their ability to enforce financial governance metrics into their pricing models and licensing metrics.

- The ease with which customers can hold them accountable for agreed on SLAs. Buyers are evaluating ways through which they can hold vendors accountable for the promised SLAs (in terms of uptime, turnaround times to issues, bug fixes, migrations, and so on). Providers must demonstrate ways through which customers can escalate and attain credits/discounts when SLAs are not met.
- Finally, data management as a discipline needs to track, predict and preempt the overall costs associated with cloud integration workloads, as D&A teams become more distributed across various domains and are increasingly placed in various lines of business. This makes it important that data management leaders are able to associate the cost of running data integration workloads with their value. They also need to have control over the allocation of processing capacity to workloads they deem to be important through optimal analysis of performance and system metadata. Vendors will therefore need to offer tools that can enable financial governance and automate aspects of this through FinOps approaches, which should also be reflected in their pricing and governance models.

Market Responsiveness and Track Record

Ability to respond, change direction, be flexible and achieve competitive success as opportunities develop, competitors act, customer needs evolve, and market dynamics change. This criterion also considers the provider's history of responsiveness to changing market demands.

We evaluate each vendor in the following areas:

- We look evidence of how the vendor has "course corrected" to changing buyer requirements in terms of use cases, upcoming capabilities, pricing models and support requirements.
    - As an example, managed services options for maintaining data pipelines and handling schema drifts are in demand, particularly by business teams and citizen integrators. Providers that can enable these requests are therefore selected over others that are still focused on IT teams alone.
- The requirements to enable data fabric designs are also increasing. We therefore look for vendors that are adding features (including those on the roadmap) and formulating

partnerships (such as the ones with graph providers and semantics technologies) to enable more comprehensive data fabrics.

- Even though solutions that provide low/no-code UIs are preferred, Gartner also hears requests from data engineering teams for tools that can open "portals" for custom coding and importing scripts created in languages such as R and Python for highly advanced transformations.

- The market is seeking vendors with expertise in transitioning away from environments that seem to have lost traction — for example, Hadoop — and moving toward other popular data/sources and targets, such as cloud database platforms as a service and cloud applications.

- Preference is given to vendors that have the ability to respond to current market demands for solutions that support AI/ML to automate complex and repetitive data integration tasks — for example, data transformation, orchestration, parts of data modeling and data delivery.
- Preference is given to those vendors that provide open ecosystems to support independent data integration (which does not depend on any cloud infrastructure specifically, or a DBMS, or a proprietary data exchange format, or data storage format).

Marketing Execution

The clarity, quality, creativity and efficacy of programs designed to deliver the organization's message in order to influence the market, promote the brand, increase awareness of products and establish a positive identification in the minds of customers. This "mind share" can be driven by a combination of publicity, promotional activity, thought leadership, social media, referrals and sales activities.

We evaluate each vendor in the following areas:

- Brand recall value has a high premium in a mature market like data integration.

- The ability to develop a means of converting community "chatter" and excitement to support delivery and go-to-market campaigns.

- The overall effectiveness of the vendor's marketing efforts, which impact its mind share, market share and account penetration.

- 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.
- 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 ability to support and become part of open community channels for sharing best practices, connectors/code/mappings/other assets, or supporting open metadata sharing standards.
- Finally, how the vendor is rated and perceived on community review and evaluation channels like Gartner Peer Insights.

Customer Experience

Products and services and/or programs that enable customers to achieve anticipated results with the products evaluated. Specifically, this includes quality supplier/buyer interactions, technical support, or account support. This may also include ancillary tools, customer support programs, availability of user groups, service-level agreements, etc.

We evaluate each vendor in the following areas:

- We assess the level of satisfaction expressed by customers with the vendor's product support and professional services support.
- We look at customers' overall relationship with the vendor, their experience while upgrading software versions, the learning curve for new users given the training resources available, and customer perceptions of the value of the vendor's data integration tools relative to cost and expectations.
- We evaluate the ability of the vendor to maintain parity between cloud and on-premises software. We also measure after-sales support, migration ease, ease of deployment and overall maintenance and support.
- We evaluate customer feedback on a vendor's ability to meet roadmap deliverables, technical knowledge sharing, skills enablement, augmentation and training.
- We look at various platforms for data points that include, but are not limited to, our interactions with end users in inquiries, Peer Insights data, surveys, customer reference

calls, touchpoints across various Gartner and external events, community chatter and vendor briefing data.

- 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 success in “innovative” market delivery in reviewing the customer experience.

