# Magic Quadrant für Datenintegrationstools

Veröffentlicht 18 August 2020 - ID G00450251 - 104 min lesen

By Analysts Ehtisham Zaidi, Eric Thoo, Nick Heudecker, Sharat Menon, Robert Thanaraj

Der Markt für Datenintegrationstools sieht eine neue Dynamik, die durch dringende Anforderungen an Hybrid-/Multicloud-Datenmanagement, erweiterte Datenintegration und Data Fabric-Designs angetrieben wird. Diese Bewertung von 20 Anbietern wird Daten- und Analyseleitern dabei helfen, ihre Datenintegrationsanforderungen optimal zu erfüllen.

## Strategische Planungsannahmen

Bis 2025 werden über 80 % der Unternehmen mehr als einen Cloud Dienstanbieter (CSP) für ihre Daten- und Analyseanwendungsfälle verwenden, was es für sie entscheidend macht, eine unabhängige und CSP-neutrale Integrationstechnologie zu priorisieren, um Anbietersperren zu vermeiden.

Bis 2023 wird das erweiterte Datenmanagement die Abhängigkeit von IT-Spezialisten für sich wiederholende und wirkungsarme Datenmanagementaufgaben verringern und somit bis zu 20 % ihrer Zeit für Zusammenarbeit, Schulungen und höherwertige Datenmanagementaufgaben freimachen.

Bis 2023 werden Unternehmen, die Daten-Fabrics nutzen, um Datenmanagementprozesse dynamisch zu verbinden, zu optimieren und zu automatisieren, die Zeit für die integrierte Datenbereitstellung um 30 % verkürzen.

Bis 2022 wird die Anwendung von Graphenverarbeitungs- und Graphendatenbanken um 100 % wachsen, um die Datenintegration zu beschleunigen und eine adaptivere Datenwissenschaft zu ermöglichen.

## Marktdefinition/Beschreibung

Gartner definiert **die Datenintegration** als eine Disziplin, die die Praktiken, Architekturtechniken und Tools umfasst, die es Unternehmen ermöglichen, Daten über das gesamte Spektrum der Datentypen hinweg zu erfassen, zu transformieren, zu kombinieren und zu bereitstellen. Diese Integration erfolgt im Unternehmen und darüber hinaus – partnerübergreifend sowie über Datenquellen und Anwendungsfälle von Drittanbietern –, um die Datenverbrauchsanforderungen aller Anwendungen und Geschäftsprozesse zu erfüllen. Dies beinhaltet alle Technologien, die Datenintegrationsanforderungen unabhängig von der aktuellen Marktnomenklatur unterstützen (z. B. Datenerfassung, Datentransformation, Datenverarbeitung, Datenpipelining, Datenreplikation,

Datensynchronisierung, Datenvirtualisierung, Datenintegration, Datendienste, Datendatenfabricen, Datenengineering und vieles mehr).

Beispiele für beliebte Datenintegrationsnutzungsszenarien sind (aber nicht beschränkt auf):

- **Datenintegration für optimierte Analysen –** Zugriff auf, Warteschlange oder Extraktion von Daten aus betriebssystemellen Systemen, Transformieren und Zusammenführen dieser Daten logisch oder physisch und Bereitstellung durch einen integrierten Ansatz für Analysezwecke.
- **Unterstützung des Master Data Management (MDM) –** Ermöglicht die Konnektivität und Integration von Daten, die wichtige Geschäftseinheiten und Domänen wie Kunden, Produkte und Mitarbeiter darstellen. Datenintegrationstools können verwendet werden, um die Datenzugriffs- und Synchronisierungsprozesse zu erstellen, um verschiedene MDM-Tools und -Initiativen zu unterstützen.
- **Datenkonsistenz zwischen betrieblichen Anwendungen –** Datenintegrationstools bieten die Möglichkeit, die Konsistenz auf Datenbankebene sowohl unternehmensintern als auch unternehmensintern sowie bidirektional oder unidirektional sicherzustellen.
- **Gemeinsame Nutzung** von Daten über Unternehmen – Unternehmen sind zunehmend verpflichtet, Daten an externe Handelspartner (Kunden, Lieferanten, Geschäftspartner und andere) bereitzustellen und von ihnen zu empfangen. Dieses Nutzungsszenario hat während der COVID-19-Pandemie Aufmerksamkeit erregt, da Unternehmen sich auf ihre Datenintegrationstools verlassen, um integrierte Daten sowohl intern als auch extern für das Überleben von Unternehmen zur Verfügung zu stellen.
- **Data Services-Orchestrierung –** Die Möglichkeit, alle Aspekte der Laufzeitdatenintegrationsfunktionalität als Datendienste bereitzustellen (z. B. können bereitgestellte Funktionen über eine Webdienstschnittstelle oder eine API oder über Microservices aufgerufen werden).
- **Unterstützung für Datenmigration und -konsolidierung –** Datenintegrationstools decken zunehmend die Datenverschiebungs- und Transformationsanforderungen der Datenmigration und -konsolidierung ab, z. B. den Austausch von Legacyanwendungen oder die Migration in neue Computerumgebungen. Dieser Anwendungsfall gewinnt an Zugkraft, da Unternehmen versuchen, ihre Datenbestände über mehrere CSPs oder für Hybridnutzungsszenarien in die Cloud zu verschieben.

Datenintegrationstools sind erforderlich, um viele der Kernfunktionen der Datenintegration auszuführen, die auf jedes der oben genannten Szenarien angewendet werden können. (Eine detaillierte Liste und Erläuterung aller Bewertungskomponenten und Kernfunktionen von Tools im Datenintegrationsmarkt finden Sie unter Anmerkung 2).

Der Markt für Datenintegrationstools besteht aus Anbietern, die Softwareprodukte anbieten, die den Aufbau und die Implementierung von Datenzugriffs- und Datenbereitstellungsinfrastrukturen

für eine Vielzahl von Integrations-Anwendungsszenarien ermöglichen.

**Diese Definition umfasst keine Open-Source-Frameworks, allgemeine Entwicklungsplattformen oder Programmierschnittstellen. Solche Datenintegrationsframeworks oder -plattformen, die "allgemein" sind, und solche, die von Entwicklern stark angepasst werden müssen, um sie für bestimmte Datenintegrationsszenarien zu entwickeln, sind von diesem Magic Quadrant ausgeschlossen. Zu den in diesem Magic Quadrant bewerteten Anbietern gehört mindestens ein kommerzielles Standardtool, das speziell für die Datenintegration und -transformation entwickelt wurde.**

## Magic Quadrant

### Abbildung 1. Magic Quadrant für Datenintegrationstools


Quelle: Gartner (August 2020)

## Stärken und Vorsichtsmaßnahmen der Anbieter

### Actian

Actian ist ein Nischenspieler in diesem Magic Quadrant; in der vorherigen Iteration dieser Forschung, es war auch ein Nischen-Player. Actian mit Sitz in Palo Alto, Kalifornien, USA, bietet das DataConnect-Produkt als primäres Datenintegrationstool an. Seine Geschäftstätigkeit ist geografisch vielfältig und hat einen Kundenstamm von über 8.000. Die Kunden von Actian sind in der Regel B2B- und OEM-Partner von Unternehmen.

### Stärken

- **Stabiler Performer:** Actians leichtes und kleines Datenintegrationstool hat einen treuen Kundenstamm. Referenzkunden nannten die Benutzerfreundlichkeit, Zuverlässigkeit und Die Fähigkeit des Anbieters, komplexe Workloads zu bewältigen. Referenzen hoben auch die einfache Einbettung von Actian DataConnect in Anwendungen hervor.

- **Flexibilität bei der Preisgestaltung und Bereitstellung:** Actian hat einen vereinfachten, vorhersehbaren Ansatz für Preis- und Bereitstellungsoptionen gewählt, der die Unterstützung unbegrenzter Benutzer zu einem konsistenten Preis und die Bereitstellung universeller Konnektivitätsoptionen umfasst. Darüber hinaus können Benutzer während der Abonnementlaufzeit jederzeit zwischen den lokalen und Cloud-Optionen von Actian migrieren.

- **Verarbeitungsoptimierung:** Das Datenintegrationstool von Actian führt Inline-Statistiken für Daten, die die Integrationsplattform kreuzen. Dies ist nach wie vor eine Stärke – die Kombination von Kapazität, Auslastung, Datenstatistiken, Datenprofilerstellung und vielen anderen Komponenten, um eine Kombination von betrieblichen Warnungen für den Systemzustand und hinsichtlich Änderungen der Daten für Benutzer und Entwickler zu erstellen.

### Vorsichtsmaßnahmen

- **Fehlende rollenbasierte Bereitstellung:** DataConnect konzentriert sich auf traditionelle Datenintegrationsexperten, die Integration als Teil einer Anwendungsentwicklung oder in der Kapazität der Unterstützung von Data Engineering bereitstellen. Die Roadmap von Actian unterstreicht jedoch die Benutzerfreundlichkeit durch die Einführung neuer Connectors und Vorlagen, einen visuellen Abfrage-Generator und eine vereinfachte Schnittstelle zur Unterstützung unterschiedlicher Rollen (einschließlich Bürgerintegratoren). Dies ergänzt die Vorlagen und den Assistenten für den geführten Workflow, die Ende 2019 hinzugefügt wurden.

- **Die installierte Basis dient in erster Linie der Massen-/Batchintegration:** Basierend auf Referenz-Kundenantworten bleibt der primäre Datenintegrationsstil von Actian massen-/batch-orientiert. Während die meisten Organisationen mit der Massen-/Batch-basierten Datenintegration beginnen, kann dies ein einschränkender Faktor sein, wenn sie Massen/Batch-Vorgänge – durch Extraktion, Transformation und Laden (ETL) – mit anderen modernen Integrationsstilen wie Datenvirtualisierung kombinieren oder interoperieren müssen. Actian muss seine Datenbereitstellungsmethoden um Datenvorbereitung, Datenvirtualisierung und andere moderne Datenintegrationsstile erweitern, wenn es in neue Bereitstellungskanäle expandieren möchte.

- **Diversifizierender Produktfokus:** Actian entwickelt sein DataConnect-Portfolio weiter, um eng mit der Entwicklung seines Cloud-basierten Data Warehouse – Avalanche – in Einklang zu stehen. Während Avalanche Connect auf DataConnect basiert, kann das neue Avalanche Connect-Angebot Entwicklungs- und Marketingressourcen von DataConnect abrufen.

### Adeptia

Adeptia ist ein Nischenspieler in diesem Magic Quadrant; in der vorherigen Iteration dieser Forschung, es war auch ein Nischen-Player. Adeptia mit Sitz in Chicago, Illinois, USA, bietet Adeptia Connect als Datenintegrationsprodukt an. Die Aktivitäten konzentrieren sich hauptsächlich auf Nordamerika und EMEA. Es hat einen Kundenstamm von über 1.500 in verschiedenen Branchen, von denen die meisten in den Bereichen Versicherungen, Finanzen und Fertigung sitzen.

#### Stärken

- **Integriertes Produkt:** Die Datenintegrationstechnologie von Adeptia konzentriert sich auf die Ermöglichung der gemeinsamen Nutzung von Daten, Self-Service-Integration und Geschäftspartnermanagement. Mit Funktionen, die über ein einziges Produkt verfügbar sind, der engen Integration der zugrunde liegenden Komponenten und der Fähigkeit, eine schnelle Implementierung zu unterstützen, entsprechen sie den Anforderungen geschäftsorientierter Rollen und den B2B-Datenintegrationsanforderungen.

- **Verbesserte Benutzerfreundlichkeit durch bessere Datenvorbereitung und erweiterte Integration:** Die Effizienz von Adeptia Connect bei der Erstellung von Datenpipelines wird durch Verbesserungen bei der großen dateibasierten Datenerfassung, die Überwachung des Datenflusses des Remote-Zugriffs und die Verwendung von Machine Learning (ML) zur Unterstützung der Datenzuordnung verstärkt. Pläne für die Verwendung von erweiterter Datenintegration zum Generieren von Datenschemas und zum Introspektivieren von Metadaten aus Datenquellen erhöhen den Fokus auf die Bereitstellung von Geschäftsrollen für Integrationsaufgaben.

- **Günstige Kundenbindung:** Eine positive Gesamterfahrung mit Adeptia spiegelt sich im Kundenfeedback, im Vorverkauf und in der Implementierungslieferung wider. Die Auswahl der Datenintegrationstools des Anbieters wird häufig durch eine starke bestehende Beziehung beeinflusst.

#### Vorsichtsmaßnahmen

- **Begrenzte Reichweite:** Einige Referenzkunden berichteten von Herausforderungen mit Adeptia Connect in Bezug auf Interoperabilität und integrierte Nutzung mit verwandten Technologien für MDM. Einige Verweise berichteten über Bedenken hinsichtlich der Interoperabilität der Lösung mit externen Datenverantwortlichkeits-, Governance-, Datenqualitäts- und Metadatenmanagement-Tools.

- **Begrenzter Marktanteil:** Während das Bewusstsein für Adeptia wächst, besteht nach wie vor ein Mangel an Unternehmensneigung, die Datenintegrationstools dieses Anbieters in komplexen Implementierungen zu verwenden, was ein Hindernis für die unternehmensweite Einführung schafft. Nutzer des Kundenabfragedienstes von Gartner erwähnen diesen Anbieter in Wettbewerbssituationen nur selten.

- **Implementierungsleitlinie und -unterstützung:** Bereitstellungen in immer komplexeren Szenarien wecken die Erwartungen der Kunden an die Implementierungsunterstützung und Anleitung von Adeptia für Best Practices. Referenzkunden identifizierten jedoch die einfache Diagnose, Problemlösung und Upgrade-Handling als Bereiche, die verbessert werden müssen.

### Denodo

Denodo ist ein Anführer in diesem magischen Quadranten; in der vorherigen Iteration dieser Forschung war es ein Challenger. Denodo hat seinen Sitz in Palo Alto, Kalifornien, USA. Das Datenintegrationsangebot Denodo Platform konzentriert sich auf die Bereitstellung logischer Daten-Fabric-Funktionen. Denodo ist geografisch vielfältig und hat einen Kundenstamm von über 800 vor allem in den Bereichen Finanzdienstleistungen, Fertigung und Gesundheitswesen.

