Magic Quadrant pour les solutions de qualité des données

1 novembre 2022 - ID G00759931 - 66

Par Ankush Jain, Melody Chien

L'augmentation de la qualité des données, stimulée par l'IA/ML, les métadonnées, la convergence et les intégrations entre les offres de gestion des données pour une automatisation utile, continue d'être le principal moteur de ce marché. Cette recherche permet aux responsables des données et de l'analyse de comprendre le paysage des fournisseurs et de faire le meilleur choix.

Hypothèses de planification stratégique

- D'ici 2024, 90 % des décisions d'achat de technologies de qualité des données seront axées sur la facilité d'utilisation, l'automatisation, l'efficacité opérationnelle et l'interopérabilité en tant que facteurs de décision critiques.

- D'ici 2025, 80 % des fournisseurs de qualité de données grand public étendront leurs capacités de produits pour fournir de meilleures informations sur les données en découvrant des modèles, des tendances et des relations de données, en plus de la résolution des erreurs.

- D'ici 2026, 20 % des grandes entreprises utiliseront une plate-forme unique de gouvernance des données et de l'analyse pour unifier et automatiser les programmes de gouvernance discrets.

Définition/description du marché

Les solutions dequalité D ata (DQ) sont l'ensemble des processus et des technologies permettant d'identifier, de comprendre, de prévenir, de faire remonter et de corriger les problèmes dans les données qui soutiennent une prise de décision et une gouvernance efficaces dans tous les processus métier. Les solutions packagées spécialement conçues disponibles sur le marché comprennent une gamme de fonctions critiques, telles que le profilage des données, l'analyse, la normalisation, le nettoyage, la mise en correspondance, la surveillance, la création de règles et l'analyse, ainsi que le flux de travail intégré, les bases de connaissances et la collaboration.

Ces solutions packagées aident à mettre en œuvre et à soutenir la pratique de l'assurance qualité d ata, principalement intégrée dans le cadre d'une stratégie plus large de données et d'analyse (D & A). Les dirigeants de D&A doivent continuellement améliorer la qualité des données pour

assurer l'utilisation et l'applicabilité des données pour divers cas d'utilisation existants et à venir, tels que:

- Analyse et science des données

- Ingénierie des données

- Gouvernance D&A

- Gestion des données de référence

- Qualité des données opérationnelles/transactionnelles

- Intelligence artificielle et apprentissage automatique (IA/ML)

Des descriptions détaillées de ces cas d'utilisation courants et émergents qui nécessitent divers degrés d'augmentation et d'automatisation, ainsi que d'autres fonctionnalités de base en matière de qualité des données, sont disponibles dans le document complémentaire Critical Capabilities for Data Quality Solutions.

Les leaders D&A ont plusieurs choix lorsqu'il s'agit de choisir une solution de qualité des données. Le principal moteur de ce marché, à la fois dans un passé récent et futur, est l'augmentation de la qualité des données (ADQ). ADQ est principalement motivé par les technologies liées aux métadonnées et à l'IA et est le facteur clé de l'innovation technologique en matière de qualité des données. Les fournisseurs dominent le marché avec des capacités accrues pour automatiser les processus traditionnels de qualité des données ou introduire de nouvelles façons d'améliorer les meilleures pratiques en matière de qualité des données. Chacune de ces technologies peut fonctionner indépendamment et en coopération pour créer des effets de réseau, qui peuvent ensuite être utilisés pour accroître l'automatisation et l'efficacité de la qualité des données. Par exemple, des graphiques de connaissances peuvent être créés en fonction des métadonnées capturées. Les modèles à travers le graphique peuvent ensuite être exploités pour identifier les informations dupliquées. Les données dupliquées auront des connexions partagées avec d'autres entités. Ce chevauchement des relations peut alerter les gestionnaires pour qu'ils étudient les possibilités de duplication des données ou les possibilités de transformation des données.

La qualité accrue des données est une capacité émergente qui peut améliorer les processus de qualité des données de routine et répétitifs grâce à l'analyse des métadonnées, aux graphiques de connaissances et aux techniques d'IA. Il en résulte une amélioration de la précision des données, des suggestions prescriptives et l'automatisation des flux de travail de processus pour un retour sur investissement rapide avec un coût opérationnel réduit.

Gartner's research on this market focuses on technologies and approaches delivering on the current and future needs of end users. Gartner defines the data quality solutions market as stand-alone software products that address the following data quality capabilities that organizations need in their data management solutions portfolio:

- **Connectivity:** Access and apply data quality across a wide range of data sources, including internal/external, at rest/streaming, on-premises/cloud, relational/nonrelational data sources.

- **Profiling:** The statistical analysis of diverse datasets including structured/unstructured and on-premises/cloud datasets to give business users insight into the quality of data and enable them to identify data quality issues.

- **Analytics and interactive visualization:** Interactive analytical workflow and visual output of statistical analysis to help business and IT users identify, understand and monitor data quality issues and discover patterns and trends over time, such as through reports, scorecards, dashboards and mobile devices.

- **Monitoring and detection:** Perform data quality monitoring based on preconfigured, custom-built monitoring rules, or adaptive rules, and alert violations. Automatically detect outliers, anomalies, patterns and drifts. Also provide monitoring dashboard, log files or audit trail for compliance requirements.

- **Parsing, standardizing and cleansing:** The decomposition and formatting of diverse datasets based on government, industry or local standards, business rules, knowledge bases, metadata, and machine learning (ML). Modification of data values to comply with domain restrictions, integrity constraints or other business rules.

- **Matching, linking and merging:** Match, link and merge related data entries within or across diverse datasets using a variety of traditional and new approaches such as rules, algorithms, metadata, artificial intelligence (AI) and ML.

- **Metadata and lineage:** Collect and discover metadata or import metadata from third-party tools. Build or import lineage to perform rapid root cause analysis of data quality issues and impact analysis of remediation. Apply active metadata findings, make use of metadata-based rule recommendations and associations, data discovery and cataloging, metrics view against critical data elements.

- **Multidomain support:** Address multiple data subject areas (such as various master data domains and vertical industry domains) and offer depth of packaged support (such as prebuilt data quality rules) for these subject areas.

- **Address validation/geocoding:** Support location-related data standardization and cleansing. Completion for partial data in real-time or batch process.

- **Data curation and enrichment:** Integrate externally sourced data to improve completeness and add value.

- **Rule management and data validation:** Design, create and deploy business rules for specific data values. The rules can be called within the solution or by third-party applications for data validation purposes, which can be done in batch or real-time mode.

- **Business-driven workflow and issue resolution:** The processes and user interface to manage the data quality issue resolution through the stewardship workflow. Easily identify, quarantine, assign, escalate and resolve data quality issues, facilitated by collaboration, pervasive monitoring and case management.

- **DataOps support:** Collaboration of data management practice focused on improving the communication, integration and automation of data flows between data managers and data consumers across an organization.

- **Deployment environment:** Styles of deployment, hardware, operating system and maintenance options for deploying data quality operations.

- **Architecture and integration:** Commonality, consistency, and interoperability among various components of the data quality toolset (including third-party tools) and other data management solutions or components.

- **Automation and augmentation:** Automate data quality processes or perform in-depth and rapid execution of data quality actions on datasets by leveraging advanced technologies such as AI/ML, knowledge graph, active metadata or NLP to minimize manual effort.

- **Multipersona usability:** Suitability of the solution to engage and support the various roles (especially nontechnical business roles) required in a data quality initiative: data engineers, stewards, DQ analysts, data architects, DI analysts, business analysts, data preparation and so on.

Driven by transition to more augmented data quality features, the market for data quality solutions is further consolidating with adjacent data and analytics markets such as metadata management, governance platforms, data integration tools and master data management (MDM) solutions. As a result, D&A leaders and practitioners expect seamless interoperability between these products driven by consolidation and sharing of metadata. Evaluating and selecting solutions has become less specialized and now requires greater collaboration with business leaders and different personas who plan to use them for various use cases.

Magic Quadrant

Figure 1: Magic Quadrant for Data Quality Solutions

Source: Gartner (November 2022)

Vendor Strengths and Cautions

Ataccama

Ataccama is a Leader in this Magic Quadrant. Its data quality product is Ataccama ONE Data Quality Suite. Ataccama has more than 460 customers for this product line and several thousand for its two freemium products, Ataccama DQ Analyzer and Ataccama ONE Profiler. Its operations are mostly in EMEA and North America, with clients primarily in the financial services, healthcare and manufacturing sectors.

