

Magic Quadrant for Data Management Solutions for Analytics

Published: 20 February 2017 **ID:** G00302535

Analyst(s): Roxane Edjlali, Adam M. Ronthal, Rick Greenwald, Mark A. Beyer, Donald Feinberg

Summary

Disruption is accelerating in this market, with more demand for broad solutions that address multiple data types and offer distributed processing and repository. Cloud solutions are also gaining traction. We help data and analytics leaders to weigh up the vendors in an increasingly dynamic space.

Market Definition/Description

This document was revised on 28 February 2017. The document you are viewing is the corrected version. For more information, see the Corrections

(http://www.gartner.com/technology/about/policies/current_corrections.jsp) page on gartner.com.

Organizations now require data management solutions for analytics that are capable of managing and processing internal and external data of diverse types in diverse formats, in combination with data from traditional internal sources. Data may even include interaction and observational data — from Internet of Things (IoT) sensors, for example — as well as nonrelational data such as text, images, sound and video. These requirements are placing new demands on software in this market as customers look for features and functions that represent a significant augmentation of their existing enterprise data warehouse strategies. Moreover, expectations are now turning to the cloud as an alternative deployment option, because of its flexibility, agility and operational pricing models. As the use of a combined cloud and on-premises hybrid is quickly becoming the norm, so organizations expect vendors to support them in enabling such deployments. Finally, the traditional data warehouse use case, while still the most common, is declining in importance. Among Gartner clients, traditional data warehouse inquiries are now fewer than those for the logical data warehouse (LDW; see Note 1 for a definition). This trend was first described in 2014 (in "The Data Warehouse DBMS Market's 'Big' Shift"), and is reflected in the change in name for the Magic Quadrant for 2017 (from "Magic Quadrant for Data Warehouse and Data Management Solutions for Analytics" in 2016). This change has resulted in an expansion of the types of vendors included and in the inclusion criteria becoming more challenging and therefore more difficult to meet.

For this Magic Quadrant, a data management solution for analytics (DMSA) is defined as a complete software system that supports and manages data in one or many file management systems (most commonly a database or multiple databases). These solutions include specific optimization strategies designed for supporting analytical processing, including (but not

limited to) relational processing, nonrelational processing (such as graph processing), and machine learning or programming languages such as Python or R. Data is not necessarily stored in a relational structure, and can use multiple models (relational, document, key-value, text, graph, geospatial and others).

Our definitions also state that:

A DMSA is a system for storing, accessing, processing and delivering data intended for one or more of the four primary use cases Gartner identifies that support analytics (see Note 2).

A DMSA is not a specific class or type of DBMS.

A DMSA may consist of many different data management technologies in combination. However, any offering or combination of offerings must, at its core, exhibit the capability to provide access to the data under management by open-access tools via APIs; for example, via Open Database Connectivity (ODBC), Java Database Connectivity (JDBC), Object Linking and Embedding Database (OLEDB), and others.

A DMSA must support data availability to independent front-end application software, include mechanisms to isolate workload requirements and to control various parameters of end-user access within managed instances of data.

A DMSA must manage the storage and access of data residing in a type of storage medium, which may include (but is not limited to) hard-disk drives, flash memory, solid-state drives and DRAM.

There are many different delivery models, such as stand-alone DBMS software, certified configurations, database platform as a service (dbPaaS) offerings and data warehouse appliances. These are evaluated together in the analysis of each vendor.

Magic Quadrant

Figure 1. Magic Quadrant for Data Management Solutions for Analytics



Source: Gartner (February 2017)

Vendor Strengths and Cautions

1010data

1010data (<https://www.1010data.com/>) 's offering consists of an integrated DBMS and business intelligence (BI) solution. Most of its customers are in the financial services, retail/consumer packaged goods, telecom, government and healthcare sectors.

STRENGTHS

Performance and ease of use: Reference customers report that 1010data's platform has exceptional performance, good ease of use, fast development of ad hoc analysis, and addresses large volumes of data without any significant difficulty. In addition, business users report direct usage of the entire platform.

Customer loyalty and penetration: Most of 1010data's customers utilize the platform for two or more internal business units or departments. The product exhibits good longevity in its account base (the existing base expands in parallel with the broader DBMS market), with most reporting utilization for at least two years and beyond (slightly less than half of its customer references use the product in the range of five to 10 years).

Managed service delivery: Because 1010data is primarily a managed service, users are generally on the latest release (including a steady stream of functional enhancements) and are fully supported. As one of the first managed service data warehouse providers, 1010data has a long track record in supporting its customers and maintaining their technology and delivery models.

CAUTIONS

Below-average perceived value: 1010data's reference customer feedback reveals an average number of issues encountered and also rates the aspects of doing business as average relative to the other providers in the Magic Quadrant. These same reference customers indicate that 1010data is ranked relatively low in terms of value achieved.

Small player: 1010data exhibits longer sales cycles and slower revenue growth when prospects increasingly compare it to cloud infrastructure as a service (IaaS) and platform as a service (PaaS) providers such as Amazon Web Services (AWS) and (more recently) Microsoft Azure. While these competing offers are actually no more or less proprietary than 1010data, the market perception is that they are more open.