#### Stärken

- **Unterstützung für Die Data Fabric-Architektur:** Denodo bietet seinen Kunden die notwendigen Produktfunktionen für die Automatisierung des Data Fabric-Designs mit seinen Kernplattformkomponenten – einem einheitlichen semantischen Katalog, einer dynamischen Abfrageoptimierungs-Engine und Laufzeitmetadaten-basierten ML-Algorithmen. Das Data Fabric-Design basiert auf Datenvirtualisierung, um Geschäftsanwendern schnell integrierte Daten bereitzustellen, um schnellere Ergebnisse zu erzielen.

- **Kontextbezogene Informationsbereitstellung:** Der Denodo-Katalog dient als zentrale Anlaufstelle für die Durchsetzung von Sicherheits- und Governance-Richtlinien über Unternehmensdatenressourcen und die Eliminierung des direkten Zugriffs auf physische Datenspeicher. Die integrierten Datenvorbereitungsfunktionen vereinfachen selbstdienende analytische Aktivitäten für Die Bürger. Das eingebettete ML-Notebook ermöglicht es Spezialisierten, ihre Erzählungen zu erstellen und zu teilen, indem Abfragen, Skripte und Grafiken kombiniert werden.

- **Umfangreiche Markteinführungspartnerschaften und -ergebnisse:** Mit über 250 Systemintegratoren (wie Accenture, Cognizant und Deloitte) verfügt Denodo über einen breit gezurteilen Pool an qualifizierten Fachkräften. Referenzkunden bewerteten es mit mehreren Service- und Support-Metriken wie qualitätder Endbenutzerschulungen, rechtzeitiger Reaktion durch technischen Support und Zuverlässigkeit des Vertriebsteams.

#### Vorsichtsmaßnahmen

- **Overhead für das Infrastrukturmanagement:** Die Referenzkunden von Denodo wiesen auf ein angemessenes Maß an kontinuierlichen Anstrengungen bei der Verwaltung der zugrunde

liegenden Infrastruktur und der Anwendung von Produktaktualisierungen hin. Denodos neueste Version – Denodo Platform Version 8.0 – behebt einige dieser Probleme mit transparenten Ressourcenverwaltungsfunktionen wie Lastenausgleich und autoscaling.

- **Fehlende leicht erzwungene Leitplanken für** die Leistung: Referenzkunden mit komplexen verteilten Architekturen wiesen darauf hin, dass zusätzliche Prozesse implementiert werden müssen, um sicherzustellen, dass ihre verteilten Abfragen innerhalb der Plattform ausgeführt werden. Diese zusätzlichen Prozesse scheinen auf eine unzureichende Leistungsdiagnose zurückzuführen zu sein.

- **Herausforderungen bei DataOps und Orchestrierung:** Referenzkunden wiesen auf mehrere periphere Probleme hin, wie z. B. mehrere Clientkomponenten, mangelnde Unterstützung für erweiterte Datenqualitätstransformationen, Probleme bei der Git-Integration, Planungsprobleme, manuelle Formatierung von Datumsattributen und Probleme bei der Integration von Analysetools. Denodo v.8 behebt einige dieser Probleme mit einer einzigen Konsole, integriertem Planer und verbesserter Integration mit Analysetools.

### Fivetran

Fivetran debütiert als Nischenspieler in diesem Magic Quadrant. Sie hat ihren Sitz in Oakland, Kalifornien, USA. Das Datenintegrationsangebot, auch Fivetran genannt, konzentriert sich auf die Bereitstellung von Cloud-First-Datenintegrationsfunktionen. Die Aktivitäten des Anbieters konzentrieren sich in erster Linie auf Nordamerika mit einer wachsenden Präsenz in EMEA und APAC. Es hat einen Kundenstamm von über 1.200 in den Bereichen Software-Dienstleistungen, kommerzielle und professionelle Dienstleistungen und Verbraucherdienstleistungen.

#### Stärken

- **Zweck für cloud-first Data Integration:** Fivetran ist ein vollständig verwaltetes Extrakt-, Load-, Transformations-Angebot (ELT), das SaaS und lokale Datenspeicher mit Cloud-DBMs integriert. Die Benutzerfreundlichkeit, die einfache Unterstützung, Automatisierungsfunktionen wie Schemasynchronisierung, selbstheilende Pipelines mit automatischer Wiederherstellung und verwaltete Reaktion machen die Erfassung von Cloud-Daten sowohl für Dateningenieure als auch für Analysten relativ einfach. Referenzkunden sagten, dass sie Fivetran als zusätzliches Datenintegrationstool verwenden, um Cloud-Aufnahmelücken mit ihrem traditionellen Toolset zu schließen.

- **Niedrige Gesamtbetriebskosten (TCO):** Fivetran reduziert die Cloud-Erfassungszeit und -wartung, indem Schemadrifts zwischen Quellen und Zielen verwaltet werden, wodurch es für Geschäftsteams einfach ist, die Integration mit minimalem IT-Support durchzuführen. Fivetran wendet ein verbrauchsbasiertes Preismodell auf die Anzahl der verarbeiteten Zeilen an. Seine Fähigkeit, die Schemadrift-Erkennung vollständig mit Auflösungsmanagement in Client-Umgebungen zu verwalten, verbessert die Zuverlässigkeit der Datenreplikation und senkt die Kosten für seine Kunden.

- **Go-to-Market-Partnerschaften konzentrierten sich auf den Kundenerfolg:** Fivetran arbeitet mit großen Cloud-Dienstleistern (z. B. AWS, Google Cloud Platform und Microsoft Azure), Anbietern von Cloud-Software (z. B. Snowflake, Databricks, Looker, Tableau) und Systemintegratoren (z. B. Slalom, Accenture) zusammen, um gezielte Kundenerfolgsprogramme zu erstellen. Referenzkunden bestätigten die Bedeutung von Fivetran bei der Erreichung seiner Cloud-Strategie "Land and Expand".

### Vorsichtsmaßnahmen

- **Funktionslücken:** Obwohl Fivetran seinen Kunden über 160 Connectors zur Verfügung stellt, unterstützt es derzeit begrenzte Datenintegrationsstile (z. B. Datenvirtualisierung wird nicht unterstützt), eingeschränkte Quellkonnektivität (z. B. kann derzeit nicht von API-Management-Plattformen oder NoSQL-Repositorys gelesen werden), eingeschränkte Zielkonnektivität (z. B. Cloud-Objektspeicherziele wie AWS S3 sind immer noch Teil der Roadmap) und begrenzte Daten-Governance- und Datenqualitätsfunktionen für Unternehmen.

- **Keine Unterstützung für lokale Datenintegrationsworkloads:** Fivetran eignet sich für die Synchronisierung cloudbasierter oder lokal erarbeiteter Datenquellen mit Cloud-Zielen, unterstützt aber derzeit keine Datenintegrationsanforderungen, die nur innerhalb des lokalen Unternehmensgeländes vorgesehen sind. Wenn solche Organisationen jedoch bereit sind, in die Cloud zu wechseln, wird Fivetran zu einem schnellen Enabler.

- **Begrenztes Marktbewusstsein:** Basierend auf den Daten von Gartner-Anfragen und unserem Vorschlagsüberprüfungsservice tritt Fivetran nicht häufig in Wettbewerbssituationen auf. Die gezielten Kundenerfolgsprogramme und Go-to-Market-Partnerschaften sollten jedoch das Marktbewusstsein rechtzeitig verbessern.

## Hitachi Vantara

Hitachi Vantara ist ein Nischenspieler in diesem Magic Quadrant; in der vorherigen Iteration dieser Forschung, es war auch ein Nischen-Player. Hitachi Vantara mit Sitz in Santa Clara, Kalifornien, USA, bietet Pentaho Data Integration (PDI) und Lumada Data Catalog (aus der Akquisition von Waterline Data) im Rahmen des Lumada Data Services-Portfolios an, das Analyse- und Datenmanagementtechnologien umfasst. Die Geschäftstätigkeit des Anbieters ist geografisch vielfältig und verfügt über einen Kundenstamm von über 880 Organisationen, vor allem in den Bereichen Technologie und Finanzdienstleistungen.

### Stärken

- **Erweiterung des Produktportfolios:** PDI wird von Referenzkunden für die Bereitstellung von Funktionen zur Massen-/Batchdatenverschiebung anerkannt, die nun die umfassenderen Datenverwaltungsfunktionen von Lumada Data Services ergänzen, einschließlich Datenkatalogisierung und Data Lake Management. Eine containerisierte Version der Pentaho-Engine wird bei der Verwendung von PDI unterstützt, was Benutzern jetzt mehr Optionen bei der Bereitstellung von Funktionen als Teil von Lumada ermöglicht.

- **Edge-to-Core-to-Cloud-Datenfokus:** Mit einem Fokus auf Datenintegrationsanforderungen im Back-Office-Kern, am Edge und in Multicloud-Umgebungen erweitert Hitachi Vantara seine einsatzbezogene Nutzung zur Verbesserung der Erfassung von Maschinendaten, Metadatenerfassung, IoT-Datenintegration und -Analyse im Internet der Dinge. Referenzkunden lobten das Lumada Data Catalog-Tool für seine Tagging-, Inventarisierungs- und Suchfunktionen sowie für seine Fähigkeit, die Metadaten (innerhalb des Katalogs) mit PDI zu teilen, wodurch die Produktivität von Dateningenieuren verbessert wurde, die IoT-Daten integrieren möchten.

- **Anpassung an DataOps-Anforderungen:** Hitachi Vantara erweitert seinen Integrationsumfang für die Automatisierung von Datenpipelines, Data Governance und DataOps-Unterstützung für den Zugriff auf und die Verteilung von Daten, indem es ermöglicht, entwickelte Artefakte einbettbar zu machen und als Microservices verfügbar zu machen.

### Vorsichtsmaßnahmen

- **Mangelnde Klarheit und Marktzug bei weiterentwickelten Angeboten:** Die kombinierten Vorteile von PDI, zusammen mit Hitachi Vantaras Portfolioentwicklung (einschließlich erworbener Datenmanagement-Technologie), werden von den Interessenten bei Wettbewerbsbewertungen oft nicht sofort verstanden. Infolgedessen bleiben einige Kunden unsicher über die breiteren Vorteile, wenn sie mit dem Anbieter in Kontakt treten.

- **Kundenerfahrung:** Einige Referenzkunden berichteten von Problemen bei der Versionsaktualisierung, was darauf hindeutet, dass Hitachi Vantara die einfache Einrichtung und Migration, die integrierte Nutzung des Portfolios, die Fehlerbehebung und den Zugriff auf technischen Support und Dokumentation des Produkts verbessern muss.

- **Implementierungsunterstützung und Wertschöpfung:** Wo Bereitstellungen mit Hitachi Vantara-Datenintegrationstechnologien immer komplexer werden, haben Referenzkunden ihre Besorgnis über die Qualität der Implementierungsunterstützung zum Ausdruck gebracht. Verbesserte Anleitungen für Übungen und Bereitstellung sind ebenfalls wünschenswert, um die erweiterte Verwendung von Tools zu vereinfachen und die Zeit zu beschleunigen, um den Wert zu verbessern.

## IVs

HVR debütiert als Nischenspieler in diesem Magic Quadrant. HVR mit Sitz in San Francisco, Kalifornien, USA, bietet ein gleichnamiges Datenintegrationsprodukt an, das hauptsächlich die protokollbasierte Änderungsdatenerfassung (CDC) unterstützt. Der Kundenstamm des Anbieters für dieses Produkt besteht aus 324 Organisationen mit über 1.000 Bereitstellungen. Die Geschäftstätigkeit liegt hauptsächlich in Nordamerika und EMEA, und seine Kunden liegen hauptsächlich im verarbeitenden Gewerbe, im Einzelhandel und im Finanzsektor.

### Stärken

- **Treuer Kundenstamm und Marktdynamik:** HVR verzeichnete 2019 eine hohe Kundenbindungsrate bei gleichzeitig deutlich wiederkehrendem Umsatzwachstum. Dies deutet auf eine hohe Kundenzufriedenheit und ein hohes Wachstum innerhalb bestehender Konten hin. Das Wachstumspotenzial wird auch durch einen Mangel an ausgereiften und unabhängigen Wettbewerbern im Datenreplikations- und Synchronisierungsbereich verstärkt.

- **Replikation von Daten über häufig verwendete Endpunkte für Cloud-Migrationen:** HVR unterhält Partnerschaften mit AWS, Microsoft, Snowflake und Google Cloud Platform für eine optimierte Zielkonnektivität und Integration mit SQL Server, Oracle und SAP für Quell- und Zielkonnektivität. HVR wird daher für Replikation und Live-Vergleich und Reparatur (für eine effektive Schemadrift-Auflösung) in diesen Systemen als Teil der Daten- und Analysemodernisierungsinitiativen seiner Kunden verwendet. Für SAP erfasst HVR Daten aus SAP HANA- und ECC-Systemen durch die Integration in SAP-Wörterbücher, um Echtzeitschemas für Pool- und Clustertabellen abzurufen, ohne sich auf ABAP oder BAPIs zu verlassen. HVR bietet auch ein Agenten-Plug-in, um einen ständig wachsenden Pool von Konnektivitätsoptionen zu unterstützen.

- **Gute Serviceerfahrung:** Die Referenzkunden von HVR bewerteten es überdurchschnittlich für die "zeitnahe und vollständige Beantwortung von Produkt- oder Servicefragen", wobei die Vertriebsingenieure als äußerst kompetent und professionell gelobt wurden.

### Vorsichtsmaßnahmen

- **Fehlende Fokussierung von Geschäftsanwendern:** Die meisten Referenzkunden wiesen darauf hin, dass ihre Ausgaben für die Datenintegration mit HVR zentral-IT-gesteuert waren und von Den Geschäftsbereichen nur minimal beteiligt waren. HVR verfügt noch nicht über eine eigenständige oder eingebettete Datenaufbereitungslösung. Sie müsste daher diese Lücke in der Funktionalität und im Go-to-Market-Ansatz schließen, um ihren gesamten adressierbaren Markt deutlich zu erweitern, zumal sie schrittweise auf die Einführung einer SaaS-Lösung auf dem Markt ausgerichtet ist.