Strengths

- **Market understanding, market responsiveness and technology innovation**: Ataccama shows superior market understanding and vision in the data quality solutions space. Ataccama has successfully enhanced its offering by adding key emerging features to its core capabilities, notably an AI-driven, metadata-based and time-series anomaly detection. Another key enhancement is a data quality rule creation with a self-learning engine to drive automation in data quality processes, all with a focus on explainable AI-core.

- **Unified and augmented data management solution supporting data fabrics/mesh:** Ataccama ONE is a comprehensive, tightly integrated platform for managing enterprise data with a focus on modern data architecture designs such as data mesh and fabrics. It provides data cataloging, data quality, MDM, reference data management, data insights and data integration. With an end-to-end experience based on activation of metadata and policy management, the solution supports data readiness across various use cases.

- **Overall customer satisfaction:** Gartner Peer Insights reviewers express high satisfaction (80% or more on average) in their experience with the vendor, the product's value for money, pricing flexibility and customer support. Ataccama clients using Gartner's inquiry service also express satisfaction in their partnership with the vendor and its willingness to work collectively toward their strategic goals.

Cautions

- **Market limitations:** Ataccama accounts for approximately 1% of the revenue in the data quality solutions market, which is low. Although Ataccama witnessed good growth in its data quality solutions revenue, its reach has been geographically limited. Most of Ataccama's customers (93%) are in North America and EMEA. Clients in other regions should carefully evaluate vendor support.

- **Challenging to use for complex data quality rules:** Ataccama reviewers on Gartner Peer Insights noted the steep learning curve in some areas, such as writing more complex rule expressions using advanced functionalities like the integrated development environment (IDE). It requires extensive training to effectively use, as well as more than modest technical ability to investigate problems if rules fail.

- **Pricing and licensing limitations:** Ataccama has moved to a subscription-only model for all new contracts. A perpetual option is available for exceptional cases only. Customers looking for perpetual pricing, especially for on-premises operations, should plan cautiously, and existing customers should consider any financial impacts. There is no standard approach for more flexible consumption-based/serverless metered pricing in cloud deployments.

Collibra

Collibra is a Visionary in this Magic Quadrant. Its data quality solution is Data Quality & Observability. Collibra is estimated to have around 105 customers mostly in North America and primarily in the financial services, healthcare and retail sectors.

Strengths

- **Mind share and market presence:** Though a new entrant to this Magic Quadrant, Collibra demonstrates market-leading mind share. It is increasingly present in competitive assessments focused on data quality requirements, as noted by users of Gartner's client inquiry service. Collibra data quality solutions are frequently mentioned by clients who are not already using Collibra Data Catalog.

- **Data observability and automated anomaly detection:** Collibra has invested in developing comprehensive profiling capabilities, upstream observability and monitoring capabilities, and anomaly detection. By combining these features, Collibra automates rule creation and execution with minimal manual intervention. Continuous monitoring also allows Collibra to ensure that rules are automatically updated with changes in the data or semantics. This reduces the process overhead while increasing responsiveness for end users, allowing them to enforce data quality rules for issue identification. However, some automated rules can be simpler in nature and can be further customized with manual intervention for more complex logic.

- **Customer experience:** Collibra Data Quality & Observability offers flexible pricing options with free trials, with targeted user interfaces for personas like data engineers (observability focus) and data stewards (governance focus). Collibra offers rapid-start packages and plans to further expand the user community, similar to its metadata management user community.

Cautions

- **Lacks some core data quality features:** Collibra offers very limited parsing, standardization and cleansing capabilities. More demanding use cases, such as master data management and data engineering, require extensive focus on data transformation and remediation (including the ability of business users to take corrective action). For these, Collibra customers must write their own scripts or use alternative options.

- **Low market share and geographically limited:** Collibra holds a tiny market share in the nearly $2 billion data quality solutions market. It is mostly deployed in North America (80% revenue) with little presence in the Asia/Pacific region. Potential customers should consider and confirm the availability of global support.

- **Lack of integrated and unified solution:** Collibra acquired OwlDQ more than a year ago, but there has been limited integration combining the capabilities of data quality and metadata management offering. Collibra markets data quality as part of its broad portfolio, but product level integration has yet to happen. Integration for lineage, governance and cataloging benefits is still part of its roadmap. End users will need to use marketplace-driven packages to consolidate some of these capabilities.

Datactics

Datactics is a Niche Player in this Magic Quadrant. Its data quality product is Self-Service Data Quality platform, which contains four components: DQ Manager, DQ Clinic, FlowDesigner and AI Server. Datactics has an estimated 115 customers mostly in EMEA and North America, and its clients are primarily in banking and government.

Strengths

- **Business-centric, customer-first approach:** Datactics positions itself as a self-service-oriented data quality solution serving business personas in data quality programs. Datactics plans to work toward self-driving data quality solutions; it plans to automate and augment most of its

data quality capabilities over the next two to three years. Its focus is on replacing complex technical features with low-code/no-code designs for business users.

- **Growing partnerships**: As a small, pure-play data quality vendor, Datactics does not have immediate plans to add more in-depth functionality for data governance and metadata management. However, over the last year it has used partners for these requirements, such as partnering with Solidatus and Manta for additional metadata- and lineage-driven augmentation. The combined capabilities have been deployed at some large organizations.

- **Roadmap with AI/ML-driven features**: Datactics is investing in expanding augmentation and automation capabilities with AI/ML. Its long-term vision involves further automating profiling and data discovery, using the discovered patterns and analysis to automatically infer and apply data quality rules for issue identification and then apply possible cleansing algorithms. It plans to transition to a web-based client for more complex rule authoring.

Cautions

- **Slow sales execution, geographically limited and low market share**: Datactics has a very low market share in the nearly $2 billion data quality solutions market and grows at about 4.5% year over year compared to data quality market growth of 6.4% in 2021. Also, Datactics has added an average of just five new customers every year. More than 85% of these customers are in the EMEA region. New customers, especially those with global presence, should verify whether the vendor's support operations will meet their needs.

- **Limited mind share and lack of focus on SaaS growth**: Datactics is rarely mentioned by users of Gartner's client inquiry service and almost never seen in competitive evaluations against other vendors in this Magic Quadrant. Datactics is a small company with approximately 70 employees. It engages in direct sales, with almost all revenue coming from on-premises installations. No revenue is derived from cloud and SaaS deployments.

- **Limited industry reach**: Datactics has limited industry reach and is most active in financial services and government, with approximately 95% of its customers in these sectors. Prospective customers in other sectors should check that this vendor's technology, services and industry knowledge will meet their business requirements.

Experian

Experian is a Challenger in this Magic Quadrant. Its data quality products include Experian Aperture Data Studio and other individual data validation and enrichment products for contact data. Experian has more than 250 enterprise customers for its flagship data quality suite, Experian Aperture Data Studio, and several thousand for its data validation offerings. Its operations are geographically diversified, and its clients are primarily in the financial services, public and retail sectors.

Strengths

- **Market share expansion**: Experian continues to grow strongly in the data quality solutions market. It continues to command a strong market share for its data quality products and

services, with over 70% of its data quality-related revenue from its data validation services, which can be easily used without technical expertise.

- **Marketing execution and sales execution:** Experian has been successful in transitioning its clients from its legacy data quality suite offering, Pandora, to Aperture Data Studio. The number of customers now using Aperture Data Studio has grown 55% from the previous year , with an expected correlating drop in the number of Pandora customers.

- **Improved product strategy and key partnerships:** Experian has added automated issue resolution and assignments based on anomaly tracking, ML-driven deduplication, improvements in reporting, and improvements in data enrichment using Experian Data in customer domains. It has also partnered with key metadata management platforms (like Alation, Alex, Collibra and others) to support improved insights and decisions in connection with quality processes.

Cautions
- **Limited mind share:** In client inquiries, Experian is mentioned less often in competitive evaluations against Leaders in this Magic Quadrant. Experian is a pure-play data quality vendor and limited in its offering when compared to unified data management offerings available from Leaders in this research.