Operations

The ability of the organization to meet goals and commitments. Factors include quality of the organizational structure, skills, experiences, programs, systems and other vehicles that enable the organization to operate effectively and efficiently.

We evaluate each vendor in the following areas:

- Operations are not specifically differentiating to end-user markets, but product management consistency and support/maintenance practices add to the overall customer experience and to 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 active metadata, metadata management overall and semantic tiers.
- Suppliers must demonstrate that they are well-positioned to provide ongoing support for the massive bulk/batch data movement market, and support for upcoming data delivery styles including replication, streaming, API, virtualization and messaging.
- Partner programs, skills augmentation, improvements in support and services, training materials and programs, and delivery with external service providers are all important in this evaluation criterion.
- We look for evidence of:
  - Knowledge transfer from vendor to end-user organization on an ongoing basis.
  - Training, skills augmentation of internal staff and overall enablement of data teams.
  - Investment in support, training and community building.

Ability to Execute Evaluation Criteria

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

Source: Gartner (December 2024)

Completeness of Vision

Market Understanding

Ability to understand customer needs and translate them into products and services. Vendors that show a clear vision of their market — listen, understand customer demands, and can shape or enhance market changes with their added vision.

We evaluate each vendor in the following areas:

- The ability to formulate product vision around complex multicloud and hybrid data integration capabilities. Also, the ability to participate in cloud data ecosystems and support data management architecture patterns (e.g., data mesh, data fabric, lakehouse) to enable outputs like data products.
- Provide core data integration services support, including for all types of data movement topologies, bulk/batch, replication/synchronization, streaming and virtualization, and support for the interoperating and combining of these data delivery styles.

- The ability to support multiple personas of technical (e.g., data engineers) and less-technical (e.g., business analysts, data scientists) users through use-case-aligned UIs, GenAI and self-service data preparation.
- The ability to enable augmented data integration, including active metadata analysis and DataOps practices, to automate recommendations and actions, and support agile data engineering practices.
- The ability to support optimization of data integration costs, financial governance and cloud financial management, through FinOps capabilities.
- The ability to provide “advisors” and insights into design, development, deployment and management of data integration services to support decision insights and decision automation.

Marketing Strategy

Clear, differentiated messaging consistently communicated internally, and externalized through social media, advertising, customer programs, and positioning statements.

We evaluate each vendor in the following areas:

- Evidence of good internal and external messaging to differentiate their offerings from other market categories — for example, basic ingestion tools versus more complex data integration use cases.
- Good visibility into the product portfolio, along with features across each separate tool, including possible overlaps, ways to buy, means to procure, support tiers and licensing.
- Clear messaging on trial/freemium to full enterprise offerings (with differentiators across each support/SLA level), as part of a product-led growth (PLG) strategy.
- The ability to showcase a complex portfolio through clear differentiated messaging, justifying purchase and clarifying use of each product/SKU.
- Efforts and investments (with demonstrations) in partner programs, training programs, OEM/value-added reseller (VAR), other partnerships, cloud provider partnerships, SI partnerships, and so on.
- Demonstrable proof of expansion in training, certifications and availability of talent in the market (through, for example, partner programs and training).

Sales Strategy

A sound strategy for selling that uses the appropriate networks including direct and indirect sales, marketing, service, and communication. Partners that extend the scope and depth of market reach, expertise, technologies, services and their customer base.

We rate the vendors on:

- Expansion in sales partner networks.
- Affiliate partnerships.
- Vendors' ability to become a part of different cloud data ecosystems as an independent data integration partner for the customer.
- Strategy to grow beyond existing markets, use cases, geographies and critical capabilities.
- Demonstrable evidence of improvement in communication of existing and upcoming tools/services.
- Growth through varying channels (e.g., OEMs, VARs, SIs, hyperscaler marketplaces, consulting companies, joint go-to-market, partnerships with vendors in the D&A space).

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.

We evaluate each vendor in the following areas:

- We look for tools that can deliver distributed data integration across on-premises, cloud, multicloud and edge ecosystems.
- Support for the creation and management of AI-ready data assets.
- Tools exhibiting improvement in automation-oriented capabilities, including GenAI capabilities.
- While advanced capabilities are needed, tools must not drop existing and "traditional" requirements of data integration, including bulk/batch capabilities, supporting hybrid/on-premises sources and targets or supporting developers.