- **Es fehlt die Unterstützung für komplexe ETL-Workloads:** HVR verfügt nicht über Funktionen zur Adressierung anderer Datenbereitstellungsstile als der Replikation. Für komplexe Massen-/Batch-ETL-Workloads gibt es Partnerschaften mit Talend und Matillion. Die eingeschränkte Fähigkeit, Metadaten bidirektional mit anderen Tools auszutauschen, bedeutet jedoch, dass Kunden HVR mit anderen Datenintegrationstools zusammenfügen müssen, um die gesamte Bandbreite der Datenintegrationsstile zu ermöglichen. Dies spiegelt sich in der Tatsache wider, dass Referenzkunden HVR deutlich unter dem Durchschnitt für "Grad der Offenheit" bewertethaben, obwohl der Anbieter eine erweiterbare Architektur hat.

- **Extrem eingeschränkte Unterstützung von Daten- und Analytics-Governance:** HVR verfügt nicht über eine sofort einsatzbereite Integration mit Tools, die Daten-Governance-Funktionen wie Datenqualität, Datenkatalogisierung, Informationsverantwortung und Datenschutz bereitstellen. Sie stellt diese Fähigkeiten auch nicht selbst zur Verfügung.

### Ibm

IBM ist führend in diesem Magic Quadrant; in der vorherigen Iteration dieser Forschung war es auch ein Leader. IBM Information Server for IBM Cloud Pak for Data, IBM Cloud Pak for Integration, IBM Data Replication, IBM Data Virtualization Manager für z/OS, IBM InfoSphere Information Server on Cloud und IBM InfoSphere DataStage on Cloud zielen auf eine Reihe von Unternehmensdatenintegrationsszenarien ab. Der Kundenstamm des Kreditors für diesen Produktsatz besteht aus mehr als 11.000 Organisationen. Seine Aktivitäten sind global, und seine Kunden sind in der Regel Enterprise B2B und B2C Organisationen.

#### Stärken

- **Support und Performance:** Referenzkunden lobten die Unterstützung von IBM für seine Datenintegrationstools, insbesondere die Zugänglichkeit von Experten und die Reaktionsfähigkeit des Vertriebsteams. Referenzkunden betonten wiederholt die Skalierbarkeit und Leistung des Produkts.
- **Diverse data integration delivery styles and architectures:** Reference customers said they use IBM's products for traditional and complex upcoming data delivery styles, although the products see less use for data virtualization. The range of connectors, including for cloud object stores and a range of DBMS options, were praised by multiple reference customers for their ability to deliver on hybrid (on-premises and cloud) integration scenarios.
- **Repackaged integration capabilities:** IBM has made its data integration suite more consumable through the introduction of Cloud Pak for Data (based on Red Hat OpenShift), an offering that delivers all key data integration functionality. IBM's vision with Cloud Pak for Data is a composable, microservices-based architecture allowing multiple roles access to several data integration delivery styles, including data virtualization, streaming data integration and data replication/synchronization. Cloud Pak for Data can be deployed on all major CSP infrastructure and is not limited to IBM Cloud.

#### Cautions

- **New packaging confusion:** IBM recently repackaged its data integration tools into IBM Information Server for IBM Cloud Pak for Data, transitioning from the InfoSphere Information Server platform. Early questions from Gartner clients indicate confusion about what IBM Information Server for IBM Cloud Pak for Data comprises, what it costs, where it can be run and how it is deployed.
- **Architecture and upgrade complexity:** Several surveyed reference customers cited challenges upgrading some or all of their IBM data integration footprint, with some requesting a focus on container-based upgrade options. While IBM's recently introduced Cloud Pak for Data is built on a containerized base, it will require migration to a new offering, which will not ease upgrade challenges.
- **Modern deployment and usage patterns:** IBM received some of the lowest reference survey scores for its self-service capabilities and its ability to deliver data integration capabilities as

cloud services.

### Informatica

Informatica is a Leader in this Magic Quadrant; in the previous iteration of this research, it was also a Leader. Informatica is headquartered in Redwood City, California, U.S. It offers the following data integration tools as part of its Intelligent Data Platform: Informatica Intelligent Cloud Services, which includes Cloud Data Integration, Cloud Data Integration Elastic, and Cloud Mass Ingestion; Data Integration Hub; Data Engineering Integration; Enterprise Data Preparation; Enterprise Data Catalog; Edge Data Streaming; Data Engineering Streaming; PowerCenter; and PowerExchange. Informatica has over 10,000 customers for these product lines. Its operations are geographically diversified and its clients are primarily in the financial services, healthcare and public sectors.

*Note: As of 18 July 2020, Informatica announced completed acquisitions of Compact Solutions and GreenBay Technologies. Gartner will provide additional insight and research to clients as more detail becomes available. Reflection of these acquisitions is excluded from this research as they occurred after the cut-off date for the analysis.*

#### Strengths

- **Strong data fabric and augmented data integration support:** Informatica has made significant investments in various data-fabric-enabling technologies. These include augmented data cataloging, support for knowledge graphs, support for multiple data integration styles, and an active metadata-driven AI/ML engine (CLAIRE) that supports integration design and delivery automation. Reference customers provided Informatica with the highest score for augmented data integration and active metadata support.

- **Focus on capabilities convergence and CSP-agnostic integration:** Informatica has responded well to the market's demand for convergence in data management capabilities and is delivering this through its Intelligent Data Platform. Reference customers chose Informatica for its well-integrated capabilities beyond data integration, which also include support for data quality, information stewardship, metadata management (including data cataloging) and MDM. Reference customers gave Informatica some of the best scores for its multicloud/hybrid integration capabilities and frequently selected it for its commitment to CSP-neutral data integration. Notable partnerships include Microsoft Azure, AWS, Databricks, Google Cloud Platform, Salesforce, SAP and Snowflake.

- **Ability to scale in support of complex data integration scenarios:** Informatica is frequently selected for its scalable data integration capabilities. More than one-third of reference customers reported using its tools for integrating more than 100 data sources and/or targets. Also, over one-third reported using Informatica across four or more departments/projects. Finally, over half of Informatica's reference customers stated that the vendors' data integration tools were being used across their organizations by 50 people or more.

### Cautions

- **Lack of clarity on data virtualization support:** Even though Informatica supports virtual data access through a combination of APIs and microservices (through its Data Integration Hub and Enterprise Data Catalog tools), reference customers and Gartner clients are not fully aware of those modern data virtualization capabilities. In fact, many Gartner clients enquired about competing data virtualization tools, and Informatica's own reference customers scored the vendor well below average for its data virtualization capabilities.
- **Confusion regarding stream data integration support:** Informatica supports stream data integration and data replication primarily through a combination of its Edge Data Streaming, Cloud Mass Ingestion and Data Engineering Streaming services. However, a significant number of Gartner clients inquiring about Informatica were confused or simply unaware about the capabilities, usage scenarios and/or best practices to deliver and expand streaming architectures using a combination of Informatica's existing tools. Reference customers also scored Informatica below average for its capabilities supporting message-oriented data movement.
- **Lack of awareness regarding the DataOps capability in data engineering:** A small portion of Informatica's reference customers reported lack of awareness regarding its DataOps and orchestration capabilities. These had been made available in the early releases of Informatica's Data Engineering Integration and Data Engineering Streaming products.

## Information Builders

Information Builders (ibi) is a Visionary in this Magic Quadrant; in the previous iteration of this research, it was also a Visionary. Based in New York City, New York, U.S., ibi offers the Omni-Gen data integration platform. The technology comprises a central platform plus additional tools, including iWay Service Manager, iWay DataMigrator and iWay Universal Adapter Suite. ibi operates mostly in North America and EMEA, with a customer base estimated at 900 primarily in the healthcare, financial services and public sectors.

### Strengths

- **Business solution focus:** With a history of recognized adapters and solid connectivity options, ibi offers capabilities in diverse data integration styles and implementations. Templates and use-case frameworks delivered as part of the Omni-Gen solution set position the vendor nicely for integrated business solutions implemented through comprehensive data integration. This works especially when buyers are looking to forgo larger solution competitors in favor of smaller, more-focused vendors providing data integration functionality at a lower TCO.
- **Amplified data management support:** The applicability of ibi's product portfolio is broadened by its Omni-Gen platform evolution. This includes augmented data quality and enablement of data engineering for analytics through embedded data preparation capability, alongside data integration tooling to capitalize on data science and ML demand trends.

- **Adoption driven by strategic relationships:** ibi offers high-quality customer engagement. This is captured in customer feedback, which identifies the vendor's sales process and overall relationship as strong positives. Buyers often view ibi as a strategic partner and cite this as a key reason for selecting its data integration tools.

#### Cautions

- **Limited market traction:** Adoption of ibi's data integration tooling continues, although with a relative lack of mainstream recognition. Over the last year, leadership changes at the company have generated some uncertainty among prospective customers about the impact on product and market focus. This potentially constrains prospects' confidence regarding the vendor's market traction.
- **Product complexity:** While reference customers view the range of functionality and flexibility of ibi's products as positives, many cited product complexities, a longer learning curve and insufficient documentation as areas in which the vendor could improve.
- **Limited appeal to less-technical roles:** ibi's data integration tools are used by many data integration tool specialists, but not by those in less-technical roles. Going forward, the vendor's roadmap involves improving data preparation for business-oriented roles, easing the reuse of data and artifacts, and bringing multiple, disparate features into integrated offerings.

### Matillion

Matillion debuts as a Niche Player in this Magic Quadrant. It has dual headquarters in Manchester, U.K. and Denver, Colorado, U.S. It offers two data integration products — Matillion ETL and Matillion Data Loader — both focused on cloud-first data integration capabilities. Its operations are primarily focused in North America and EMEA. Matillion has a customer base of over 800 in the business services, software and manufacturing sectors.

#### Strengths

- **Impactful differentiators for cloud deployment:** Cloud data integration capabilities like migratable business logic, dynamic ETL, multicloud data movement and high-availability clustering have enabled several reference customers to select Matillion over established vendors for their cloud integration use cases.
- **Targeted tooling for various integration personas:** Matillion ETL is intended for technical users like data engineers, who should also be familiar with cloud-native services like networking, DevOps and configuring data workflows. Matillion Data Loader is positioned for analysts and citizen integrators. This role focus helps organizations to balance business exploration against governed delivery of data pipelines.
- **Flexible pricing model in support of cloud data integration scalability:** Matillion ETL applies a consumption-based pricing model (based on compute units) via the customer-preferred cloud

marketplace. This allows organizations that are adopting a cloud data management strategy to start small and expand as needed.

### Cautions

- **Limited support for diverse data integration styles and capabilities:** Matillion does not support stream data integration. Currently, it provides only limited data governance and data quality capabilities for enterprise customers (for example, data lineage improvements and support for PII data are part of its roadmap). Also, Matillion currently does not find enough traction for popular operational use cases like MDM for its data integration tooling among its reference customers.
- **No support for on-premises target systems:** Matillion is not suitable for those organizations whose data integration targets include on-premises data warehouses — i.e., currently, Matillion's usage is limited to cloud data warehouses' target systems.
- **Room to improve current product features:** Reference customers raised concerns around various product features including scheduling, parallel working, data streaming, varying UI standards, logging and process status communication. However, in general, reference customers praised the vendor for its regular upgrades through which it is improving critical functionality.

## Microsoft

Microsoft is a Challenger in this Magic Quadrant; in the last iteration of this research, it was also a Challenger. Its SQL Server Integration Services (SSIS) tool targets on-premises data integration tasks, while Azure Data Factory (ADF) targets Azure-based data integration tasks. The number of paid customers for its data integration tools is not disclosed by Microsoft. Its operations are geographically diverse, and its clients range from small and midsize businesses to enterprises.

### Strengths

- **Usability and TCO:** When asked why they chose Microsoft in competitive situations, reference customers cited its overall low TCO, speed of implementation, ease of use and ability to integrate with other Microsoft SQL Server capabilities/Azure data services as primary reasons.
- **Cloud-centric vision:** Microsoft's Azure Synapse product, which provides a unified data and analytics environment in the cloud, heavily leverages Azure's data integration functionality in a low-/no-code environment. This will appeal to Microsoft's large incumbent customer base as they migrate to the Azure environment.
- **Focus on a metadata-driven future:** While historically an area of weakness for Microsoft, the company is investing significantly in its metadata capabilities across its data integration tooling to give users a consistent experience in the Azure environment. Microsoft is making significant enhancements to the Azure Data Catalog that will allow users to inventory distributed data assets, track lineage and provide insights (such as recommendations and

alerts). The vendor is also investing in improved support for metadata (and insights) sharing between its data catalog and its data integration tools, thereby improving developer productivity.

### Cautions

- **Support for multicloud data integration:** Microsoft received the second-lowest score from reference customers for its ability to support multicloud data integration. Its focus has been on the Azure cloud environment and getting data into that environment, but its stance increasingly conflicts with the needs of many enterprises for bidirectional data integration capabilities.
- **Pipeline complexity and support for complex data pipelines:** Reference customers cited challenges implementing complex data integration pipelines with ADF — particularly when they vary or combine data delivery capabilities. ADF and SSIS largely target ETL and data replication workloads with limited support for virtualization, streaming and messaging services. Microsoft stated that it is investing in improvements to simplify data pipelines.
- **Scaling and troubleshooting challenges:** Multiple reference survey respondents and Gartner Peer Insights reviewers commented that ADF and SSIS had challenges scaling ETL workloads over millions of rows of data and that error messages are cryptic, making debugging and troubleshooting unnecessarily challenging.

## Oracle

Oracle is a Leader in this Magic Quadrant; in the previous iteration of this research, it was also a Leader. Based in Redwood Shores, California, U.S., Oracle offers Oracle GoldenGate, Oracle Data Integration Suite, Oracle Big Data SQL, Oracle Integration Cloud, Oracle Cloud Infrastructure (OCI) Data Integration, and OCI Data Flow. Oracle's customer base for these products is over 12,000 organizations. Its operations are geographically diversified, and its clients are primarily in the financial services, manufacturing and public sectors.