- **Customer domain-oriented:** Experian excels at data validation services for customer data, but Gartner client interactions and other data points show limited adjacent use cases. Even though the vendor's products can support various data quality scenarios, usage lags for emerging and more complex use cases like data mesh, logical data warehouses and AI/ML development.

- **Limited augmented data quality capabilities:** Experian has started to work toward making use of ML in some of its features. However, some of these features still lack maturity in augmentation and automation; specifically, those that make more integrated use of ML over metadata for rule inference, rule association and remediation suggestions, especially for complex data quality issues.

IBM

IBM is a Leader in this Magic Quadrant. Its data quality products are IBM Watson Knowledge Catalog

(WKC), QualityStage, Match360 powered by IBM Cloud Pak for Data, and InfoSphere Information

Analyzer and QualityStage powered by InfoSphere Information Server. IBM has 2,500-plus customers for these product lines. Its operations are geographically diversified, and it has clients in various sectors.

Strengths
- **Unified and integrated platform:** IBM WKC offers a comprehensive and tightly integrated platform for managing enterprise data governance, quality and collaboration. IBM differentiates itself by providing an end-to-end data management experience rooted in metadata and policy

management. The solution also embeds several accelerators at no extra cost that can support extensive governance and compliance-driven data quality needs.

- **Augmented data quality capabilities, support for data mesh/fabrics:** WKC uses integrated metadata management functions to automatically assign data quality rules as part of catalog curation activities. Combined with out-of-the-box automation rules, data stewards can address common data quality issues with minimal setup and configuration. The innovation in metadata and AI-/ML-driven automation in its cloud offering, along with fully integrated quality and governance capabilities, make IBM a leading vendor in competitive evaluations.

- **DataOps:** IBM WKC is well-positioned to support end-to-end DataOps and AI life cycle scenarios. IBM has improved its vision with the recent acquisition of data observability provider Databand.ai and a close partnership with lineage provider MANTA. IBM customers can now find, track and prevent issues closer to the source and for a more streamlined operational approach to managing data.

Cautions

- **Unclear product migration information:** Several Gartner clients have mentioned a lack of clarity, detailed roadmap, status of current feature migration and detailed documentation on migration plans. Clients continue to seek more automated migration support from IBM for moving to WKC. Clients planning to migrate to WKC should seek better clarity on IBM's level of support and information on new or upcoming enhancements and innovations.

- **Reliance on legacy products:** WKC does not yet fully support streaming data at the level as offered by legacy InfoSphere QualityStage. Prospective customers should note that as of now some of these enhancements, especially end-to-end usability improvements, are part of the WKC roadmap.

- **Complicated, expensive licensing:** IBM customers continue to cite challenges in understanding complicated licensing. Additionally, relatively high costs impede adoption of IBM tools by new customers. This has led to some of IBM's prospects viewing its solution as complex and only suitable for large enterprises with highly skilled staff. Gartner client inquiry data indicates that small or midsize enterprises rarely consider WKC unless they are current IBM customers. While IBM has introduced consumption-based pricing, uptake is slow.

Informatica

Informatica is a Leader in this Magic Quadrant. Its data quality products are Informatica Data Quality (IDQ), Intelligent Data Management Cloud (IDMC, which includes Cloud Data Quality and Cloud Data Governance) and Informatica Data as a Service. Informatica has 5,700-plus customers for its offerings. Its operations are geographically diversified, and it has clients in various sectors.

Strengths

- **Market understanding, market responsiveness and leading mind share:** Informatica demonstrates tremendous understanding of the data quality market, with agile response to market disruptions. Informatica has delivered several innovations like intelligent automation

making use of metadata activation, support for hybrid cloud deployments and others. Informatica is frequently mentioned in Gartner inquiries based on competitive evaluations of data quality solutions by enterprises of all sizes.

- **Augmented data quality:** IDMC offers a fully integrated data quality and data governance platform with extensive automation supported by CLAIRE intelligent capabilities. Informatica has worked on adding and enhancing more automation features. With DQ Insights, it makes use of AI to spot anomalies and convert to reusable rules across IDMC. Informatica has also enhanced its data classification and metadata-driven insights for intelligent rules association and automation of remediation workflows.

- **Automated migration factory:** Informatica has been working extensively to support its existing customers' migration to the cloud for various data management needs including data quality. Customers planning to migrate to cloud can use the automated capabilities and assess what portion of the overall migration effort can be automated, reduced or made more efficient.

Cautions

- **On-premises future roadmap and innovation awareness:** Informatica continues to support its customers using on-premises product IDQ and data as a service (DaaS); however, many are unaware of the latest product innovations added to these offerings. Informatica's significant product updates are targeted toward IDMC. Some customers are not aware of the automated migration factory and expect product innovations to be added to IDQ that have been added to CDQ to continue tackling complex data quality requirements.

- **On-premises products pricing:** Informatica customers on Gartner Peer Insights and client inquiry services continue to report pricing challenges for on-premises offerings like IDQ, often combined with MDM offerings. Though IDMC has a simple, user-friendly approach to pricing, customers have concerns over migrating existing pricing arrangements to the cloud.

- **Growing competition:** Gartner client inquiry service users note that a focus on more comprehensive end-to-end data management platforms in the cloud, as Informatica has, can be overwhelming for midsize to small buyers. The data quality market has seen rapid innovation further extended by data observability, with smaller vendors offering competitive pricing and less complex offerings for rapid adoption. Clients can benefit from awareness of IDMC's availability as a module without the weight of the entire platform.

Innovative Systems

Innovative Systems is a Challenger in this Magic Quadrant. Its data quality products include the Enlighten Data Quality Suite and FinScan, both part of the Synchronos Enterprise Customer Platform. Innovative Systems has over 1,000 customers for these products. Its operations are mostly in the Americas and EMEA, and its clients are primarily in banking and securities, insurance, and media.

Strengths

- **Core data quality functionality:** Innovative Systems offers robust data quality features coupled with a focus on maintaining rich knowledge bases which augment its precise

matching/linking/merging, and parsing, standardizing and cleansing capabilities. Its solution uses knowledge bases in automated fashion enabling higher levels of accuracy, precision, and speed especially for business-facing users and applications.

- **Streaming data support especially in the financial domain:** Innovative Systems supports key data quality functions for low- to zero-latency data streams. FinScan offerings are designed to apply DQ functions to cleanse, merge and deduplicate customer records in real-time transactions. Innovative Systems plans to build additional capabilities to process payments natively in the SWIFT network and to apply for SWIFT certification.

- **Ease of doing business:** Customers continue to report ease of doing business with the vendor, especially the presales and postsales experience, product technical support, dedicated professional services, and guidance for end users. Innovative Systems continues to maintain a loyal customer base and has long-term strategic relationships with many of them.

Cautions

- **Limited augmented data quality features:** Even though Innovative Systems claims to have improved its use of AI through a crowdsourced AI approach, their product lags in innovation. Solution does not make use of supervised and unsupervised AI/ML to speed up DQ processing with semantic tagging, automated classification, inference of rules, automated anomaly detection and active usage of metadata.

- **Limited market presence:** Innovative Systems has been part of the data quality solutions market for many years, with recurring revenue and a growing customer base, but has still not captured significant market presence. There is a lack of targeted messaging and support for major trends like data mesh/fabrics, DataOps and data observability. Gartner inquiry data indicates it is seldom mentioned in competitive assessments.

- **Limited industry and domain experience:** Innovative Systems derives close to 80% of its revenue from the banking, securities and insurance sectors. Although financial services is one of the most demanding industries for data quality initiatives, prospective customers in other sectors should confirm that this vendor's technology and services will fully meet their business requirements. Additionally, Gartner sees Innovative Systems' product suite mostly used for contact data validation and enrichment.

Melissa

Melissa is a Niche Player in this Magic Quadrant. Its data quality products include Contact Zone, Data Quality Suite, Sentient Suite and Unison. It also offers specific data quality modules for healthcare and life sciences. Melissa has more than 1,000 customers for these products. Its operations are mostly in North America, and its clients are primarily in the media, communications and healthcare industries.

Strengths

- **Product and market strategy:** Melissa has worked on improving its data quality product and market strategy. Improvements include ease of use with more drag-and-drop and no-code

design features, use of metadata, knowledge graphs for greater DQ automation, and AI/ML-based reasoning for augmented use of data quality features.