- The ability to understand, convert, utilize and analyze “active” metadata in support of data fabric designs, enabling profiling, machine learning, evaluation and comparison of assets, self-correcting process optimization, and “push” recommendations for new data assets.
- Efficient delivery of integrated data in a decentralized operating model as data products or data mesh, which can be created, maintained and operationalized as software products through subject matter experts across distributed domains.
- We assess vendors’ roadmaps, existing capabilities and their degree of openness to support interoperability (through open metadata exchange) with other D&A tools (their own or a third party’s).
- We assess roadmaps to support DataOps and FinOps, agile data engineering, cloud financial management and cost-optimization measures.
- Finally, we look for a roadmap (through demonstrable evidence) that supports seamless on-premises to cloud migration of tools or version/system migration in general. This will include changes to tool pricing.

Business Model

The design, logic and execution of the organization’s business proposition to achieve continued success.

We evaluate each vendor in the following areas:

- The overall approach the vendor takes to future-proof its business model and marketing strategy — including diversity of delivery models, packaging and pricing options, partnerships and data ecosystem relationships.
- The ability to provide both “current” requirements, through best-fit engineering tools, and future requirements, through end-to-end platforms or best-of-breed options.
- The business proposition must include the ability for end-user organizations to “try before they buy.” Tools must be able to interoperate with existing tools within the customer base, rather than requiring the replacement of all current software.
- We look for vendors to create a niche for themselves in this complex market. How vendors carve out differentiation, land-expand, grow and target specific differentiated use cases, personas, delivery models and even operating models is evaluated.

- Most importantly, for a data integration tool to do well, it must not exist in isolation. How does it work with the overall cloud data ecosystem? Does it interoperate and share metadata with CSP ecosystems?

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.

We evaluate each vendor in the following areas:

- We look at the degree of emphasis the vendor places on vertical solutions, and the vendor's depth of vertical market expertise.
- We look for tools that are able to create a specific industry model, ontology or knowledge graph based on an industry-specific taxonomy.
- Vertical/domain-specific solution accelerators, KPIs, best practices and other industry starter templates might be favored by some buyers, in addition to industry/domain experts being a part of the professional services provided.
- Finally under this criterion, we look at a vendor's ability to provide vertical knowledge and expertise in presales, sales, implementation and support for targeted verticals.

Innovation

Direct, related, complementary, and synergistic layouts of resources, expertise or capital for investment, consolidation, defensive or preemptive purposes.

We evaluate each vendor in the following areas:

- Support requirements for the emerging management and integration of GenAI (RAG, etc.), and AI-ready data assets.
- Managing location-agnostic capability in data integration — the ability to not have to move or replicate data necessarily, but to connect to data in place when feasible and take the processing to the data (rather than vice versa) to execute integration.
- As data becomes highly distributed, data integration activities are also required to become easily distributable to any data location. Each vendor should be able to recommend/determine when data needs to be moved for optimal processing and deliver

workloads to the most optimal execution engines through containerized services and microservices architecture.

- Integration execution engines should deploy and run on-premises and in the multicloud, and switch between them.
- Converging data and application integration approaches are now expected.
- ML-based automation using internal analytics on all kinds of collected metadata to support integration activities is another area of improvement that the market currently demands.
- The growing diversity of users indicates a much higher demand for administrative, auditing, monitoring and even governance controls that utilize job audit statistics.
- Graph analysis and GenAI assistants to determine user classification and optimization "hints" are also increasingly demanded, supporting less-technical user personas.
- Provide financial governance (FinOps) to enable business units, CDOs and CFOs to support cost optimization as workloads move to the cloud.
- Support for enabling data products and delivering and managing them across distributed domains, ensuring their governance and allowing management of these products and their infrastructure.
- Finally, because the increase in the number of data pipelines is inevitable, organizations are expecting providers to have DataOps-oriented capabilities that can support CI/CD; project management capabilities such as Git and Jenkins; automated testing/validation; the ability to handle various environments in an agile manner; sandboxes on demand and management of them; and agile pipelines creation, reuse, execution and management.

Geographic Strategy

The provider's strategy is 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.

We evaluate each vendor in the following areas:

- The ability to provide customers with local support, via VARs, other resellers, and channel, vertical/industry and SI partnerships.

- The ability to provide continuity of support across regions.
- The ability to monitor, via their development platforms, where data originates with jurisdictional cognizance, and where it is eventually delivered.
- The ability to address possible violations of national laws due to data movement.
- The vendor's strategy for expanding into markets beyond its home region/country and its approach to achieving global presence (for example, direct local presence and use of resellers/distributors) are crucial for capitalizing on global demands for data integration capabilities and expertise.
- Finally, each vendor is evaluated on their level of support and performance in different geographies, and their ability to support after-sales maintenance, etc.