### Strengths

- **Well-positioned for supporting data fabrics:** Oracle has a strong product portfolio for supporting data fabric designs. This includes OCI Data Catalog, Linked Open Data Cloud for domain-specific knowledge graphs, NLP/NLG support for Oracle Analytics Cloud, low-code graph-based discovery with support for semantic enrichment, serverless data integration through OCI Data Integration, and microservices architecture support. Oracle is therefore in a very good position to support its customers' data fabric designs.
- **Focus on data management capabilities convergence:** Oracle's strategy for delivering key data management capabilities including data integration, data quality and data cataloging through a consolidated platform is praised by reference customers. Its data integration solutions — such as Oracle GoldenGate for replication and streaming and Big Data SQL for data virtualization — are available via the OCI Marketplace. This aids Oracle in its vision of providing a converged data management platform.

- **Enterprise-grade products reflected through market growth:** Oracle's reference customers scored it very high for "timely and complete" response to product or service questions, as well as "performance, scalability and throughput" of its products. Its data integration products continue to be well-regarded for their robustness and reliability, especially for dealing with complex operational workloads. This is evidenced by Oracle's solid data integration tools market growth.

### Cautions

- **Continued concerns with pricing:** Oracle received the lowest mean score of all vendors from reference customers for "pricing method" and "overall pricing." This is evidenced by the fact that none of its reference customers selected Oracle for TCO; organizations seem to prefer Oracle when cost is not the top priority. To partially address this challenge, Oracle Data Integrator (ODI) is now available in the OCI Marketplace free of charge for loading data into Oracle Database Cloud targets.

- **Limited usage beyond traditional use cases:** None of Oracle's reference customers reported using its data integration products for modern use cases such as IoT data integration and data services orchestration. Also, the proportion of reference customers utilizing Oracle's products to support a combination of more than one data delivery style is lower compared with other Leaders in this Magic Quadrant, with usage skewed heavily toward bulk/batch data movement, replication and synchronization.

- **Lack of a multicloud strategy:** While Oracle has embraced the use of microservices for a broader range of deployment options, and some customers operate these deployments in non-Oracle clouds, Oracle's plans for natively deploying its data integration services on other CSP infrastructures are currently limited. The interconnect partnership with Microsoft is certainly a step in the right direction.

## Precisely (formerly Syncsort)

Precisely (previously known as Syncsort) is a Challenger in this Magic Quadrant; in the previous iteration of this research, it was a Niche Player. Based in Pearl River, New York, U.S., the vendor offers the Precisely Connect product set (which includes Connect ETL and Connect CDC), Precisely Ironstream, Syncsort MFX and the Spectrum tools portfolio. Precisely's customer base for this product set is over 2,300 organizations. Its operations are mostly in North America and EMEA, and its clients are primarily in the financial services, insurance and healthcare sectors.

### Strengths

- **Focused on data management infrastructure modernization:** As organizations look to modernize their data and analytics landscapes, moving off mainframes is a common migration scenario. Precisely excels at extracting data from multiple mainframe systems, and the acquisition of SQData has given it access to IBM Information Management System (IMS) data as well. For data target connectivity, Precisely has entered strategic partnerships with Cloudera,

Snowflake and Databricks, among others. It is therefore well-positioned to support the popular cloud migration use case.

- **Strong technical and after-sales support:** Most reference customers praised Precisely for its strong support services. This included praise for the knowledge base of its support teams. For timely and complete response to product or service questions, and for quality of technical support and documentation, Precisely received the highest scores possible from reference customers.

- **Significantly wider data management coverage:** Through the acquisitions of Trillium Software and Pitney Bowes Software & Data, Precisely now provides well-integrated solutions for use cases needing a combination of data integration and data quality capabilities. Data enrichment and location intelligence capabilities provided by Pitney Bowes Spectrum further help differentiate Precisely from other vendors focused only on data integration.

#### Cautions

- **Limited active metadata support:** Use of metadata to support data management tasks is becoming increasingly mainstream among end users. Precisely received the lowest score among all vendors evaluated for this capability, especially due to the rapidly increasing complexity of its product portfolio. While Precisely does provide passive metadata management capabilities such as data lineage and impact analysis, it needs to expand its support for active metadata based insights, for more augmented data integration design and delivery.

- **Complex portfolio:** Precisely has undergone a major rebranding exercise to bring all of Syncsort's tooling and acquired solutions under a single banner. But some confusion remains around the use of distinct capabilities across the four key categories that Precisely calls "Integrate, Verify, Locate and Enrich." Precisely needs to chart out a long-term vision of how it expects customers to use its multiple point tools across various data integration use cases.

- **Need for more flexible deployment options:** Precisely cannot as yet be deployed as a complete integration platform as a service (iPaaS) solution. For customers needing to use the Connect product family for data integration, they currently must use Trillium as the PaaS and deploy the Connect tool on top of that. Precisely cannot be deployed as a serverless metered service or as a managed service. Containerization of individual capabilities within the Connect portfolio is also not yet available.

### Qlik (Attunity)

Qlik is a Challenger in this Magic Quadrant; in the last iteration of this research, it was also a Challenger. Based in King of Prussia, Pennsylvania, U.S., Qlik targets a range of data replication and metadata management tasks through its Qlik Replicate, Qlik Compose, Qlik Enterprise Manager and Qlik Catalog products. Its customer base for this product set is more than 3,000

organizations globally. Qlik's operations are predominantly based in North America and EMEA and its clients tend to be enterprises.

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

#### Strengths

- **Proven in data replication scenarios:** Qlik continues to be evaluated and selected by clients in most competitive situations for targeted data replication needs. Reference customers gave positive feedback for the vendor's robust CDC-based replication capabilities, and Qlik appeared very frequently in competitive evaluations for data replication across the whole reference survey.
- **Ease of use:** Reference customers repeatedly cited Qlik's ease of use — including ease of installation and configuration, automated code generation and nonintrusive implementation — as the top reasons for selecting the vendor in competitive situations.
- **Portfolio expansion and market momentum:** Qlik has successfully integrated its Attunity acquisition with the rest of its portfolio, including Qlik Data Catalyst (now Qlik Catalog) and Qlik Sense to support a consistent experience — from data ingestion/integration, to data cataloging, to real-time and batch analytics. Through its partner momentum (for example, strong OEM support with popular cloud infrastructure providers such as AWS and Microsoft Azure), Qlik has registered significant revenue and market mind share growth year over year.

#### Cautions

- **Limited traction beyond core scenarios:** Use of Qlik for data integration styles other than CDC-based data replication is yet to find traction. Qlik lacks proven deployments in data virtualization scenarios, which could be an issue for customers looking to use it to combine and interoperate between these data delivery styles (e.g., data replication with data virtualization).
- **Multicloud support:** Reference customers stated that support for connectors was uneven across all cloud infrastructure vendors, resulting in more complex multicloud integration scenarios than desired by end users.
- **Uneven deployment and support record:** Some reference customers called out the need to escalate support requests to elicit a response and resolution from Qlik. This was specifically cited around new product deployments and upgrades.

### Safe Software

Safe Software debuts as a Niche Player in this Magic Quadrant. Based in Surrey, British Columbia, Canada, Safe Software offers the FME platform, which includes FME Desktop and FME Server. The vendor's customer base for this product set is more than 6,500 organizations (excluding OEMs). Its operations are mostly in North America and EMEA, and its clients are primarily in the

government, utilities, energy (oil and gas) and architecture, engineering and construction (AEC) sectors.

### Strengths

- **Strength in spatial data integration:** Safe Software's FME continues to be evaluated and selected by reference customers for its strength in geospatial data ingestion, integration and sharing. Reference customers praised the vendor's extensive collection of spatial (with native support for Esri, GIS and OSIsoft) and 3D connectors with prebuilt transformations that provide relevance for location intelligence, interenterprise data sharing and predictive maintenance use cases for asset-heavy industries.
- **Strong execution scores and customer relationships:** Reference customers praised Safe Software for its overall ease of installation and setup, development and deployment activities, ongoing support, and quality of end-user training and support documentation. Safe Software's scores were in the top quartile of vendors in this Magic Quadrant for all these execution metrics in our customer reference survey.
- **Strength in self-service data preparation:** Safe Software provides FME Server Apps — a set of prebuilt integrations that allow citizen integrators and subject matter experts to develop and share data artifacts across organizations. Reference customers praised the vendor's low-code-centric UI and data preparation capabilities, which allow non-IT users to share integrated data with partners.

### Cautions

- **Limited traction beyond core usage scenarios:** While Safe Software does support all data integration use cases, its popularity and perceived strength seem to be extremely skewed toward scenarios involving spatial data. This is also evidenced by Gartner's client inquiry and proposal review services, where the vendor is rarely evaluated in competitive situations for data integration scenarios not involving spatial data or location intelligence.
- **Limited metadata management and active metadata support:** Safe Software has limited support for metadata synchronization or exchange with third-party data profiling, relationship discovery or information stewardship tools. As of the time of writing this research, the vendor's FME tooling does not support impact analysis reports or in-stream data profiling capabilities. Also, the ability to collect and analyze metadata (in order to activate metadata), to inform and even automate parts of integration, design and delivery is currently limited in FME.
- **Limited support for combining different data delivery styles:** Safe Software's FME platform currently lacks extensive support for integration style interoperability — i.e., support for combining and switching between two different data delivery styles (such as bulk/batch and data virtualization) through extensive data orchestration capabilities.

## SAP

SAP is a Leader in this Magic Quadrant; in the previous iteration of this research, it was also a Leader. Based in Walldorf, Germany, SAP's key data integration offerings — SAP Data Services, SAP Data Intelligence, SAP Cloud Platform Integration Suite, along with integration capabilities within the SAP HANA platform — are focused on delivering a comprehensive integration. Its operations are geographically diverse, with a customer base of over 70,000, and the top three industries it supports are oil and gas, consumer products, and retail.

### Strengths

- **Comprehensive data integration solutions:** SAP's data integration product set helps its customers realize an end-to-end integration strategy by combining and switching between multiple integration styles. Its Integration Solution Advisory Methodology guides customers in creating their own integration strategy.
- **Seamless data orchestration capabilities for various personas:** SAP Data Intelligence empowers integration specialists to build and manage cross-system data orchestration (such as data flows for ETL jobs, Apache Spark jobs, SAP modules or non-SAP external processes with scheduling and monitoring capabilities). Equally, the vendor enables citizen integrators with its simple design and self-serving capabilities to operationalize business user learnings.
- **Extensive go-to-market partnerships and high-quality customer service:** SAP has over 3,000 partners worldwide — distributors, OEMs, implementation and service partners. Reference customers confirmed a timely and complete response to product or service questions, while they also approved of SAP's quality of technical support and documentation.

### Cautions

- **New implementations and upgrades require a high level of expertise and SAP support:** Some reference customers indicated a lack of skilled expertise in the market for newer products like SAP Data Intelligence (formerly SAP Data Hub), where product support is limited to SAP directly. As a result, this market skills gap adds to customer implementation costs and ongoing support.
- **Limited integration with non-SAP products:** Some reference customers indicated that SAP's data integration product set has not been successful in replacing alternative tools, citing compatibility reasons. To an extent, this depends on the complexity of the customer data ecosystem. The introduction of SAP Cloud Platform Open Connectors (offered in SAP Cloud Platform Integration Suite) has started to address this issue.
- **Licensing and contract challenges:** Reference customers expressed concerns around navigating SAP's pricing strategy, licensing and contract negotiation process. However, in general they are positive about the long-term engagement with SAP.

## SAS

SAS is a Leader in this Magic Quadrant; in the previous iteration of this research, it was also a Leader. Based in Cary, North Carolina, U.S., SAS offers the following data integration products: SAS Data Management, SAS Data Integration Studio, SAS Federation Server, SAS/ACCESS, SAS Data Loader for Hadoop, SAS Data Preparation, and SAS Event Stream Processing. The vendor's customer base for this product set is around 15,000 organizations. Its operations are geographically diversified, and its clients are primarily in the banking, government and services sectors.

### Strengths

- **Enhanced support for diverse user personas:** SAS Studio Flow is a new web-based, low-code functionality, introduced within SAS Viya (the vendors' cloud-ready and microservices-based platform) for improving ease of use for developers designing ETL jobs. SAS Data Preparation, with its capability to augment and automate repetitive data wrangling tasks, provides additional support for citizen integrators. These tools place SAS in a good position to leverage the recent trends toward data democratization.

- **Improved augmentation and orchestration:** SAS has embedded ML-driven functions in its data integration, data quality and data preparation solutions for autorecommending next best actions. For orchestration, SAS Viya is available on containers and utilizes APIs for various hybrid and multicloud deployment scenarios.

- **Strong ecosystem of partners and peers:** Reference customers scored SAS well above average on availability of high-quality, third-party resources (e.g., implementation services, service providers) and the quality of its peer user community. While SAS's own technical and after-sales support has been praised in the past, this quality permeating through to a wider ecosystem of partners and peers bodes well for increased expansion opportunities.

### Cautions

- **Skewed support for SAS-based solutions:** SAS's data integration tools are still focused on supporting the company's own technology stack, such as SAS solutions for risk, fraud and retail. While tighter integration with SAS Viya modules is good for enhancing the SAS ecosystem experience, more focus on non-SAS solutions is expected. SAS has started addressing this through its collaboration with ODPi Egeria for bidirectional metadata exchange with non-SAS systems.

- **Lingering market perception of an "analytics only" focus:** While SAS has tried to educate the market on the utility of its data integration tools for operational use cases, the consequent shift in market perception has not yet materialized. This could be due to lack of awareness, as most SAS reference customers continue to use its data integration tools mainly in support of their analytics and/or data science use cases and less so for operational use cases like MDM, for example. In fact, data warehousing and data integration for supporting data science continue to be the top two use cases supported by SAS for its reference customers.

- **Absence of a holistic data management platform approach:** SAS provides many point solutions for data integration and data management. However, it lacks a narrative that explains how these myriad solutions can be used together to serve business-centric use cases. Other Leaders in this Magic Quadrant have already laid down this vision, either through delivering data integration services via a data fabric architecture, or through continuous delivery of data integration services via a DataOps framework. SAS therefore lags these Leaders in charting this vision.

### SnapLogic

SnapLogic is a Visionary in this Magic Quadrant; in the previous iteration of this research, it was also a Visionary. Headquartered in San Mateo, California, U.S., SnapLogic offers the SnapLogic Intelligent Integration Platform as its data integration offering. It operates mostly in North America and EMEA, with a customer base estimated at 930 organizations from diverse sectors including technology, financial services, healthcare, transportation and retail.