- **Sophisticated identity verification process**: Melissa's electronic identity verification (eIDV) solution is one of the vendor's key strengths. It offers more comprehensive data validation and verification services in the customer domain than its competitors. It offers complete identity verification, starting with address verification and capabilities to match all contact elements to the address, verify identification documents and screen individuals across numerous watchlists. Validated identity can be corroborated and trusted against multiple sources from a dataset of billions of active and historical records from global authorities.

- **Customer experience**: Reviewers on Gartner Peer Insights expressed high satisfaction with their overall experience with the vendor, citing pricing and contract flexibility and quality of customer support. They also emphasize the timeliness of both support required and software updates.

Cautions

- **Limited mind share**: Melissa is less frequently mentioned by users of Gartner's client inquiry service and is rarely seen in competitive evaluations against other vendors in this research. From a data quality perspective, Melissa is mostly mentioned for contact data validation and verification needs, but not for broad enterprise-level data quality implementations.

- **Limited geographic coverage**: Melissa generates an estimated 95% of its revenue from North America, and it has minimal presence in the rest of the world. Even though Melissa now has local geographic presence in the U.K., Germany, Singapore and India, prospective clients should verify Melissa's level of local support for more complex and demanding requirements.

- **Innovation and functionality**: Melissa has enhanced some data quality tasks (for example, ML-based record matching and enrichment). However, the vendor still lags its competitors in using emerging technologies, such as active metadata and natural language processing (NLP), to build automation in broader data quality areas. Organizations looking for data quality solutions for more demanding use cases, such as data governance, should exercise caution. Larger organizations looking for a full suite and more unified data management platforms also should validate the available features and level of support available.

MIOsoft

MIOsoft is a Niche Player in this Magic Quadrant. Its data quality products are its MIOvantage platform and MIOvantage Data Quality Explorer. MIOsoft has an estimated 485 customers for these products. Its operations are mostly in Asia, with presence in North America and EMEA. Its clients are mostly in telecommunications, insurance and the public sector.

Strengths

- **Streaming data support with improved scalability and performance**: MIOsoft's primary focus is delivering a highly performant and scalable platform for large organizations or complex use cases. It invests as much as 70% of its revenue on R&D (one of the highest levels in the data quality solutions market). MIOsoft's products are powerful when doing real-time data

processing, including application of data quality logic on streaming data or Internet of Things (IoT) data.

- **Customer relationships:** MIOsoft continues to serve its loyal customer base with dedicated support and long-standing partnerships. Customers praise the presales and postsales support and implementation experiences. MIOsoft is engaging with more OEM partnerships to make some components available as open-source. Customer response to these activities has been positive, leading to several thousand new installations of its data quality products.

- **Sophisticated entity resolution rules:** MIOsoft continues to invest, innovate and advance its full graphical conflict resolution rule (CRR) engine, including continuous entity resolution on streaming data, with features for selecting and building the elements that will constitute the result. Customers praise the flexible, multientity and multidomain entity resolution that can discover relationships between entities, even when these relationships are not explicitly identified in the data.

Cautions

- **Lags in emerging technologies:** MIOsoft is innovative in some data quality areas, such as ML-based record matching, entity resolution, streaming data processing, and performance and scalability. However, the vendor still lags its competitors in using emerging technologies such as metadata, lineage, ML-backed monitoring, workflow automation, issue resolution and NLP to build automation in broader data quality areas.

- **Lack of mind share:** MIOsoft is rarely mentioned by users of Gartner's client inquiry service and almost never seen in competitive evaluations against other vendors in this research. Customers should note that MIOsoft is a small company with a significant portion of its revenue derived through indirect sales.

- **Minimal marketing strategy model and limited social media presence:** MIOsoft has followed a traditional marketing strategy approach even after 20 years in this market. It continues to grow primarily through references and word of mouth. MIOsoft's presence on external platforms like YouTube, LinkedIn and other dedicated platforms has declined or is nonexistent. MIOsoft also doesn't offer formal user groups in any geography.

Precisely

Precisely is a Visionary in this Magic Quadrant. Its data quality products include the Precisely Trillium Quality, Precisely Spectrum Quality and Precisely Data360 DQ+ product families. Precisely has over 4,300 customers for these products. Its operations are geographically diversified, and it has clients in various sectors.

Strengths

- **Market vision:** Precisely has created a detailed future vision and roadmaps for all three product lines. Several of its key focus areas include enhancing Data360's metadata capabilities for improving, augmenting and automating data quality functions across its other product lines, which lack comprehensive metadata capabilities.

- **Market share and combined portfolio:** Precisely holds over 9% market share by revenue in the nearly $2 billion data quality solutions market. Its expanded market presence and market share is due to a combined portfolio of three DQ product lines from two acquisitions over the last three years. This has allowed Precisely to engage with new market opportunities and experiment with more cross-selling and upselling.

- **Flexibility of use:** Reviewers on Gartner Peer Insights express high satisfaction with the ease of use of the Data360 data quality suite. They specifically call out drag-and-drop features for key data quality functionality from profiling to visualization, rule creation and updating, and tracking of metadata changes. These features enable business user engagement, especially in data governance and data science use cases.

Cautions
- **Limited product integration:** Precisely has three distinct data quality product lines integrated at the API and metadata level. It has not achieved a level of integration to make most optimum use of features that depend on metadata activation, semantic tagging, AI/ML usage for pattern analysis, fingerprinting of data, and automated anomaly/outlier detection. Precisely is working on further integration but customers are uncertain of how the products today will be extended.

- **Unclear availability of SaaS-based Data Integrity Suite:** While Precisely has shared its vision for SaaS-based Data Integrity Suite, as of spring 2022, it had not clearly articulated the availability and development of the Data Quality module. Prospective customers will need to check with the vendor for detailed product roadmaps and to learn what is available today.

- **Limited net new sales focus:** Precisely has expanded in the data management space through acquisitions, offering a wide range of opportunities to expand across a broader customer set. As a result, sales focus has been centered on cross-selling to adjacent product lines, rather than addition of new clients (65 new enterprise data quality customers added in 2021). While Precisely is expanding the almost one-tenth of the market share that it has in its favor, growth into net new accounts is relative to market and existing account penetration.

Redpoint

Redpoint is a Challenger in this Magic Quadrant. Its data quality products are Redpoint rgOne, Redpoint Data Management and Redpoint In Situ. The vendor has an estimated 325 customers for these products. Its operations are mostly in North America and EMEA, and its clients are primarily in the retail, financial services and healthcare sectors.

Strengths
- **Integrated customer data offerings:** Integrated with its customer data platform, Redpoint Data Management provides real-time automated data quality processing across many different channels and data sources. Redpoint supports entity matching and resolutions to create and maintain a real-time, single customer view for marketing, sales and operational use cases.

- **Product and marketing strategy:** Redpoint has worked extensively on its product strategy of making product deployments available in a wide range of options, from on-premises to cloud.

Redpoint recently launched In Situ, which provides core data quality capabilities via a cloud-native data quality-as-a-service model. The platform calculates a trust index for every data object, providing a view of data quality over time, recommendations for improvement, and enhanced visualization for DQ status and readiness.

- **Performance, scalability and flexibility:** Reviewers on Gartner's Peer Insights forum praise the performance and scalability of Redpoint's data quality products, specifically in supporting real-time data processing in Hadoop environments with billions of transaction records.

Cautions

- **Limited market presence:** Redpoint has a small customer base in this market, which is reflected in its relatively low revenue market share. It had a small increase in customer numbers in 2021, and remains relatively less-known in the market. It is rarely mentioned by users of Gartner's client inquiry service. A small company, Redpoint extends its reach through partnerships with global systems integrators (SIs). Potential customers should assess the vendor's resources to sustain product innovation and to support existing and new customers in this market.

- **Lack of augmented data quality vision:** Redpoint lags behind its competitors in using emerging technologies (such as metadata analysis and semantic tagging for rule recommendation, unsupervised AI for anomaly detection and rule inference, automated issue resolution using supervised ML, knowledge graphs, etc.) to build automation in broader data quality areas. Customers should note that Redpoint Automated Machine Learning is an add-on to its Data Management platform, and it is only natively embedded in rgOne.