Completeness of Vision Evaluation Criteria

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

Source: Gartner (December 2024)

Quadrant Descriptions

Leaders

Leaders are front-runners in their capability to support combinations of different data delivery styles, such as combining and switching between ETL/ELT pipelines, data replication and synchronization, streaming data and data virtualization, based on use-case demands to support all data integration requirements. Leaders support more advanced data management patterns, like active metadata and data fabric, data mesh, lakehouse architectures and delivery of data products. These vendors also support AI and GenAI through both embedded capabilities and enabling integration patterns like LLMs leveraging RAG. Most of these vendors also support complex hybrid and multicloud deployments, and participate in data ecosystems.

Leaders exhibit significant market mind share and recognize the need for new and emerging market demands — often providing new functional capabilities in their products ahead of demand. They do this by identifying new types of business problems that data integration tools can significantly impact, thereby providing business value. Leaders have an established market presence, significant size and a multinational presence. Leaders effectively establish their data integration tools to support both traditional and new data integration patterns to capitalize on market demand. Many Leaders see continued loss of market share to CSPs or smaller vendors that offer more focused capabilities, attractive pricing or go-to-market strategies targeting business users.

Challengers

Challengers exhibit a strong understanding of the current data integration market demand and exhibit both the credibility and viability to deliver. Some Challengers are mature in specific capabilities and use cases, which enables them to deliver targeted use cases faster and with a better overall TCO than other vendors (sometimes even Leaders). Challengers have developed best practices for leveraging their strongest product capability in new delivery models.

Challengers generally have substantial customer bases. They exhibit strong market presence, although implementations may be of a single project nature or reflect multiple projects of a single type, or they reflect use cases to support integration within specific CSP environments. Challengers also participate in data ecosystems, filling gaps for data

integration workloads, and often provide more flexibility than native CSP data integration tools (being able to support multicloud and hybrid scenarios).

Challengers continue to capitalize on the cloud data ecosystem trend and realize that the major growth in the market is attributed to multicloud data integration and delivery. Challengers strive to deliver most data integration capabilities at competitive price points and often use innovative pricing and licensing, packaging and go-to-market strategies to exert pressure on Leaders and win large deals.

Visionaries

Visionaries demonstrate a strong understanding of emerging technology and business trends, or they focus on a specific market need that is far outside of common practices, while also possessing capabilities that are expected to grow in demand. The Visionaries in this Magic Quadrant have focused early on futuristic or emerging capabilities and go-to-market strategies to capitalize on their unique offerings.

Visionaries sometimes lack market mind share or credibility beyond their established customer base, their main use cases or very specific application domains/verticals. Visionaries may still be ramping up partnerships with SIs, consulting companies and other partners, and may lack the installed base and global presence of larger vendors. Visionaries may also lack partnerships and/or tight integrations with other incumbent data management vendors, such as in the areas of third-party metadata management, data governance or data quality solutions. Finally, Visionaries may be established players in adjacent data management markets that do not have the level of execution of the Leaders in this one.

Niche Players

As the data integration tools market is mature, Niche Players generally do not exhibit gaps in primary functionality or features. Instead, they are simply challenged to improve 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 may not appear frequently in competitive situations for comprehensive and/or enterprise-class data integration deployments. Many have strong offerings for a specific range of data integration problems. These include a set of specific data delivery styles, application domains or use-case scenarios (such as geospatial data integration or location

intelligence). Some Niche Players deliver substantial value for their customers in the associated segment.

Niche Players often exhibit advantages in pricing within their established footprint and in vertical or horizontal solutions, and sometimes cannot complement an organization's other data management technologies. Niche Players are known for solving one part of the data integration problem well through a targeted solution and may be a "good enough" solution for organizations with less complex needs. However, they may be lacking maturity in certain features that display market vision.

Context

Data and analytics leaders must navigate a market brimming with products that claim to solve a range of data integration problem types. However, not all vendor solutions have experience in, or evenly provide, all the relevant capabilities needed across Gartner's key data integration use cases (see the companion **Critical Capabilities for Data Integration Tools**). Data integration functions enable ingestion, sharing and consumption of data across all organizational and systemic boundaries. Therefore, organizations are increasingly seeking a comprehensive range of improved data integration capabilities to modernize their data, analytics and AI infrastructures.