#### Strengths

- **Accelerating productivity through a cloud-first delivery model:** Since the inception of its business, SnapLogic has successfully pursued a targeted market segment focused on cloud-based integration use cases. Intelligent Integration Platform provides a mix of hybrid deployments catering for data and application integration, data engineering and integration automation requirements. A favorable TCO, productivity and time to value, and offerings provided on a single platform are key factors for buyer adoption.

- **Augmented data integration and sharing support:** SnapLogic's focus on augmented data integration provides automated guidance embedded to simplify the development of integration flows, with plans for enhancing the combination of ML and NLP for data preparation. The "flow" functionality introduced sets out to enhance augmented user experience using NLP, aimed at easing community sharing and creation of integration artifacts. These integration artifacts are useful for enhancing productivity, typically for incremental journeys of data sharing and process automation spanning customer, partner and employee domains.

- **Strong existing customer relationships:** Customer feedback during vendor evaluations revealed a positive overall experience with SnapLogic. Selection of the vendor's data integration tools is often influenced by a need for iPaaS.

#### Cautions

- **Challenges in deployment and diagnostic guidance:** Some reference customers identified debugging and resolution of error messages as challenges and stated that they require better diagnostic support and more consistency in the quality of technical support. Improvements to operations monitoring consoles and increasing the ease for operationalizing developed artifacts (for continuous delivery of integration jobs) are also desired by some reference customers. SnapLogic has made moves to address some of these issues by adding automated

customer regression testing, allowing customers to test within the endpoints, networks or other infrastructure environments of their own.

- **Limited resonance for data management support:** SnapLogic's current credentials are built more on its focus on and use in cloud and hybrid cloud service integration scenarios, rather than supporting broader data management solutions. This may adversely affect some competitive evaluations of its data integration tools.

- **Limited market mind share:** SnapLogic is one of the least evaluated data integration tool vendors in this Magic Quadrant, based on the reference customer survey. SnapLogic is looking to address this by leveraging co-marketing activities with new ISV partners like AWS, Snowflake and SAP, while also adding industry-specific solutions in target markets like human capital management (Workday, SAP) and spend management (Coupa, SAP).

### Talend

Talend is a Leader in this Magic Quadrant; in the previous iteration of this research, it was also a Leader. Talend is headquartered in Redwood City, California, U.S. Its data integration tools include: Talend Data Fabric; Talend Cloud; Talend Data Catalog; Talend Data Management; Talend Big Data; Talend Data Services; Talend Open Studio; and Talend Data Preparation and Data Stewardship (as add-on components). Talend has an estimated 4,500 licensed customers for this product line. Its operations are geographically diversified, and its clients represent companies in a variety of sectors such as media and services, financial services, and manufacturing.

#### Strengths

- **Focused on database-agnostic and multicloud hybrid data integration:** Talend's execution is focused on independent data integration across a multicloud and hybrid ecosystem. Reference customers cited Talend's strong support of multiple CSP infrastructure and DBMS-agnostic data integration as key reasons for selecting the vendor in competitive situations.

- **Strong capabilities for data engineering support:** Talend has invested significantly in the Talend Pipeline Designer, which allows data engineers to profile, prepare and transform data pipelines using the scalability of Apache Spark (as an underlying execution engine) and Apache Beam (for processing complex data pipelines). Integrated support for API services and management (through the acquisition of Restlet) makes it easy for data engineers to deliver integrated data views as loosely coupled microservices enabling effective data services orchestration.

- **Comprehensive portfolio in support of best-of-breed data management requirements:** Talend delivers a comprehensive portfolio of capabilities needed to deliver a data fabric architecture. These capabilities include data integration, application integration, data preparation, data cataloging, data stewardship and API management capabilities. It has expanded its data integration portfolio through the recent acquisitions of Stitch (for data ingestion into cloud repositories) along with organic development of its data catalog tool (for inventorying

connected data assets). These strategic investments position the vendor well for best-of-breed data management requirements.

### Cautions

- **DataOps support challenges:** Some reference customers reported challenges with the overall orchestration and operationalization of complex data pipelines in Talend's data integration tools. Reference customers requested better continuous integration (CI)/continuous delivery (CD), monitoring, visual impact analysis and Git support for integration flows developed in Talend.
- **Challenges with online documentation and community support:** Since Talend provides various features through commercially supported options of open-source projects (such as Kafka, Spark and Beam), its customers rely on extensive community support, as well as documentation support for scaling their projects. This is where some reference customers raised concerns over a lack of adequate community support and online documentation from Talend.
- **Limited traction for data virtualization:** Talend provides the ability to support distributed caching and federated queries (through SPARQL and APIs). But reference customers and Gartner clients currently exhibit limited understanding of which tool to use and how to deliver integrated views of data (without physical data movement or consolidation) using Talend's data integration tools portfolio. Talend's reference customers scored the vendor well below average for its data virtualization capabilities.

## TIBCO Software

TIBCO Software is a Challenger in this Magic Quadrant; in the previous iteration of this research, it was a Niche Player. TIBCO is based in Palo Alto, California, U.S. Its main data integration tools are TIBCO Data Virtualization (TDV), TIBCO Cloud Integration, TIBCO Cloud Events, TIBCO Messaging, TIBCO Streaming, and TIBCO Foresight. The vendor's customer base for this product set is more than 5,000 organizations. Its operations are geographically diversified, and its clients include companies in the financial services, telecommunications and manufacturing sectors.

### Strengths

- **Strength in real-time and edge data integration scenarios:** Reference customers praised TIBCO for its strength in integrating data at the edge. TIBCO provides a comprehensive portfolio of tools in support of event stream processing and edge data integration scenarios. This includes TIBCO Streaming (for stream data integration), TIBCO Messaging (for low latency and high throughput messaging) and TIBCO Eclipse Mosquitto Distribution (for connecting IoT devices using MQTT to TIBCO Messaging service).
- **"Connected Intelligence" positioning resonating with market vision:** TIBCO's strategy for data integration includes maintaining a balance between connecting and collecting data. Reference customers praised TIBCO Data Virtualization for its mature data virtualization capabilities for

reduced cost and time to data delivery. TIBCO has also invested in a new metadata management service (TIBCO Cloud Metadata), which enables data cataloging and active metadata analysis.

- **Strategic acquisitions improving integrated data management capabilities:** Certain strategic acquisitions in recent years have strengthened TIBCO's data management portfolio. These include Scribe Software in 2018, which provided ETL, iPaaS and data synchronization capabilities. TIBCO also acquired Orchestra Networks in 2018 and integrated its MDM and data governance capabilities into its broader Connected Intelligence strategy. Finally, TIBCO's 2019 acquisition of SnappyData (an in-memory DBMS) has improved the performance and caching capabilities of TDV.

### Cautions

- **Limited data preparation capabilities:** Even though TIBCO embeds self-service data preparation within TIBCO Spotfire, reference customers requested more comprehensive business-user-focused data preparation capabilities within its data integration tools portfolio. To alleviate this concern, TIBCO is adding a web-based self-service data provisioning UI to its data virtualization offering that will add data preparation and integration capabilities for business users. Reference customers rated TIBCO below market average for its data preparation functionality.

- **Pricing models and overall cost:** Some reference customers highlighted that TIBCO's pricing model is expensive and complicated to understand and forecast. They scored the vendor below the survey average for its pricing method and overall pricing (perceived value versus overall cost). To alleviate some of these concerns, TIBCO has transitioned from a perpetual to a subscription license model based on customer demand.

- **Portfolio complexity leading to integration style interoperability concerns:** TIBCO has a broad data integration tools portfolio (which has grown over the years through acquisitions) to address key data integration capabilities and use cases. Reference customers requested more clarity on how to orchestrate and switch between different data delivery styles across TIBCO's individual data integration tools. TIBCO has started to alleviate some of these concerns through its Connected Intelligence platform strategy. However, it needs to do more to simplify and better integrate its portfolio, where the focus is more on use cases rather than individual integration capabilities.

## 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

- Fivetran
- HVR
- Matillion
- Precisely (Syncsort) — formerly included as Syncsort before rebranding
- Safe Software

### Dropped

None

## Inclusion and Exclusion Criteria

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

- Vendors whose customer reference base could not administrate and orchestrate compatible data integration technologies for use in **at least three of seven key data delivery styles were excluded**. These styles are bulk/batch, message-oriented data movement, data services orchestration, data replication, data synchronization, stream data integration and data virtualization. We describe them (alphabetically with no reference to order of importance) as follows:
  - **Bulk/batch data movement** — Includes single pass or multipass/step processing that incorporates the entire contents of the data file after an initial input or read of the file is completed from a given source or multiple sources. All processes take place on multiple records within the data integration application before the records are released for any other data consuming application.
  - **Data services orchestration** — The ability to deploy any of the other data integration styles, but with the specific capability to interoperate with application services (logic flows, interfaces, end-user interfaces, and so on). Also, the ability to pass instructions to, and receive instructions from, those other services on the data services bus.
  - **Data virtualization** — The utilization of logical views of data, which may or may not be cached in various forms within the data integration application server or systems/memory managed by that application server. Data virtualization may or may not include redefinition of the sourced data.
  - **Data replication** — This involves simple copying of data and schema from one location to another, always in a physical repository. Replication can be a basis for all other types of data integration but, specifically, does not change the form, structure or content of the data it moves.

- **Data synchronization** — Can utilize any other form of data integration. However, it specifically focuses on establishing and maintaining consistency between two separate and independently managed create, read, update, delete (CRUD) instances of a shared, logically consistent data model for an operational data consistency use case. Synchronization also maintains and resolves instances of data collision, with the capability to establish embedded decision rules for resolving such collisions.
- **Message-oriented data movement** — This data delivery style utilizes a single record in an encapsulated object. This may or may not include:
  - Internally defined structure (XML)
  - Externally defined structures (electronic data interchange)
  - A single record that delivers data for action to the data integration process
  - Another source that delivers data for action to the data integration process
- **Stream data integration** — Data consists of datasets that follow a consistent content and structure over long periods of time and large numbers of records that effectively report status changes for the connected device or application, or continuously update records with new values. Stream data integration includes the ability to incorporate event models, inferred row-to-row integrity, and variations of either those models or the inferred integrity with alternative outcomes. These outcomes may or may not be aggregated/parsed into separate event streams from the same continuous stream. The logic for this approach is embedded in the data stream processing code.

- **Range of connectivity/adapter support (sources and targets)** — Native access to relational and nonrelational DBMS products, plus access to nonrelational legacy data structures, flat files, XML and message queues, cloud-based data asset types (including data of SaaS applications and cloud data stores) and streaming data.
- **Mode of connectivity/adapter support for interaction** — Support for bulk/batch, change data capture (CDC), and real-time/near-real-time connectivity/interaction (in a continuous or streaming data environment, for example).
- **Data transformation support** — Included and packaged capabilities to support basic (e.g., data type conversions, string manipulation, calculations), intermediate (e.g., lookup/replace, aggregations, summarizations, slowly-changing dimension management, matching) and complex (complex parsing, combining data and content sources, text mining, media mining, pattern/events in big data, other) transformations. Vendor solutions must support complex transformations such as integration with data quality and MDM tools (vendor's own or external) to access changes in data validation directly from the metadata within those solutions.

- **Metadata and data modelling support** — Automated metadata discovery (such as profiling new data sources for consistency with existing sources), lineage and impact analysis reporting, the ability to synchronize metadata across multiple instances of the tool. Also, an open metadata repository, including mechanisms for bidirectional sharing of metadata with other tools. **As a desirable advantage, some data integration tooling may provide machine-learning-enhanced metadata discovery, and internal analytics to enhance human data management and integration requirements using both passive and active metadata collection, sharing and analysis**. This would assist organizations in activating passive metadata and then utilizing active metadata to inform and even automate parts of data integration design and delivery (see Note 1 on “active metadata”).

- **User- or role-specific** variations in the development interface that are capable of supporting self-service data preparation. These mechanisms may include business-user-centric, low-code UIs, embedded templates, embedded transformations, automated schema drift resolutions, workflow automation, version modification (via internal library management or other mechanisms), and quality assurance capabilities — either via audit/monitor metadata (manual) or through embedded workflows (administrator tools).

- **Design and development support** — Graphical design/development environment and team development capabilities, such as version control and collaboration. This includes multiple versions running in disparate platforms and multiple instances of services deployments in production environments, as well as alternative or collaborating development environments. Increasingly, organizations also expect DataOps support — i.e., the ability to create a collaborative environment for developers and data engineers to work with data consumers in an iterative and agile manner. This is delivered through capabilities that support continuous collaboration and automation of data flows for repeatable and optimized data integration. Tools need to support CI/CD and continuous testing capabilities to deliver this at the very least.

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

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

- **Service enablement support** — The ability to deploy functionality as services, including manners in which functionality can be called via a data service or web services interface. Hence the ability to deliver integrated views of data via APIs and/or microservices is critical.

- The ability to **support the delivery of data integration functionality as cloud services**. This could be done through a hosted, containerized PaaS, IaaS or SaaS delivery mechanism. **The ability to perform integration across a hybrid and multicloud ecosystem is highly desired.**

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

- **Revenue** — Generate at least $30 million of their annual software revenue from data integration tools — i.e., from perpetual license with maintenance, or subscription with support (which would include payment only for data integration software). Or, the vendor must maintain at least 300 maintenance-paying/subscription-paying customers in production environments for its data integration tools.
- **Geography** — Support data integration tool customers in at least two of the following geographic regions or specific national markets: North America, South America, EMEA and Asia/Pacific.
- **Presence** — Demonstrate enough market presence, reviewed and assessed through internal Gartner search, external search engines, Gartner inquiry interest, technical press presence and activity in user groups or posts. A relative lack of market presence could be determined as a reason to exclude a product/service offering.

**Vendors that focus on narrow use cases that are too specific for broader market application are excluded.** In the past, some vendor/supplier tools were excluded because:

- They focused on **only one horizontal data subject area;** for example, the integration of customer-identifying data.
- They focused on **only a single vertical industry.**
- **They served only their own, internally managed data models and/or architectures** (this includes tools that only ingest data to a single proprietary data repository). Or, they provided data integration as an embedded functionality for use within the confinement of their own analytics/BI tools, data science platform or database offering. These vendors use their data integration tools only to ingest/integrate data into their own repository or within the confinement of their own broader tool/platform or ecosystem.
- **Vendors that only provide support to open-source platforms**/frameworks or development platforms that 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).
- **Vendors that provide adapters or drivers** to various data and analytics sources and targets, thereby indirectly supporting data integration, but these vendors do not market a stand-alone data integration tool(s).