- **Narrow focus on party data:** Redpoint actively supports party data in the customer domain, focusing on data quality for golden records management and MDM use cases. Gartner sees limited use and mention of Redpoint beyond customer and party data validation and enrichment. Prospects should be aware that the majority of Redpoint's data quality suite implementation is for customer party data and related transactional data, and assess its capabilities for other data domains.

SAP

SAP is a Leader in this Magic Quadrant. Its data quality products include SAP Information Steward, SAP Data Services and SAP Data Intelligence Cloud. SAP has an estimated 27,100 customers for these product lines. Its operations are geographically diversified, with clients in various sectors.

Strengths

- **Market share leader:** With close to a quarter of the data quality solutions market share (by revenue), SAP has been the market leader for more than a decade. SAP has a large established customer base derived from its strength in ERP software, strong brand recognition and an extensive partnership network with numerous technology providers, SIs and marketing partners across the globe. The majority of its expansion last year came from EMEA.

- **Unified data management:** SAP Data Intelligence Cloud, which is built on a containerized architecture, continues to be central to SAP data quality products. SAP integrates data quality offerings with its data integration, metadata cataloging, governance and orchestration capabilities. Its comprehensive portfolio provides added value to SAP customers with an integrated data management solution for data fabric and data mesh needs that also integrates with its business applications and processes.

- **Hybrid data management strategy:** SAP allows its customers to combine and switch between multiple implementation styles across on-premises and cloud. Most SAP customers are currently using SAP Information Steward and SAP Data Services on-premises/private cloud. SAP Data Intelligence Cloud expands this capability by offering hybrid deployment as part of SAP Business Technology Platform (BTP). Customers appreciate the reuse of Information Steward rules, data quality transformations, sharing of glossary terms within Data Intelligence Cloud, and the flexibility of executing jobs on-premises or in the cloud.

Cautions
- **Slower uptake of SAP Data Intelligence Cloud:** While SAP has a vision to have a single, unified, augmented cloud-based data orchestration and ML service that serves most data management needs, only a handful of SAP customers are using Data Intelligence Cloud for data quality. Customers should be aware that SAP Data Intelligence Cloud is the flagship platform that SAP prioritizes for investment.

- **Integration and implementation in non-SAP systems:** SAP's data quality products feature strong integration within SAP ecosystems, but some reviewers on Gartner Peer Insights report that connecting to non-SAP platforms remains challenging. It may require additional technical support or training from SAP or external service providers to be successful.

- **Challenges with user experience:** SAP's data quality products offer a long list of features that can transform data. However, some customers comment that it is a complicated process to configure and use them. Typically, training is required to use the tools effectively and efficiently, the cost for which is relatively high compared with competitors.

SAS

SAS is a Leader in this Magic Quadrant. Its data quality products are SAS Data Management, SAS Data Quality, SAS Data Loader for Hadoop and SAS Data Governance; and on SAS Viya: SAS Information Governance, SAS Data Preparation and SAS Event Stream Processing. SAS has an estimated 2,670 customers for these product streams. Its operations are geographically diversified with clients in various sectors.

Strengths
- **Enhanced product:** SAS Information Governance on SAS Viya now brings innovation to the data quality solutions space in monthly releases. Enhancements include intelligent tagging, improved text analytics, semantic assignment and classification, and directly building and applying rules as part of the cataloging feature. Some of this product development aligns with the SAS Cognitive Data Quality roadmap.

- **Extensible SAS Quality Knowledge Base:** SAS Quality Knowledge Base (QKB) is a repository with a host of data quality logic that can be fully customized, used and reused across all SAS' offerings. The QKB is available for 50 different locales and in 35 languages representing 46 countries. The QKB provides access to data quality functions throughout the entire SAS stacks of runtime environments, enabling the business to develop the logic once and use it as needed.

- **Metadata, lineage and open-standards roadmap:** SAS has been actively participating and investing in projects like Egeria, ensuring extension of its metadata and lineage capabilities. SAS connectors to Egeria are now free to the Egeria and SAS communities. SAS is exploring other open-standard projects like OpenLineage and Marquez to enable improved consumption and activation of metadata captured to be shared with all participating platforms.

Cautions

- **Unclear product positioning:** SAS offers multiple products for on-premises and cloud/hybrid scale requirements for data quality without proper differentiation or clear-cut documentation on their use. Some of this confusion is also attributed to comparison of SAS 9.4 engine versus SAS Viya. SAS has started providing more comprehensive updates on its newer offerings (SAS Information Governance on SAS Viya) but a majority of its data quality business is still based on legacy offerings.

- **Reduced mind share:** Gartner client inquiry data reveals a reduction in mind share for SAS data quality solutions. SAS also continues to receive relatively fewer reviews on Gartner Peer Insights compared to many of its closest competitors. Additional focus on emerging market trends like augmented data quality capabilities would help with market visibility.

- **Lagging innovation:** SAS lags its competitors in product innovation. It lacks native support for several key emerging technologies, such as data quality issue collaboration and resolution and auto inference or creation of data quality rules by leveraging active metadata and machine learning algorithms to provide best next actions. These enhancements are planned for SAS Information Governance.

Syniti

Syniti is a Visionary in this Magic Quadrant. Its data quality product is Syniti Knowledge Platform (SKP). Syniti has an estimated 518 customers for this product. Its operations are mostly in North America, EMEA and Asia, and its clients are primarily in the life sciences, retail, manufacturing and food sectors.

Strengths

- **Unified platform:** Syniti offers a unified platform based on user needs and has planned enhancements and innovation across distributed profiling, metadata scanning, data matching, ML-driven smart rule suggestion and overall cloud data quality automation. This integrated solution already provides cloud-native offerings, built-in knowledge graph capability, natural language narratives for reporting and process automation to support various data quality use cases.

- **360Science acquisition:** Syniti acquired 360Science in 2021 and has integrated its matchit capabilities to SKP. 360Science enhances SKP's data matching capabilities, both real-time and non-real-time, in party and other nonparty business data.

- **Strong partnerships:** Syniti maintains a strong relationship with SAP, with a global reseller agreement and elite partnership status in many SAP programs. In addition, Syniti provides cloud-based data quality, data migration solutions and implementation services for SAP ecosystems with common metadata exchange between the two. Syniti is building global alliances with other software partners (such as SIs and ERP vendors), leveraging its strengths in system integration and implementation.

Cautions
- **Lack of market visibility:** Though part of the data quality solutions market for over two decades and with a strong SAP consulting background, Syniti struggles to maintain market presence and market share. Syniti has very limited presence outside of SAP ecosystems and is rarely mentioned in Gartner client inquiries for competitive data quality vendor selection. This shows Syniti still derives most of its business from SAP-dependent customer bases. Prospective customers seeking more general solutions should carefully assess Syniti's capabilities outside of SAP-centric ecosystem deployments.

- **Missing key features:** SKP lacks certain data quality features, such as packaged functionality to address specific requirements of data generated by IoT solutions. There is also a lack of out-of-the-box automatic data standardization, where the tool can identify similar values within a data attribute and recommend and apply a correct standard value using ML.

- **Integration challenges with third-party data management tools:** Although the tooling is modular, including an open API, the only prebuilt interfaces include Collibra and SAP Information Steward. Integration beyond these products will require development of new interface connectors, potentially limiting options to leverage the automation features.

Talend

Talend is a Leader in this Magic Quadrant. Its data quality products include Talend Data Fabric and Talend Data Catalog with an estimated 2,000 licensed customers for these product lines. It also has two free open-source versions with limited functionality — Talend Open Studio for Data Quality and Talend Data Preparation Free Desktop. Its operations are geographically diversified, and its clients are primarily in media, financial services and manufacturing.

Strengths
- **Market growth:** Talend continues to grow strongly with a more than 30% annual revenue increase from last year. Talend holds a 5% market share in the nearly $2 billion data quality solutions market. Talend has successfully grown its cloud business; the revenue from cloud subscription and services increased to 25% of its total data quality revenue, up from 16% last year.

- **Market-aligned offerings:** Talend Data Fabric increasingly appears in competitive situations seen by Gartner. The product offers an integrated, end-to-end solution that includes capabilities in data quality, data catalog, data preparation and data integration, with MDM offered as an optional module. Talend received the highest number of reviews on Gartner Peer Insights over the past 18 months.