The tools in the data integration market continue to support basic to advanced data integration styles in support of common use cases. These styles include a mix of traditional extraction, transformation and loading (ETL), transformations in place (ELT), data virtualization, and ingestion, analysis and delivery of high-velocity streaming data. In addition to enabling various data integration patterns, tools must now support hybrid, multicloud environments, participate in larger data ecosystems (with data governance, metadata, MDM tools, etc.) and leverage AI-driven capabilities to optimize pipeline and user experiences, and support client implementations of GenAI. Data integration tools must also support multiple user personas — such as data engineers and less-technical users like business analysts. These tools enable modern data management practices like data fabrics, data mesh, lakehouses and data products, all while being expected to enable better DataOps and FinOps practices.

While basic data integration functions remain a core component in data architectures, the evolution of data integration tool capabilities, multicloud/hybrid deployment models and

participation in data ecosystems mean vendor selection should be based on balanced support for use cases, as well as integration with current and target data management architectures.

Market Overview

The market for data integration tools continues to evolve and is supported by strong revenue growth and continued rapid adoption. The market grew at 9.8% in 2023, compared with 9.1% in 2022 (see **Market Share: Data and Analytics Software, Worldwide, 2023**). A conservative estimate for the five-year compound annual growth rate (CAGR) for the 2023 to 2028 time frame forecast is 7.2% (see **Forecast: Enterprise Infrastructure Software, Worldwide, 2022-2028, 3Q24 Update**). The reduction in future growth outlook is due to expected revenue gradually shifting from stand-alone tools (for data integration, data quality, metadata management, etc.) to converged platforms for data, analytics and AI. Data and analytics leaders are realizing that data integration is a foundational component of their data management strategies.

Continued loss of market share by leaders and market growth driven by modern data integration requirements and cloud data ecosystems: The top five vendors in this market (based on market share) had a collective market share of 71% in 2017. This number has been steadily dropping over the years, and in 2023, the collective market share was only 53%. One of the main reasons for this is that, as organizations migrate data integration to the cloud, they're evaluating modern cloud-native/SaaS vendors and public cloud hyperscalers (CSPs), including native cloud data ecosystems (see **Secure a Competitive Edge With Data Ecosystems: A Comprehensive Guide**). Another reason is that smaller vendors with more focused offerings, or those that target business users through innovative offerings, pricing, packaging and go-to-market strategies, continue to disrupt larger vendors. Vendors gaining market share have a common theme: They focus on leadership in specific data integration styles such as data virtualization, data replication or streaming; and/or they focus on data integration delivered as a native and managed cloud service. Vendors need to find the right balance between all-encompassing platform solutions and easily accessible point solutions to keep pace (see **Market Share Analysis: Data Management Software (Excluding DBMS), Worldwide, 2023**).

Tools must enable multiple personas and modern data management architectures, support hybrid and multicloud deployments, and leverage GenAI: Data engineers are not the only

users of data integration tools; less-technical users (for example, business analysts, line-of-business owners, power users and data scientists) also need low/no-code options to build data pipelines. Data management practices using data fabric, data mesh, lakehouse and data product designs, all require data integration capabilities, including more complex analysis of metadata to drive automation and optimization. Data integration tools must also support complex hybrid and multicloud deployment models. Finally, the use of GenAI assistant is now standard in most data integration tools, allowing creation, observability and management of data pipelines through NLQ type interfaces (see **Emerging Tech Impact Radar: Generative AI**).

Support for DataOps and FinOps practices are critical: DataOps practices focus on optimization of data pipelines through CI/CD and collaboration best practices. Most data integration tools support some level of observability and options to optimize pipelines. There is also a growing set of Data Observability and DataOps vendors that focus on supporting these practices, often in conjunction with data integration tools (see **Market Guide for Data Observability Tools** and **Market Guide for DataOps Tools**). FinOps or cloud financial management is also critical, in that most clients report significant challenges managing cloud costs. Data integration pipelines are notoriously difficult to measure in terms of overall cost, and often lead to higher-than-expected compute costs, especially when POCs scale into production. FinOps for data integration comprises the capabilities that enable D&A leaders to iteratively control spending, understand product performance and make choices regarding price-to-performance trade-offs (see **The State of FinOps for Data and Analytics, 2024**). This results in optimal allocation of resources in the cloud for efficient cost optimization. Data integration tooling is still relatively immature in FinOps management, but many vendors are improving their overall FinOps capabilities — usually tied to overall observability capabilities.

⊕ Evidence

⊕ Evaluation Criteria Definitions