## Evaluation Criteria

### Ability to Execute

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

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

### Product/Service

Core goods and services that compete in and/or serve the defined market, including current product and service capabilities, quality, feature sets and skills.

These goods and services can be offered natively or through OEM agreements/partnerships (as defined in the Market Definition/Description section or described below). Product strategy is expected to be in support of traditional integration needs filling current gaps, weaknesses and opportunities to capitalize on less-advanced demand trends in this market.

Given the requirement for data integration tools to support diverse environments for data, delivery models and platform-mix perspective, we also assess vendors on the degree of openness of their technology and product strategy. Some consumers are prepared to accept less-capable products from many different suppliers and assemble them together on their own. Connecting data integration activities to data quality and governance-related capabilities (such as MDM) becomes integral to supporting all use cases that can share high-quality data and lineage metadata, with runtime management and monitoring support.

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

- How well the vendor supports the range of distinguishing data integration functionalities required by the market
- How this functionality is delivered
- Support for established and emerging deployment models
- Overall usability and consumption of the tools

### Overall Viability

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

We view the likelihood of the organization continuing to offer and invest in the product, as well as the product's position in the current portfolio. Overall vendor viability is reviewed and utilized by

end-user organizations and developers in determining a supplier's capability to deliver ongoing production support. Importantly, open-source solutions are measured here by the strength of their community and the overall capability of the governing body to guide the roadmap and manage open-source projects.

The appropriateness of the vendor's financial resources, the continuity of its people and its technological consistency affect the practical success of the business unit or organization in generating business results.

### Sales Execution/Pricing

The vendor's capabilities in all presales activities and the structure that supports them.

This includes deal management, pricing and negotiation, presales support, and the overall effectiveness of the sales channel. Organizations increasingly seek "modularity" or the capability to isolate specific required functions in data integration that are then reflected in their implementation approach and cost allocation.

The focus on pricing by vertical — which allows for pricing by use case, role, and volumetric and performance metrics (all considered applicable for different market needs) — has increased in 2020. In addition, pricing by feature, deployment model, user persona and functionality is increasingly sought to allow for flexible use cases within familiar toolsets. The effectiveness of the vendor's pricing model in light of current customer demand trends and spending patterns, as well as the effectiveness of its direct and indirect sales channels, was scored as part of the evaluation.

### Market Responsiveness/Track Record

The vendor's ability to respond, change direction, be flexible and achieve competitive success as opportunities develop, competitors act, customer needs evolve, and market dynamics change.

This criterion also considers the vendor's history of responsiveness to changing market demands. Market track record is itself one measure of market responsiveness and, in this case, data integration tools are much like other infrastructure-focused solutions. Often, organizations demand data virtualization, message-oriented data movement, replication and synchronization, and stream data integration support. However, traditional bulk/batch processing is still the predominant demand. Not only do most solutions overlap; the market is now demanding a capability to deliver all forms of integration to differently skilled implementers, with everything from simple data preparation through self-service data integration to enterprise-class systems.

The degree to which the vendor has demonstrated the ability to respond successfully to market demand for data integration capabilities over an extended period, as well as how well the vendor acted on the vision of prior years, is also evaluated.

### Marketing Execution

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

Marketing execution was traditionally considered to be the positioning and declarations of a supplier, but now end-user organizations use it frequently as a gauge of how in-tune supplier roadmaps are with overall market demand. Suppliers need to be aware of emerging best practices for data management infrastructure, and if they and their customers can specifically benefit from specialized horizontal or vertical capabilities, geographically targeted approaches or partner-supported implementation practices. Providers must develop a means of converting community "chatter" and excitement to support delivery and go-to-market campaigns.

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

### Customer Experience

Products and services and/or programs that enable customers to achieve anticipated results with the products evaluated.

Specifically, this includes quality supplier interactions with buyers, technical support or account support. It may also include ancillary tools, customer support programs, the availability of user groups and service-level agreements (SLAs).

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

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

### Operations

The ability of the organization to meet goals and commitments.

Factors for this criterion include the quality of the organizational structure, skills, experiences, programs, systems and other vehicles that enable the organization to operate effectively and

efficiently. Operations are not specifically differentiating to end-user markets, but product management consistency and support/maintenance practices add to the overall customer experience 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 data services, metadata management and semantic tiers. Also, they must demonstrate that they are well-positioned to provide ongoing support for the massive bulk/batch data movement market.

**Table 1: Ability to Execute Evaluation Criteria**

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

Source: Gartner (August 2020)

## Completeness of Vision

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

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

### Market Understanding

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

Vendors that show a clear vision of their market will listen to and understand customer demands and can shape or enhance market changes with their added vision. A visionary market understanding recognizes the importance of advanced data management and integration to support both operational and analytics data use cases.

Application management and data management must both address the concept of role-based development. “Citizen” integrators will want rapid access to data without concerns about production optimization. Analytic assistance for data auditing, profiling, qualifying and conformance/alignment will be crucial. However, metadata-driven warnings will be needed, as well as template library management to support citizen integrators’ efforts. The degree to which the vendor leads the market in new directions (in terms of technologies, products, services or otherwise) is key, alongside its ability to adapt to significant market changes and disruptions.

### Marketing Strategy

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

Marketing is now experience-based and not as susceptible to presentations and collateral development from suppliers. In addition, suppliers must develop a means of converting community “chatter” and excitement into support delivery and go-to-market campaigns. Redesign and redeployment when going into broader implementations is considered suboptimal, so a flow from trial versions into pilot and then production is desired.

### Sales Strategy

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

This criterion covers the alignment of the vendor’s sales model with the ways in which customers’ preferred buying approaches will evolve over time. Organizations now expect differentiated pricing based on their use cases, as well as rapid conversion to scalable pricing models when new demands are introduced. Base pricing must include development and test environments that are minimal- or no-cost — or represent a convertible investment when going into production. Additionally, cost models that address mass production environments versus “citizen” or data-science-driven efforts require flexible pricing.

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

### Offering (Product) Strategy

An approach to product development and delivery that emphasizes market differentiation, functionality, methodology and features, as they map to current and future requirements.

Existing markets and use cases have begun to shift toward more distributed data integration needs, increasing the demand for self-healing and wizards/tutors to recognize new sources and information asset types. Product strategy vision includes the roadmap for continued support of traditional integration needs — filling current gaps and weaknesses as well as opportunities to capitalize on advanced demand trends. There is now significant increased expectation on “active” metadata understanding, conversion, utilization and analysis of this metadata (see Note 1). This active metadata is used in profiling, machine learning, evaluation of assets and comparison with existing integration upon connection. Self-correcting optimization in processes is now important and expected. Utilizing metadata to assist in user “push” recommendations for new data assets, and to create semantic knowledge graphs to assist with data fabric design that enables a more consistent (and application-neutral) semantic model for integration, is considered a differentiator.

In addition, given the requirement for data integration tools to support diverse environments for data, delivery models and platform-mix perspective, we assess vendors on the degree of openness of their technology and product strategy.

### Business Model

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

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

### Vertical/Industry Strategy

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

### Innovation

Direct, related, complementary and synergistic layouts of resources, expertise or capital for investment, consolidation, defensive or preemptive purposes. The current innovation demands in the market are centered on managing location-agnostic capability in data integration — i.e., the ability to not have to move or replicate data necessarily but to connect to data in-place when feasible and take the processing to the data (rather than vice versa) to execute integration.

Integration should run on-premises and in the cloud, and switch between them. As data becomes highly distributed, data integration activities are also required to become easily distributable to any data location, or recommend/determine when data needs to be moved for optimal processing. As data management use cases gain in importance to focus on transient data (traditionally the forte of message-oriented technologies), demand for converging data and application integration approaches is now expected.

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

**Geographic Strategy**

The vendor’s strategy for directing resources, skills and offerings to meet the specific needs of geographies outside the “home” or native geography — either directly or through partners, channels and subsidiaries — as is appropriate for that geography and market.

User organizations are seeking local support with differing levels of confidence in the various approaches possible (i.e., VARs, resellers, channel partners, OEM offerings and distributors). They are also seeking continuity of support across regions. Data tracing will become a key requirement in the geographic distribution of data. Development platforms must include the ability to monitor where data originates with jurisdictional cognizance, and where it is eventually delivered. Violating national laws through data movement must be addressed, and policy-level controls are expected to safeguard the citizen developer and the cloud deployment.

The vendor’s strategy for expanding into markets beyond its home region/country and its approach to achieving global presence (e.g., direct local presence and use of resellers/distributors) are crucial for capitalizing on global demands for data integration capabilities and expertise.

Table 2: Completeness of Vision Evaluation Criteria

| Evaluation Criteria ↓ | Weighting ↓ |
|---|---|
| Market Understanding | High |
| Marketing Strategy | High |

| Evaluation Criteria ↓ | Weighting ↓ |
|---|---|
| Sales Strategy | Medium |
| Offering (Product) Strategy | High |
| Business Model | Medium |
| Vertical/Industry Strategy | Low |
| Innovation | High |
| Geographic Strategy | Low |

Source: Gartner (August 2020)

## Quadrant Descriptions

### Leaders

Leaders in the data integration tool market are front-runners in their capability to support the combination of these different data delivery styles (e.g., the ability to combine and switch between ETL, replication and virtualization based on their use-case demands). Over 50% of organizations now expect their data integration tools to deliver this integration style interoperability. Leaders exhibit significant market mind share, and resources skilled in their tools are readily available. These vendors recognize the need for new and emerging market demands — often providing new functional capabilities in their products ahead of demand — by identifying new types of business problems to which data integration tools can bring significant value. Examples of deployments that span multiple projects and types of use cases are common among Leaders' customers. Leaders have an established market presence, significant size and a multinational presence — either directly or through a parent company.

In 2020, Leaders in this market have started delivering on the data fabric promise — i.e., their ability to balance collecting data with connecting to data. They automate the process of collecting all types of metadata (not just passive) and then represent the metadata (and sometimes data) in a graph (to preserve context). This is then followed by improving the data modeling process by enriching the models with agreed upon semantics. Finally, these vendors embed AI/ML toolkits, which utilize active metadata (as input) to start automating (or at the very least informing) various aspects of data integration design and delivery. Most vendors in the Leaders quadrant provide all the necessary capabilities to deliver the data fabric, although some might require significant customization.

Leaders have been advancing their metadata capabilities, with some introducing highly dynamic optimization and advanced design assistance functions. They have been extending their

capabilities to allow for ML over this active metadata, to assist developers with various degrees of support and automation in integration design and implementation. Leaders are adept at providing tools that can support both hybrid integration and multicloud integration options, bridging the data silos that exist across on-premises and multicloud ecosystems. Leaders allow organizations to remain independent in data integration as they look to deploy workloads across multiple CSPs and allow organizations to effectively provision cloud ecosystems.

Leaders are strong in establishing their data integration tools as enterprise standard in at least one primary use case, with the capabilities to deliver in multiple use cases. Their data integration tools become a critical component of the modern data management infrastructure. Their tools support both traditional and new data integration patterns in order to capitalize on market demand.

### Challengers

In 2020, there was a shift in market realization that data delivery styles other than bulk/batch (such as replication, streaming or data virtualization) are no longer differentiating but more a "must have." In line with this market shift, Challengers have been making significant strides in delivering these capabilities within a broader metadata-driven data integration toolset. And even as Challengers, these vendors are expected to have good, executable technology covering these trends (a facet previously associated with Leaders only).

In 2020, the Challengers listed in this research constitute vendors that exhibit a strong understanding of the current data integration market demand and exhibit both the credibility and viability to deliver on that demand. Some Challengers are extremely mature on specific core capabilities, which enables them to deliver targeted use cases faster and with a better overall TCO than other vendors. These vendors have developed best practices for leveraging their strongest product capability in new delivery models. For example, the ability to productize and market data replication as a key strength for targeted use cases such as cloud data migration; or the ability to productize and market data virtualization for faster turnaround time to analytics.

Challengers generally have substantial customer bases and established presences. They exhibit strong market presence, although implementations may be of a single-project nature or reflect multiple projects of a single type. For example, predominantly data virtualization, data replication or ETL-oriented use cases; or use cases specific to IoT data or geospatial data type. Gartner realizes that, while the market vision is for the vendors to deliver all data delivery styles across a multicloud/hybrid environment and support all popular data sources/targets, many customers could have specialized and targeted demands for their most urgent or upcoming projects. Therefore, we also recognize vendors in the Challengers quadrant that (if needed) can scale to support most data integration use cases, but can also customize their offerings for specific use cases, data types, data sources/targets, execution environments or specific CSPs in demand.

Overall, the market is pushing Challengers to embrace the market vision of a data fabric, which ensures that they utilize metadata and perform ML over this metadata to deliver solutions that can automate various data integration tasks. These tasks include automated profiling, repetitive transformations, data preparation, performance optimization, query optimization, scaling, tuning,

and movement of workloads to data stores and engines best suited for processing. Overall, this move toward enabling data fabric architectures will be a key area in 2020 that will determine which Challengers can move into the Leaders quadrant next year.

### Visionaries

Visionaries demonstrate a strong understanding of emerging technology and business trends or focus on a specific market need that is far outside of common practices, while also possessing capabilities that are expected to grow in demand. In 2020, the Visionaries in this Magic Quadrant have focused early on alternative go-to-market strategies or specific capabilities to capitalize on their capacity to leverage either:

- Augmented data integration through the data fabric design
- Serverless integration tooling that supports a multicloud and hybrid cloud integration architecture, *or*
- The growth in demand for connectors and delivery of various data integration functionalities as loosely coupled API/microservices
- Seamless orchestration of various data integration components and delivery styles through DataOps techniques

In addition, a significant driver of vision for the market this year has been the ability of tools to connect to and analyze all forms of metadata — both passive and, increasingly, active metadata (see Note 1). With this, tools can provide key statistics to developers and citizen integrators that aid with integration design and, increasingly, integration automation. The Visionaries are doing all these things. Additionally, the development of specific solution designs that mix repeatable delivery templates with appropriate custom-deployed options represents the potential to address specific new markets demanding data integration in highly specific vertical offerings. Visionaries are leading the push toward the utilization of graphs, semantics, knowledge graphs and AI/ML for significant automation in both data integration design, delivery and maintenance.Visionaries sometimes lack market mind share or credibility beyond their customer base or single application domain. They may also fail to provide a comprehensive set of product capabilities — including those that focus on a single data integration style — and simply import, export or leverage that primary data integration style to deliver customer use cases. They may be new entrants lacking the installed base and global presence of larger vendors. Finally, Visionaries may be large, established players in related markets that have only recently placed an emphasis on data integration tools.