- **AI-driven augmentation and automation:** Talend has effectively leveraged AI in matching, parsing and cleansing activities. It introduced trainable and explainable matching algorithms with in-depth features. This improves user productivity on common data quality tasks related to data remediation and enrichment. Data profiling can support attribute and dependency analysis at the column level. It can automatically calculate trust scores utilizing metrics on metadata and data quality scores to support profile comparison and trend analysis.

Cautions

- **Limited use of AI- or metadata-driven automation in monitoring tasks and rule inference/creation:** Talend's ability to implement and deploy rules requires manual effort, but the identification of data quality issues is automated. However, Talend supports rule creation by providing a rule repository and interface for business users to create and apply data quality rules manually.

- **Pricing challenges:** Talend Data Fabric is packaged and sold primarily as an integration solution, where customers pay one price based on entitled users for the entire product. Customers express concerns around cost-efficiency if they only require the data quality components. Customers should verify pricing for their planned use of the solution and whether the vendor can support more modular needs without the weight of the entire platform.

- **Integration with third-party tools and UI complexity:** Gartner Peer Insights reviewers report that the most common issues identified while working with Talend relate to integration of Talend's offering with other third-party tools. The other common issue was complexity of user interfaces.

TIBCO Software

TIBCO Software is a Visionary in this Magic Quadrant. It offers TIBCO Omni-Gen DQ (rebranded after its acquisition of Information Builders) and TIBCO DQ. TIBCO has an estimated over 320 customers for these products. Its operations are mostly in North America and EMEA, and its clients are primarily in financial services, healthcare and the public sector.

Strengths

- **Strong vision for a unified data platform:** With the launch of TIBCO DQ 5.0, TIBCO added several features that include AI-assisted workflows, anomaly detection and rule generation as well as a knowledge hub of reusable components for profiling, validating, and fixing enterprise data. This allows users to deploy data quality anywhere needed with reusable knowledge and speeds up automation of data quality repetitive tasks. Some of these capabilities form the core of data fabric and mesh implementations.

- **Extensive partner network:** TIBCO continues to partner closely with all major cloud service providers, several SIs, and hundreds of value-added reseller (VAR) and technology partners. It has expanded its global presence to over 30 countries. This allows TIBCO to benefit in the areas of implementation, training and industrial expert services, which extends the vendor's reach to various sectors and regions. TIBCO has also scaled the reach and sales delivery of Omni-Gen.

- **Cloud-first approach:** TIBCO leverages a cloud-first approach with extensive focus on updating internal architecture. It is investing in other cloud features such as integration with TIBCO Cloud Metadata. TIBCO maintains a focus on moving its customers to cloud and supporting hybrid deployments, for which it provides fully hosted and managed services on all major cloud providers.

Cautions

- **Additional roadmap visibility required:** TIBCO needs to articulate the relationship more clearly between TIBCO DQ and Omni-Gen DQ. The vendor states that the relationship between the two is a natural release progression and no migration between releases is required. Initial communication to customers had begun as of this research period, and broad market communication of the migration path will continue.

- **Limited mind share:** TIBCO has a relatively small customer base in the data quality segment, leading to its relatively low market share. It is less frequently mentioned by users of Gartner's client inquiry service, and is rarely seen in competitive evaluations against other vendors in this research.

- **Lack of out-of-box data quality features:** TIBCO DQ currently does not provide out-of-the-box data quality features such as automated issue resolution, metadata-based rules automation, or inference and machine learning-based matching. However, some of these features could be implemented through TIBCO's data science and ML tools using the vendor's data science professional service team.

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

Collibra: Collibra has been included as a first-time entrant. Collibra acquired OwlDQ in February 2021. The integration of OwlDQ into the Collibra Data Intelligence Cloud introduces a new offering — Collibra Data Quality and Observability.

Dropped

No vendors have been dropped.

Inclusion and Exclusion Criteria

For Gartner clients, Magic Quadrant research identifies and then analyzes the most relevant providers and their products in a market. Gartner uses by default an upper limit of 20 providers to support the identification of the most relevant providers in a market.

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

To qualify for inclusion, vendors must meet all following inclusion criteria:

- Offer stand-alone software solutions that are positioned, marketed and sold specifically for general-purpose data quality applications. Vendors that provide several data quality product components or unified data management platforms must demonstrate that these are integrated, and collectively meet the full inclusion criteria for this Magic Quadrant.

- Deliver critical data quality functions. At minimum: data profiling; analytics and interactive visualization; data monitoring and detection; rule management and data validation; parsing, standardization and cleansing; matching, linking and merging; metadata and lineage; business-driven workflow and issue resolution; multidomain support; automation and augmentation.

- Support augmentation of the critical data quality functions listed above by demonstrating AI/ML features to improve utilization and ease of use.

- Support the above functions in both scheduled (batch) and interactive (real-time) modes.

- Enable large-scale deployment via server-based and cloud-based runtime architectures that can support concurrent users and applications. Cloud-based/SaaS version should support all critical data quality functions independently as mentioned in above criteria.

- Support multiple data domains and diverse use cases across different industries including support for real-time streaming data across these data domains.

- Maintain an installed base of at least 100 production paying customers (different companies/organizational entities) for their flagship data quality product (not individual smaller modules or capabilities). Alternatively: Achieve at least $5 million in total recognized revenue (per GAAP definition) for software (license, maintenance, and subscription and not any services) relating to data quality solutions in the calendar year 2021.

- The customer base for production deployment must include customers in multiple countries and in at least two of the following regions: North America, South America, EMEA, and Asia/Pacific.

- Support data quality functionality with packaged capabilities to process data in at least two different languages. The user interface should support multiple languages as well.

- Include a complete solution addressing administration and management, as well as end-user-facing functionality, for four or more of the following types of users: data steward, data

architect, data quality analyst, data engineer, data integration analyst, data scientist, business intelligence analyst and a casual user.

- Provide out-of-box and prebuilt data quality rules for the purpose of data profiling and monitoring, cleansing, standardization and transformation, based on common industrial practices.

- Support integrability or interoperability with other data management solutions such as metadata management, master data management or data integration solutions from third-party tools.

- Provide direct sales and support operations, or a partner providing sales and support operations in at least two of the following regions: North America, South America, EMEA and Asia/Pacific.

The following types of vendor were excluded from this Magic Quadrant, even if their products met the above criteria:

- Vendors that meet the above criteria but are limited to deployments in a single, specific application environment, industry or data domain.

- Vendors that support limited data quality functionalities or that address very specific data quality problems (for example, address cleansing and validation) are excluded because they do not provide the complete suite of data quality functionality expected of today's data quality solutions.

- Vendors that support only on-premises deployment and have no option for cloud-based deployment on any public cloud environment (for example, AWS, Azure or Google Cloud).

- Vendors that operate in only a single country and support only one language.

- Vendors that lack the integrability or interoperability with other data management solutions such as metadata, MDM or data integration solutions.

Honorable Mentions

Below are vendors that are frequently mentioned in data quality inquiries but did not meet the inclusion criteria for this Magic Quadrant. Depending on business goals, resourcing objectives, deployment preferences, geography or other factors, vendors listed here may provide viable alternatives. This is not an exhaustive list:

- ChainSys

- Datastreams

- Irion

- Oracle

Evaluation Criteria

Ability to Execute

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

Gartner evaluates vendors' Ability to Execute in the data quality solutions market by using the following criteria:

- **Product or service:** The vendor's core goods and services that compete in and/or serve the defined market. Included are current product and service capabilities, quality, feature sets, skills and so on. Products and services can be offered natively or through OEM agreements/partnerships, as defined in the latest Market Definition and detailed further in the subcriteria. Vendors are rated based on their ability to address current market needs, which also includes serving the much-needed augmented data quality capabilities.

- **Overall 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 offering and investing in the product (or products). The vendor's financial strength (as assessed by revenue growth, profitability and cash flow) and the strength and stability of its people and organizational structure are considered. This criterion reflects buyers' increased openness to considering newer, less-established and smaller providers with differentiated offerings.

- **Sales execution/pricing:** The organization's capabilities in all presales activities and the structure that supports them. Included are deal management, pricing and negotiation, presales support, and the overall effectiveness of the sales channel (direct and indirect sales). Most importantly we evaluate the flexibility, effectiveness, ease of adoption and market suitability of the vendor's pricing and licensing model in light of current and future customer demand trends and spending patterns (for example, operating expenditure and flexible pricing).

- **Market responsiveness/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. We evaluate the degree to which the vendor has demonstrated the ability to respond successfully to market demand for data quality capabilities over an extended period.

- **Marketing execution:** The clarity, quality, creativity and efficacy of programs designed to deliver the organization's message. This messaging is intended to influence the market, promote the brand and the business, increase brand awareness, 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, partnerships, promotional initiatives, thought leadership, social media, referrals and sales activities. We evaluate the overall effectiveness of a vendor's marketing efforts, the degree to which it has generated mind share and the magnitude of the market share achieved as a result.

- **Customer experience:** Relationships, products and services/programs that enable clients to be successful with the products evaluated. Specifically, we include the quality of technical and account support that customers receive. We may also include ancillary tools, customer support programs, availability of user groups, SLAs and so on. We evaluate the level of satisfaction expressed by customers with a vendor's product support and professional services. We also assess their overall relationship with the vendor, as well as customer perceptions of the value of the vendor's data quality solution relative to costs and expectations.

- **Operations:** The vendor's ability to consistently meet its goals and commitments. Factors considered include the quality of the organizational structure, skills, experiences, programs, the stability of key staff and other means that enable the vendor to operate effectively and efficiently.

Table 1: Ability to Execute Evaluation Criteria

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

Evaluation Criteria ↓	Weighting ↓

Source: Gartner (November 2022)

Completeness of Vision

The evaluation covers current and future market direction, innovation, customer needs and competitive forces, as well as how well they correspond to Gartner's view of the market.

Gartner assesses vendors' Completeness of Vision in the data quality solutions market by using the following criteria:

- **Market understanding:** This is the degree to which the vendor leads the market in new directions (in terms of technologies, products, services or use cases). Gartner evaluates the vendor's ability to adapt to significant market changes and disruptions, such as by providing data quality process automation with metadata and ML-driven rule recommendation, or by supporting real-time data quality processing on streaming data. Also considered is the degree to which vendors are aligned with the significant trend of convergence with other data-management-related markets — specifically, the markets for data integration tools, metadata and MDM solutions.

- **Marketing strategy:** We look for clear, differentiated messages, consistently communicated internally and externally through channels, social media, advertising, customer programs and positioning statements. Also considered are the degree to which the vendor's marketing approach aligns with and/or exploits emerging trends (such as adaptive data and analytics governance and business-centric data quality programs) and the overall direction of the market.

- **Sales strategy:** We look for a sound strategy for selling products that uses an appropriate network of direct and indirect sales resources; partnerships; and marketing, service and communication affiliates. The goal is to extend the scope and depth of the vendor's market reach, skills, expertise, technologies, services and customer base. We particularly assess the use of partnerships. A sound sales strategy also aligns sales models with customers' preferred buying approaches, such as freemium programs and subscription-based pricing.

- **Offering (product) strategy:** This criterion concerns the vendor's product development and delivery approach, emphasizing differentiation, functionality, product portfolio, methodology and features as they map to current and future requirements. It also covers the degree to which the vendor's product roadmap reflects demand trends, fills current gaps or weaknesses, and emphasizes competitive differentiation. Also considered are the breadth of the vendor's strategy regarding a range of product and service delivery models, from traditional on-premises deployment to SaaS and cloud-based models.

- **Business model:** This criterion concerns the design, logic and execution of the organization's business proposition for revenue growth and sustained success. We consider the vendor's overall approach to executing its strategy for the data quality solutions market. This approach includes delivery models, funding models (public or private), development strategies, packaging and pricing options, and partnership types (such as joint marketing, reselling, OEM and system integration/implementation).

- **Vertical/industry strategy:** We assess the vendor's strategy to direct resources, skills and offerings to meet the specific needs of individual market segments, including vertical markets. The degree of emphasis that the vendor places on vertical-market solutions is considered, as is the depth of its vertical-market expertise and provision of prebuilt data quality rules or libraries.

- **Innovation:** We assess the extent to which the vendor demonstrates creative energy in thought leadership and in differentiating ideas and product roadmaps that could significantly extend or even reshape the market in a way that adds value for customers. Particularly, we examine how well vendors support— or plan to support — key trends with regard to personas, data diversity, latency, data quality analytics, intelligent capabilities and deployment, for example.

- **Geographic strategy:** We evaluate the vendor's strategy to direct resources, skills and offerings to meet the specific needs of geographies outside its "home" geography, either directly or through partners, channels and subsidiaries, as appropriate. We do so in light of global demand for data quality capabilities and expertise.

Table 2: Completeness of Vision Evaluation Criteria

Evaluation Criteria ↓	Weighting ↓
Market Understanding	High
Marketing Strategy	Medium
Sales Strategy	Medium
Offering (Product) Strategy	High
Business Model	Low

Evaluation Criteria ↓	Weighting ↓
Vertical/Industry Strategy	Medium
Innovation	High
Geographic Strategy	Low

Source: Gartner (November 2022)

Quadrant Descriptions

Leaders

Leaders demonstrate strength in depth across the full range of data quality functions, including the core data quality capabilities that have existed for years, and bring more automation and augmentation to data quality solutions.

Leaders exhibit a clear understanding of dynamic trends in the data quality market. They explore and execute thought-leading and differentiating ideas, and they deliver product innovation based on the market's demands. The leading driver in the data quality solutions market is augmented data quality. That is, the solution can automate data quality processes, perform in-depth and rapid execution of data quality actions on datasets, or provide additional insights of data by leveraging advanced technologies such as AI/ML, knowledge graph, active metadata or NLP to minimize manual effort.

Leaders align their product strategies with the latest market trends. These trends include focusing on a nontechnical audience, trust-based governance, growth in data diversity, low data latency, data quality analytics (not just reporting) and intelligent capabilities (such as active metadata, AI/ML and graph technologies). Other trends are new delivery options (such as cloud, hybrid cloud and IoT edge deployment), and alternative pricing and licensing models (such as open source, pay as you go and subscriptions).

Leaders address all industries, geographies, data domains and use cases. Their products support multidomain and alternative deployment options such as SaaS or microservices. They offer excellent support for business roles and easy-to-use visualization and include out-of-the-box machine learning capabilities and predictive analytics.

Leaders offer extensive support for a variety of traditional and new data sources (including cloud platforms, data lake, IoT platforms, Hadoop, and mobile devices), a trust-based governance

model, and delivery of enterprise-level data quality implementations.

Leaders have significant size, an established market presence and a multinational presence (either directly or through a parent company).

Leaders undertake clear, creative and effective marketing, which influences the market, promotes their brand and increases their mind share.

Challengers

Challengers have established presence, credibility and viability, along with robust product capabilities and solid sales and marketing execution.

Challengers may not have the same breadth of offering as Leaders and/or, in some areas, may not demonstrate as much thought leadership and innovation. For example, they may focus on a limited number of data domains (e.g., customer, product or location data). They may not exhibit clear understanding of concepts like augmented data quality and may not have features related to these as part of their vision.

Challengers may lack capabilities in areas such as real-time profiling on streaming data, ML-based anomaly detection, predictive analysis or support for new data sources.

Compared with Leaders, Challengers often exhibit less understanding of some areas of the market, and their product strategies may suffer from a lack of differentiation.

Visionaries

Visionaries are innovators. Visionaries demonstrate a strong understanding of trends in the market such as augmented data quality. They are aligned with the market in adding the features related to automation and augmentation as a part of their roadmaps. They also have a vision toward more unified platforms that allows convergence of capabilities or integration plans with more adjacent markets for automation of data quality processes.

Visionaries also focus on a nontechnical audience, trust-based governance, growth in data diversity, low data latency, data quality analytics and intelligent capabilities (such as machine learning, metadata and knowledge graph). Also included are new delivery options (such as container-based or IoT edge deployment) and alternative pricing models (such as open source and subscriptions). Visionaries' product capabilities are mostly aligned with these trends, but not as completely as Leaders.

Although Visionaries can deliver good customer experiences, they may lack the scale, market presence, brand recognition, customer base and resources of Leaders. They have good vision but are relatively slow in execution of these great vision elements.