### Niche Players

With the market now matured, Niche Players generally don’t exhibit gaps in primary market functionality or features. Instead, they are simply challenged in increasing their execution or have not identified a specific market approach that expands use cases for their technology. This

means that almost every Niche Player will be able to deliver against standard market expectations both in functionality and cost-price options.

Niche Players do not appear very frequently in competitive situations for comprehensive and/or enterprise-class data integration deployments. Many have very strong offerings for a specific range of data integration problems — for example, a set of technical environments, data delivery styles (batch, replication, streaming or virtual), application domains or use-case scenarios — and deliver substantial value for their customers in the associated segment. Niche Players now exhibit particular advantages in pricing, in their small footprint and even in vertical or horizontal solutions. This makes them ideal candidates to be a best-fit solution that complements other technology in the data management infrastructure of an organization.

Niche Players are known for solving one part of the data integration problem well through a targeted solution. This could be data migration to the cloud, or real-time data replication, or even location intelligence through geospatial data integration, or even IoT data integration support. Importantly, Niche Players in this market have demonstrated their capability to outperform dozens of tool and solution offerings that were considered and eventually excluded from this Magic Quadrant. Finally, Niche Players may be lacking maturity on certain features that display market vision. These features might include multicloud/hybrid support or support for data fabric architectures, or even the ability to inform and automate data integration design and delivery through active metadata support and analysis.

Clients should note that more than 80% of all end-user organizations still seek bulk/batch processing (even in hybrid and intercloud scenarios). This means that a highly efficient but batch-oriented data integration tool vendor could exhibit high-level execution capabilities without ever crossing to the right-hand side of the Magic Quadrant (i.e., to the Visionaries or Leaders quadrants). Niche Players all exhibit batch capabilities, from back-office and operations data through to massive volumes of sensor or IoT data.

## Context

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

Data and analytics leaders must navigate a market brimming with products that claim to solve a range of data integration problem types. However, not all vendor solutions have experience in — nor evenly provide — all of the relevant capabilities needed across our key data integration use cases (see the companion “Critical Capabilities for Data Integration Tools”). Some vendors focus heavily on providing solutions focused on just one data delivery style such as bulk/batch (through ETL), data replication (through CDC), messaging through (APIs), or virtual (through data

virtualization). But they may place less emphasis on the important capability of interoperating and combining these different data delivery styles for accomplishing key use cases.

In 2020, traditional data integration has shifted from being solely focused on bulk/batch dominance for delivery in the market. More than 60% of reference customers said they make significant use of data replication, 30% also utilize data virtualization, close to 40% reported using data integration tools for data synchronization and 22% reported using tools for stream data delivery. The telling metric for this year is that more than 50% of the surveyed organizations stated that they need their data integration tool to support "seamless combination of and switching between these different data delivery styles." This is where data and analytics leaders evaluating data integration tools must lay a strong focus. They must give strong preference to tools that can orchestrate multiple modes of data delivery and not be skewed toward supporting just one style (unless their use case demands so).

Some organizations have determined that basic functions are adequate for them and are, therefore, seeking tools with focused and targeted capabilities. As a result, they are interested in evaluating and procuring tools that are specialists in one data delivery style (e.g., data replication, self-service data preparation or data virtualization). Also, some organizations prefer tools that can support one use case (such as cloud data ingestion and migration), one data type (such IoT data integration) or one scenario (such as location intelligence through geospatial data integration focus). Such organizations can confidently start with the vendors in the Niche Players quadrant and then ensure that the shortlisted vendors can also support their upcoming use cases for the medium and long term (using the companion "Critical Capabilities for Data Integration Tools").

Organizations that seek tools that are generalists (or best of breed in data integration) and can support multiple use cases though a combination of different data integration styles can evaluate the vendors mentioned in the Challengers and Leaders Quadrants.

In addition, vendors in the Leaders quadrant are focused on new demands for automation in various aspects of data integration. These include design, ingestion, schema mapping, schema drift detection and corrections, next-best transforms, automated lineage and impact analysis, and infrastructure management and orchestration. These capabilities for augmented data integration demand a new data integration design — one that supports a balance of connect and collect data integration strategies. This design utilizes the collection and analysis of all forms of metadata (not just technical metadata) to provide insights that enable automated data integration delivery ("Data Fabrics Add Augmented Intelligence to Modernize Your Data Integration"). Data fabrics are defining the direction of this market, and while Visionaries have a strong product to meet the demand, Leaders have been able to productize this architecture and deliver it at scale due to their superior reach and presence.

Active metadata-enabled data integration is a significant driver of market vision this year. Metadata as a byproduct of the design and operations management of a data integration platform is a minimum requirement of data integration tools in 2020. Platforms and solutions are now expected to provide continuous feedback regarding the profiles, quality, location, performance optimization, lineage, use cases, access points, context, frequency of access and

content analysis of integrated data assets. As far as architects and solution designers are concerned, this feedback is long overdue. It is expected that graph analytics powered by every conceivable type of metadata (both passive and active) will provide the necessary dynamic data fabric designs for introducing ML capabilities into data integration platforms (see "How to Activate Metadata to Enable a Composable Data Fabric"). This capability for active metadata-based integration has been weighted very highly to define the vision of the market this year by Gartner analysts.

Gartner sees that the urgent need to acquire and integrate data across multiple CSPs, typically for hybrid cloud and intercloud integration, is becoming crucial to many data integration use cases. The COVID-19 pandemic will accelerate the movement to the cloud and data integration tools are leading the way to support this shift. However, integration tool vendors that either support the integration of data for just one cloud provider (typically their own cloud) or for integration scenarios utilizing their own databases or applications will fall behind on customer demand due to valid lock-in and CSP independence concerns.

An interesting data point from the reference customer survey was that close to 64% of surveyed organizations utilize data integration tools for their data replication demands, up from 58% in 2019. This increase is because organizations are looking to utilize their data integration tools to ingest or replicate the data from their operational DBMSs to cloud data warehouses supported by dbPaaS. This has been a significant driver of growth for many data integration providers (such as Fivetran, HVR, Matillion and Qlik). These providers have formed significant partnerships with CSPs like AWS and Microsoft Azure, along with popular cloud data warehouse vendors such as Snowflake and Teradata, to deliver integrated data from on-premises data stores and applications to cloud data warehouses and lakes for analytics. They are doing this through forward engineering, often as ready for consumption as integrated data with schema assigned for analytics and data science use cases.

In a continuation from 2019, organizations are seeking solutions that facilitate role-based data integration. This includes the capability to promote or manage the workflow of converting individually developed processes into enterprise-capable ones (see "Market Guide for Data Preparation Tools"). This year, analysts have evaluated data preparation as a key capability for execution metrics.

Finally, a mix of data integration approaches has remained crucial, spanning physical delivery to virtualized delivery, and bulk/batch movements to event-driven granular data propagation. In particular, when data is being constantly produced in massive quantities and is always in motion and constantly changing (e.g., IoT platforms and data lakes), attempts to collect all this data are neither practical nor viable. This is driving an increase in demand for connection to data, not just the collection of it (see "Market Guide for Data Virtualization"). In 2020, data virtualization has again been a key criterion for measuring the vision score of evaluated vendors. But the ability of the tool to deliver data virtualization in combination with other data integration styles including batch, streaming and messaging (which is necessary for optimal performance optimization) has been weighted as an execution criterion.

## Market Overview

The data integration tools market continues to push toward distributed and dynamic data management. This push is inherent in the modern data fabric architecture (see "Data Fabrics Add Augmented Intelligence to Modernize Your Data Integration"). The market has realized that those data integration tools that do not balance "collect"- with "connect"-based data management architecture strategies will always result in data silos and/or poorly integrated infrastructures. In 2020, we have witnessed a variety of trends dictating the data integration tools market. Some of these trends and directions were inevitable, while some have emerged as urgent due to the COVID-19 pandemic. Moving forward, organizations will need to monitor and exploit the major trends that are affecting enterprise requirements and vendor offerings in the data integration tool market. We highlight some of these below.

**Market leaders lose ground to smaller vendors.**

The top five vendors in this market (based on overall market share) had a collective market share of 70% in 2017. This dropped to 60% in 2019. A similar trend can be seen when analyzing the top three or even top 10 vendors. Even within the top 10, the three vendors (Talend, Qlik and Denodo) that have contributed most to market growth this year fall outside the top five. Most of the smaller vendors experiencing high growth share a common theme — they focus on leadership in specific data integration style such as data virtualization or data replication; or they focus on data integration delivered as a native cloud service; or both. Larger and established vendors will need to find the balance between all-encompassing platform solutions (through concepts such as data fabric) and easily accessible point solutions to keep pace (see "Market Share Analysis: Data Integration Tools, Worldwide, 2019").

**Market and tools consolidation are in full swing, driven by convergence in capabilities.**

Highly focused point solutions that deliver a specific data delivery style for targeted use cases remain popular (such as data ingestion and replication tooling supporting popular cloud data migration use cases). That said, the market is beginning to consolidate to support a combination and interoperability of multiple data delivery styles ("Modernize Your Data Integration Architecture for Digital Business by Combining Data Delivery Styles"). This consolidation is forcing vendors to rethink product strategy and come up with tools that can combine and deliver multiple data delivery styles for broader data and analytics use cases (such as a combination of ETL and data virtualization to support exploratory analytics and enterprise reporting).

Another vector for consolidation this year has been the surge in demand for convergence in application integration and data integration tooling. In fact, 66% of respondents to our Magic Quadrant reference survey stated that they are utilizing the embedded application integration capabilities within their data integration tools. This convergence is being led by iPaaS tools that combine data integration, application integration and API delivery/management capabilities, and deliver them through cloud delivery models (see "Aligning Application and Data Integration Delivers Synergy in Disciplines and Technologies").

Gartner is also seeing organizations increasingly demand multiple data management capabilities in a single, cohesive ecosystem (see "Modern Data and Analytics Requirements Demand a Convergence of Data Management Capabilities"). The primary reason for this is that organizations don't want to buy technology or tools alone; they want solutions for use cases and business problems, and often those use cases require data integration with other key data management capabilities. These capabilities include data ingestion, data preparation, data catalogs, and data and analytics governance (including capabilities for data quality). This convergence is due to the growing adoption of the data fabric as a design concept (see "Data Fabrics Add Augmented Intelligence to Modernize Your Data Integration"). Most vendors struggle to deliver all the capabilities to implement a true data fabric today in an integrated form. But this convergence (either through organic in-house product development or through partnerships and M&As) is helping the data fabric design concept gain more mainstream adoption ("Market Trends: The Impact of 3 Convergence Points on the Data Integration Tools Market").

**Independent data integration to prevent application/CSP/cloud ecosystem/database lock-in.**

Another major development this year has been the focus on independent and general-purpose data integration tools that don't necessitate or force the movement and persistence of data into a specific vendor repository or cloud ecosystem. This is more important than ever because embedded data integration capabilities delivered by vendors as part of a broader application (e.g., analytics and BI or CRM tool) or database, or even CSP-specific data integration solutions, might make it easy for organizations to ingest data into one database, application or CSP ecosystem. However, these same embedded integration capabilities do very little to allow organizations to integrate data across different data stores, applications or multicloud/hybrid environments. This could lead to potential vendor lock-in challenges and data silos, resulting in the inability of organizations to reuse integrated data for general purpose use cases. This is something that organizations can ill-afford due to their cost optimization initiatives, especially during the COVID-19 era.

**Cloud data integration and migration becomes a key focus area for end users and vendor growth.**

Throughout the COVID-19 pandemic, Gartner has witnessed a huge surge in data-integration-related inquiries, primarily focused on data migration to the cloud. This is developing into an extremely popular use case for data integration tool providers. Some vendors have responded by developing targeted offerings that make it easier for business teams to use their low-code UIs for data ingestion to a cloud data store, followed by the ability to catalog the data, perform self-service data preparation and finally to consume integrated and modelled data in their business applications. Other vendors have repackaged existing offerings and partnered with popular cloud data stores — such as Amazon Web Services (AWS), Google Cloud Platform, Microsoft Azure, and Snowflake — that make it easy for organizations to ingest, integrate and model data on the cloud.

**Data fabric is critical and driven by the end-user push toward augmented data integration.**

Another huge Completeness of Vision criterion that the market is demanding is augmented data integration design and delivery. Reference customer survey respondents have frequently deemed this requirement as one of prime importance. This year, 49% of respondents stated that they would like to automate data integration tasks; 37% stated that they would like to automate the data preparation drudgery; and nearly 30% stated that they would like to automate manual data ingestion (see the Evidence section for survey details). These are telling metrics that clearly showcase the focus on augmented data integration design and delivery.

The COVID-19 pandemic has only fast-tracked this strategic direction of the market. Data and analytics leaders are realizing that they cannot keep investing in manual data integration and need automation support. In fact, in our Magic Quadrant reference survey, we noticed that data integration teams (in terms of individual members) are constantly contracting — the median number of individuals in teams for this year is less than 15. And while team sizes are reducing, the amount of data and, hence, data integration requirements are growing exponentially. This gap between demand and supply is pointing toward an urgent focus on automation and augmentation.

Augmented data integration demands a renewed focus on the data fabric architecture design. A data fabric is an architecture pattern that informs and automates the design, integration and deployment of data objects regardless of deployment platforms or architectural approaches. Data fabrics utilize continuous analytics and AI/ML over all metadata assets (such as location metadata, data quality, frequency of access and lineage metadata, not just technical metadata) to provide actionable insights and recommendations on data management and integration design and deployment patterns. This approach results in faster, informed and, in some cases, completely automated data access and sharing. The data fabric requires various data management capabilities to be combined and used together. These include augmented data catalogs, semantic data enrichment, utilization of graphs for modeling and integration, and finally an insights layer that uses AI/ML toolkits over metadata graphs to provide actionable recommendations and automated data delivery. In 2020, some vendors have been able to combine all of these capabilities into productized solutions, which signifies leadership. Others are going in this direction through partnerships, M&A activity, product enhancements and, more frequently, a combination of all these.