Niche Players

Niche Players often specialize in a limited number of industries, geographic areas, market segments (such as small and midsize businesses) or data domains (such as customer data or

product data). They often have strong offerings for their chosen areas of focus and deliver substantial value for customers in those areas.

However, Niche Players typically have limited market share and presence, have limited functionalities or lack financial strength. Niche Players often have to catch up with the latest innovations, such as the IoT (connectivity and deployment), machine learning and interactive visualization.

Context

Gartner has time and again highlighted the importance of having a critical approach toward managing the health and fitness of data, as poor data quality completely destroys business value. Trusted, clean and valid data is fundamental to achieving business goals and enablers for competitive advantage. Data quality is essential whether organizations keep running their business as usual, expand their business, evaluate and mitigate risks, or comply with regulations. Data quality is fundamentally tied to governance as part of a broader data and analytics strategy, but can be applied as part of various business processes at various stages and in different contexts as well, across a variety of data and analytics use cases.

All these challenges drive the adoption of data quality solutions. Data is useful only if its quality, content and structure is documented and well-understood. The cost of dirty, insufficient and/or inaccurate data remains a substantial threat. Delivering reliable, trusted and timely data for business consumption is a continuous effort and process that can be supported with modern technologies in data quality solutions.

Use this Magic Quadrant to help you find the right vendor and product for your organization's needs. Gartner strongly advises against selecting a vendor solely because it is in the Leaders quadrant. A Challenger, Niche Player or Visionary could be the best match for your requirements. Use this Magic Quadrant in combination with the companion Critical Capabilities for Data Quality Solutions and Toolkit: RFP Template for Data Quality Solutions, as well as Gartner's client inquiry service.

In addition, data and analytics leaders should take the following actions to improve data quality best practices and to optimize the use of modern data quality solutions:

- Adopt adaptive data and analytics governance by certifying the trust levels of data sources and data itself (see Leverage a Trust Model to Reset Data and Analytics Governance and CDAOs Must Use Adaptive Governance to Succeed in Digital Business). Identify the data quality requirements for each level of trust (such as what is "good enough"), because it may not be practical or possible to achieve 100% perfect data. This will enable a more focused approach to prioritizing data quality improvement efforts and create agility and flexibility for collaboration among stakeholders.

- Gauge the technical innovation of data quality vendors by evaluating the ability of their data quality intelligence capabilities, such as machine learning-driven rule suggestions and predictive analytics, to meet the challenges posed by data of increasing diversity and

complexity. You may start by assessing your incumbent data quality vendor's existing offering and future product roadmap for augmented data quality capabilities.

- Train nontechnical business users to become more data literate so that they can meaningfully use the new automation capabilities offered to them. Ensure that they can properly interpret and manage data given business context, and decide proper use of the data while adhering to data governance and data security requirements.

- Partner with business stakeholders to evaluate and monitor solutions supported with modern data quality solutions by checking for adherence to existing governance requirements and by establishing metrics to show tangible benefits.

Market Overview

The data quality solutions market has experienced moderate growth and remains extremely vibrant due to growing needs, especially during this period of economic uncertainty. Several organizations around the world have been severely impacted and are experiencing significant strain on their governance practices and data management architectures. Organizations have realized that there is a consistent and growing need for more resilient solutions that can be easily scaled and managed with less operational overhead. Vendors are responding by working on making data quality processes and practices more automated, thus supporting self-service capabilities through the introduction and continued enhancement of augmented data quality solutions.

The data quality solutions market — often viewed as a traditional tools market — is swiftly transitioning to a more advanced augmented data quality market. Vendors have been able to test and execute on their visions by introducing capabilities that can make better use of metadata through its activation or making use of multiple AI/ML techniques with useful automation as the output. This has led the market in the quest for further enhancement of available automation and augmentation while automating and augmenting more mundane, time-consuming, resource-consuming data quality activities.

Augmented data quality represents an enhanced capability to evolve traditional data quality processes — for improved insight discovery, best-next-action suggestions and improved accuracy — using metadata, knowledge graphs and AI-related technologies. This capability has shown a significant increase in efficiency and productivity by automating process workflows, minimizing dependency on humans, and reducing time to value. This capability continues to drive the data quality solutions market to another extent: Vendors have a more competitive advantage if they invest more in augmented data quality in their product offering. Vendors who fail to pivot will be irrelevant after the market reconsolidates.

Over the next several years, the data quality solutions market will completely transition to an augmented data quality market. In anticipation of this, we have introduced several changes to the market definitions and market analysis, and even introduced newer critical capabilities (see Critical Capabilities for Data Quality Solutions) based on the movement toward ADQ. This is further supported by an analysis of demand assessment through client inquiries. Additional detail

can be found in The State of Data Quality Solutions: Augment, Automate and Simplify, and Hype Cycle for Data Management, 2022.

From a revenue growth perspective, the data quality solutions market has shown acceleration. Data quality solutions market growth increased from 3.9% in 2020 to 6.4% in 2021 (see Market Share: Data Quality Software, Worldwide, 2021). The increased growth in 2021 was seen in all data management software markets as global markets recovered from the pandemic impacts. The two vendors with the highest data quality solutions revenue, SAP and Experian, continue to grow well. Informatica also remains strong in terms of revenue, despite its revenue declining further in 2021 due to its focus on iPaaS for data quality (iPaaS is tracked as a separate market by Gartner).

Gartner has observed the following market demands from the end-user side:

Flexible and scalable deployment:

- Flexible and scalable deployment options such as on-premises, cloud, multicloud, hybrid, PaaS, SaaS, DQaaS or data quality processes at edges or gateways

- Connectivity across heterogeneous data sources and landscapes from one single data quality platform

- Scalability over data sources, landscape, architecture, data volume, use cases and latency

Persona centricity:

- Empowerment to all types of users via no-code or self-service features in data quality processes within their own business context

- Persona-based user interface with dedicated data quality workflow based on roles and responsibilities of various users

- User community collaboration enhanced by AI/ML for data quality investigation and resolution

- Greater functional depth and business context with built-in workflow, knowledge bases and recommendations to support various uses cases across different industries and disciplines

Support for various use cases and data domains:

- Extendable to new use cases including data privacy, data protection, dataOps and data fabric

- Scalable framework to leverage existing data quality practices, rules and skills for extended use cases

- Party data remains most common, but other types of data are becoming important, such as geospatial data, IoT data, streaming data and transactional data

Data quality orchestration:

- Data quality practices that are democratized to integrate data quality into business processes and applications by leveraging AI/ML, graph and active metadata

- Collaboration workflow for data steward to propagate data quality best practices

Interoperability:

- APIs used to integrate with different systems, and data quality applied across the organization to all areas (applications and processes)

- Extensibility from a point solution with limited and targeted use cases to integrated solutions with adjacent markets for streamline processes

- Close integration with metadata solutions for sharing multiple free-standing and platform- or tool-embedded data catalogs as well as data lineage

Data quality processes:

- Data quality is integral to any business operation, so data quality begins to function as an embedded and essential part of all business processes

- Centralized data quality rule development and execution anywhere with API calls

- AI-/ML-driven data quality processes to minimize human interventions and errors by identifying similarity and applying predefined solutions; AI can also improve scope of data quality by covering unknown data and discovering unknown questions

Flexible pricing and licensing options:

- Simple and predictable pricing, such as consumption-based

- Flexible pricing to scale up or scale down

- Freemium or trial-out option for initial evaluation to lower the entry barriers

In response to these demands, the supply side of the data quality solutions market has pivoted toward greater adoption of ADQ. Several vendors have been successful in adding ADQ as part of their vision, but have not yet been able to deliver on that vision. These vendors that fail to deliver on ADQ capabilities will be regarded as traditional legacy vendors and will cede market share to the more innovative vendors that can drive quick and easy adoption.

These vendors along with product innovation will also have to take into account the ease of use, flexible deployments, unification of data management capabilities or native integration with third-

party tools (like metadata management, lineage and so on) for end-to-end seamless data management experience.

Finally, vendors will need to incorporate more recent developments like data observability (see Quick Answer: What Is Data Observability?). This will be an extension of augmented data management combining features from ADQ, active metadata and DataOps. In their approach to finding data issues, data observability solutions target automated anomaly and outlier detection with the ability to reuse these algorithms for rule development. However, these solutions are not dedicated to issue resolution.

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.

- 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 Provider research unit.

- 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.

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.