**Cost optimization through modularity, tightly integrated (but loosely coupled) capability delivery and flexible deployment and licensing options are critical.**

COVID-19 has brought back focus on cost optimization in data integration. The market is now expected to deliver modular solutions through loosely coupled architectures, where specific capabilities that are urgently needed by organizations (such as data catalogs or data virtualization) can be delivered and consumed as needed. Vendors are therefore now expected to break out key capabilities from broader platforms and deliver them through loosely coupled (but highly integrated) architectures and deployment platforms (such as delivering key capabilities as microservices or through containers). The market is also favoring data integration tools that can be procured and accessed through favorable pricing models based on pay-as-you-go, freemium

and serverless metered pricing models. (These new pricing models allows customers to try functionality before buying the full platform and to pay for actual usage in popular cloud environments). This should be in addition to, and not a replacement for, performance-based pricing metrics, and a part of the market values price predictability as well.

**Business teams demand self-service data preparation for agility and data engineering teams need data preparation for improving productivity.**

Organizations need their data integration tools to provide capabilities for self-service data preparation — that is, nontechnical users such as business analytics or citizen integrators need to be able to integrate data with minimal IT support for their specific business requirements (see "Market Guide for Data Preparation Tools"). Data integration functionality provided in a "sandbox" to support analytics is of growing interest. This approach enables data to be delivered and manipulated in a physical or virtual manner for ingestion, regardless of where it resides. It also encourages experimentation with, and the building of, new models with which to use data of interest. Implementations need to support multiple types of user experience via tool interfaces that appeal not only to technical practitioners but also to people in business-facing roles, such as business analysts and end users. Offerings that promote collaboration between business and IT participants (including data engineers) are becoming important as organizations seek adaptive approaches to achieving data integration capabilities.

Data engineering teams struggling under the weight of an increasing number of data pipelines need urgent support. They also request for their data integration tools to embed modern data preparation capabilities that allow them to deliver tactical (and often repeatable) data integration tasks through automation of repeatable tasks. This allows them to focus on more strategic requirements in data integration. Finally, operationalization of self-service jobs by enabling bidirectional metadata exchange between the self-service tools (or modules) and ETL tools remains a crucial requirement.

**Organizations demand continuous and repeatable delivery of data integration through data engineering platforms and DataOps practices.**

As data management is increasingly treated as software development and roles such as the data product manager emerge, more focus is being placed on means of delivering data integration in an agile, automated and repeatable fashion. This relies on improving how data producers (data engineers, ETL developers) and data consumers (business analysts, information stewards) interact throughout the process (see "Data Engineering Is Critical to Driving Data and Analytics Success" and "Introducing DataOps Into Your Data Management Discipline").

## Acronym Key and Glossary Terms

| | |
|---|---|
| AI | artificial intelligence |
| B2B | business to business |

| | |
|---|---|
| B2C | business to consumer |
| BI | business intelligence |
| CDC | change data capture |
| CRM | customer relationship management |
| CSP | cloud service provider |
| DBMS | database management system |
| ELT | extract, load, transform |
| ETL | extract, transform, load |
| iPaaS | integration platform as a service |
| IoT | Internet of Things |
| ISV | independent software vendor |
| MDM | master data management |
| M&A | merger and acquisition |
| ML | machine learning |
| NLG | natural language generation |
| NLP | natural language processing |
| OEM | original equipment manufacturer |
| PaaS | platform as a service |
| PII | personally identifiable information |
| TCO | total cost of ownership |

| | |
|---|---|
| VAR | value-added reseller |

## Evidence

The analysis in this Magic Quadrant research is based on information from several sources, including:

- An RFI process that engaged vendors in this market. It elicited extensive data on functional capabilities, customer base demographics, financial status, pricing and other quantitative attributes.
- Interactive briefings in which vendors provided Gartner with updates on their strategy, market positioning, recent key developments and product roadmap.
- A web-based survey of reference customers identified by each vendor. This captured data on usage patterns, levels of satisfaction with major product functionality categories, various nontechnology-related vendor attributes (such as pricing, product support and overall service delivery), and more. In total, 127 organizations associated with 20 vendors across all major regions provided input on their experiences with vendors and their tools.
- Feedback about tools and vendors captured during conversations with users of Gartner's client inquiry service.
- Market share and revenue growth estimates developed by Gartner's technology and service peer feedback from Gartner Peer Insights, comprising peer-driven ratings and reviews for enterprise IT solutions and services covering over 300 technology markets and 3,000 vendors.

## Note 1
## Passive Versus Active Metadata and How to Activate Metadata

Passive metadata is metadata that is static in nature, usually emerges at design time and often requires human or manual updates. Passive metadata most often consists of simple documentation or design time technical metadata. It consists primarily of documentation, ranging from fixed schema of sources and/or targets all the way through to business definitions acquired in a glossary and maintained as a formal data dictionary.

Organizations now need continuous access, analysis and feedback on all metadata parameters (not just technical but also operational, usage and social). These include frequency of access, data lineage, performance optimization, context and data quality (based on feedback from supporting data quality/data governance/information stewardship solutions). It is expected that having graph analytics on every conceivable type of metadata will provide the necessary information for introducing ML capabilities into various data management tools (including data integration, data quality, data preparation and even DBMS). "Activating" passive metadata involves:

- Collecting all forms of metadata and illustrating them (along with their intricate relationships) in a graph
- Performing analytics on this metadata
- Using the results of this analysis as inputs for ML algorithms to assist with informing and automating data management activities

The resultant output is called active metadata.

## Note 2
## Detailed Components of the Evaluation Conditions

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

- **Connectivity/adapter capabilities (data source and target support).** The ability to interact with a range of different types of data structure, including:
  - Relational databases
  - Legacy and nonrelational databases
  - Various file formats
  - XML
  - Packaged applications such as those for CRM and supply chain management
  - SaaS and cloud-based applications and sources
  - Industry-standard message formats, such as electronic data interchange (EDI), Health Level Seven International (HL7) and Society for Worldwide Interbank Financial Telecommunication (SWIFT)
  - Parallel distributed processing environments, such as Hadoop Distributed File System (HDFS); other nonrelational-type repositories, such as graph, table-style, document store and key-value DBMSs
  - Message queues, including those provided by application integration middleware products and standards-based products (such as Java Message Service)
  - Data types of a less-structured nature, such as those associated with social media, web clickstreams, email, websites, office productivity tools and content
  - Emergent sources, such as data on in-memory repositories, mobile platforms and spatial applications

- Data integration tools must support **different modes of interaction** with this range of data structure types, including:
  - Bulk/batch acquisition and delivery
  - Granular trickle-feed acquisition and delivery
  - Change data capture (CDC) — the ability to identify and extract modified data
  - Event-based acquisition (time-based, data-value-based or links to application integration tools to interact with message request/reply, publish-subscribe and routing)

- **Data delivery capabilities.** The ability to provide data to consuming applications, processes and databases in a variety of modes, including:
  - Physical **bulk/batch data movement** between data repositories, such as processes for ETL or for ELT
  - **Data virtualization**
  - **Message-oriented encapsulation and movement of data** (via linkage with application integration tool capability)
  - **Data synchronization** when distributed datasets must resolve data collisions resulting from distinct changes in disparate copies of data to retain data consistency
  - **Replication of data** between homogeneous or heterogeneous DBMSs and schemas
  - **Stream data integration** — Stream data integration provides the high throughput and low latency that is required to handle event streams with up to hundreds of thousands of events per second, into the millions in some cases. In some scenarios, it is used for loading data from queuing services, such as Kafka or message-oriented middleware, into in-memory DBMSs, NoSQL persistent DBMSs, or AWS S3 or HDFS file systems. In other (offline) scenarios, it moves batches of data between one DBMS or file system and another. This data delivery style also supports data integration services such as filtering, transformation and enrichment as they move the data.
  - **Data services orchestration** — The ability to deploy any of the other data integration styles, but with the specific capability to interoperate with application services (logic flows, interfaces, end-user interfaces, and so on). Also, the ability to pass instructions to, and receive instructions from, those other services on the bus. Data services bus includes auditing to assist in service bus management, either internally or by-passing audit metadata to another participating service on the bus.

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

- Basic transformations, such as data-type conversions, string manipulations and simple calculations
- Transformations of intermediate complexity, such as look-up and replace operations, aggregations, summarizations, integrated time series, deterministic matching and the management of slowly changing dimensions
- Complex transformations, such as sophisticated parsing operations on free-form text, rich media and patterns/events in big data
- In addition, the tools must provide the following facilities for developing custom transformations and extending packaged transformations

- **Metadata and data modeling support.** As the increasingly important heart of data integration capabilities, metadata management and data modeling requirements include:
  - Automated discovery and acquisition of metadata from data sources, applications and other tools
  - Discernment of relationships between data models and business process models
  - Data model creation and maintenance
  - Physical-to-logical model mapping and rationalization
  - Ability to define model-to-model relationships via graphical attribute-level mapping
  - Lineage and impact analysis reporting, in graphical and tabular formats
  - An open metadata repository, with the ability to share metadata bidirectionally with other tools
  - Automated synchronization of metadata across multiple instances of the tools
  - Ability to extend the metadata repository with customer-defined metadata attributes and relationships
  - Documentation of project/program delivery definitions and design principles in support of requirements' definition activities
  - A business analyst/end-user interface to view and work with metadata
  - Embedded machine-learning-enhanced metadata discovery, and internal analytics capability to augment human data management and integration requirements using both passive and active metadata collection, sharing and analysis. This would assist organizations in activating passive metadata and then utilizing active metadata to inform and even automate parts of data integration design and delivery (see "Data Fabrics Add Augmented Intelligence to Modernize Your Data Integration")

- **Design and development environment capabilities.** Facilities for enabling the specification and construction of data integration processes, including:
  - Graphical representation of repository objects, data models and data flows
  - Management of the development process workflow, addressing requirements such as approvals and promotions
  - Granular, role-based and developer-based security
  - Team-based development capabilities, such as version control and collaboration
  - Functionality to support reuse across developers and projects, and to facilitate the identification of redundancies
  - A common or shared user interface for design and development (of diverse data delivery styles, data integration and data quality operations, cloud and on-premises environments, and so on)
  - A business analyst/end-user interface to specify and manage mapping and transformation logic through the use of end-user functionality for data integration/preparation
  - Support for testing and debugging

- Data governance support capabilities (via interoperation with data quality, profiling and mining capabilities with the vendor's or a third party's tools). Mechanisms to work with related capabilities to help with the understanding and assurance of data quality over time, including interoperability with:
  - Data profiling tools (profiling and monitoring the conditions of data quality)
  - Data mining tools (relationship discovery)
  - Data quality tools (supporting data quality improvements)
  - Information stewardship solutions
  - In-line scoring and evaluation of data moving through the processes

- **Deployment options and runtime platform capabilities.** Breadth of support for the hardware and operating systems on which data integration processes may be deployed, and the choices of delivery model — specifically:
  - Mainframe environments, such as IBM z/OS and z/Linux
  - UNIX-based environments
  - Windows environments

- Linux environments
- On-premises installation and deployment of software
- Hosted off-premises software deployment (dedicated, single-tenant implementation)
- iPaaS delivery options — i.e., data integration services consumed by the customer completely “as a service” (the vendor provides cloud infrastructure); the customer does not install or administer the software
- Cloud deployment support (requires organizations to deploy software in a cloud infrastructure); importantly, the ability to design once but deploy across multiple or even hybrid/mixed environments, on-premises, in the cloud or both
- In-memory computing environment
- Server virtualization (support for shared, virtualized implementations)
- Parallel distributed processing, such as Apache Hadoop, MapReduce, or leveraging Apache Spark or Hadoop YARN (Yet Another Resource Negotiator)

- **Operations and administration capabilities.** Facilities for enabling adequate ongoing support, management, monitoring and control of the data integration processes implemented by the tools, such as:
  - Error-handling functionality, both predefined and customizable
  - Monitoring and control of runtime processes, both via functionality in the tools and through interoperability with other IT operations technologies
  - Collection of runtime statistics to determine use and efficiency, as well as an application-style interface for visualization and evaluation
  - Security controls, for both data in-flight and administrator processes
  - A runtime architecture that ensures performance and scalability

- **Architecture and integration capabilities.** The degree of commonality, consistency and interoperability between the various components of the data integration toolset, including:
  - A minimal number of products (ideally one) supporting all data delivery modes
  - Common metadata (a single repository) and/or the ability to share metadata across all components and data delivery modes
  - A common design environment to support all data delivery modes

- The ability to switch seamlessly and transparently between delivery modes (bulk/batch versus granular real-time versus federation) with minimal rework
- Interoperability with other integration tools and applications, via certified interfaces, robust APIs and links to messaging support
- Efficient support for all data delivery modes, regardless of runtime architecture type (centralized server engine versus distributed runtime)
- The Ability to Execute data integration in cloud and on-premises environments, as appropriate, where developed artifacts can be interchanged, reused and deployed across both environments with minimal rework

- **Service enablement capabilities.** As acceptance of data service concepts continues to grow, so data integration tools must exhibit service-oriented characteristics and provide support for SOA, such as:
  - The ability to deploy all aspects of runtime functionality as data services (for example, deployed functionality can be called via a web services interface)
  - Management of publication and testing of data services
  - Interaction with service repositories and registries
  - Service enablement of development and administration environments, so that external tools and applications can dynamically modify and control the runtime behavior of the tools

## Evaluation Criteria Definitions

### Ability to Execute

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

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

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

**Market Responsiveness/Record:** Ability to respond, change direction, be flexible and achieve competitive success as opportunities develop, competitors act, customer needs evolve and

market dynamics change. This criterion also considers the vendor's history of responsiveness.

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

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

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

## Completeness of Vision

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

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

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

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

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

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

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

**Geographic Strategy:** The vendor's strategy to direct resources, skills and offerings to meet the specific needs of geographies outside the "home" or native geography, either directly or through

partners, channels and subsidiaries as appropriate for that geography and market.