Must compete against existing DBMS standard: Traditional data management solution vendors are already in use in organizations for a variety of transactional and operational data management needs. 1010data must continuously compete to be the analytics standard, leaving it vulnerable to replacement.

Amazon Web Services

Amazon Web Services (<https://aws.amazon.com/>) (AWS) is a wholly owned subsidiary of Amazon.com. AWS offers Amazon Redshift, a data warehouse service in the cloud; Amazon Simple Storage Service (S3) and Amazon Elastic MapReduce (EMR); and most recently, Amazon Athena, a serverless, metered query engine for data residing in Amazon S3.

STRENGTHS

Dominant cloud vendor: AWS is the dominant cloud vendor in this market by a significant margin, with only Microsoft Azure even close in terms of market share and presence. This dominance provides increasing network effects for all of its services, because the sources of data for a DMSA use case are more likely to reside in an AWS service than that of any other cloud vendor.

Best-fit solution approach: AWS offers several different services that can be used for different DMSA use cases. This approach allows customers to select only those capabilities needed for a defined use case, and the cloud nature of all the services helps to reduce the complexity of supporting multiple products.

Pricing: AWS has been a leader in low-cost, pay-as-you-go pricing, with discounting for optional term commitments. In addition to offering low-cost services, AWS also has a spot pricing model for EMR that allows clients to obtain additional resources at significantly lower cost in response to bids for excess capacity from other clients.

CAUTIONS

Monolithic storage and compute sizes: Amazon Redshift is available in configurations that include a specific amount of compute and storage. Customers cannot easily scale these resources independently without doing a resizing exercise, which takes several hours while data is redistributed. This can lead to excess capacity and a more cumbersome process to upgrade Redshift instances.

Cloud-only vendor: AWS solutions for DMSA are only available in the cloud. Some clients will need a hybrid environment that supports both on-premises software and cloud-based services, either as a temporary or a long-term solution. AWS does offer multiple services on the cloud, but no software for on-premises.

Integration issues: AWS takes a best-fit approach to its solutions (as detailed in the Strengths section above). By offering multiple independent services, AWS can reduce the complexity of each individual service, but this can increase the work required to integrate the separate services. At this point, integration requirements must be addressed by customers, so the proliferation of service instances can lead to significant integration issues.

Cloudera

Cloudera (<http://www.cloudera.com/>) offers Cloudera Enterprise — an Apache Hadoop distribution that combines components of Apache Hadoop and Spark, but also delivers its own components such as Cloudera Navigator for data governance, Cloudera Manager and Cloudera Director for cluster administration on-premises and in the cloud, Impala for SQL access and Kudu for analytics on transactional data. Cloudera's platform is available both on-premises and across multiple cloud environments (such as AWS, Microsoft Azure or Google Cloud Platform), including with cloud-native support for object stores.

STRENGTHS

Market presence: Among all Hadoop distributions, Cloudera is the most successful in this market based upon Gartner's published revenue numbers, partner traction and Gartner end-user clients' reported interest — demonstrating an ongoing year-over-year trend.

Cloud support: Cloudera's cloud support is progressing; it has begun, and will continue, to evolve its product to meet the requirements of cloud deployments. For example, Cloudera Director now supports the ability to spin up or down transient clusters as well as scaling up or scaling down clusters.

Technical support: Reference customers praise Cloudera for the quality of its technical support, which is essential given the limited availability of skills in the market.

CAUTIONS

Potential erosion of the core Hadoop stack: Cloudera, like other Hadoop distribution vendors, is being challenged as new processing alternatives (such as Spark) and new storage options (such as S3 for cloud object storage) offer alternatives that do not require a Hadoop stack. Cloudera is already addressing this risk by adding Spark to its distribution and offering direct access for files stored in S3.

Increased cloud competition: Demand for cloud solutions, and cloud and on-premises hybrids, is rapidly growing. Cloudera has been addressing these demands with its cloud-native capabilities and consumption-based pricing on AWS, Azure, and Google Cloud. However, its ability to drive traction on its cloud offering will be crucial and will require it to be combined with easier administration capabilities. Cloudera's reference customers point out the complexity of its UI and the expertise required for it.

Quality concerns: As the technology is being more widely used for more complex workloads and data of multiple formats, Cloudera's reference customers point out their concerns about maturity issues and bugs. Cloudera has been addressing these with enhanced scale and stability testing, as part of an overall quality initiative.

EnterpriseDB

EnterpriseDB (<http://www.enterprisedb.com/>) ships PostgreSQL with the EDB Postgres Standard subscription and EDB Postgres Advanced Server, based on the PostgreSQL open-source DBMS with the EDB Postgres Enterprise subscription.

STRENGTHS

Enhanced open-source DBMS: EnterpriseDB has added functionality to the open-source DBMS supporting clustering, high availability (HA) and disaster recovery (DR). It drives many of the large new features of PostgreSQL (including parallel query), and introduces some in Advanced Server (such as scalability improvements) before they are available in the PostgreSQL production releases.

Partners and cloud: EnterpriseDB has been building its partner ecosystem during the past few years and it is now beginning to "pay off" — with partners for database administrator (DBA) tools, application development, BI and analytics and packaged applications. Additionally, EnterpriseDB offers its dbPaaS, EDB Ark on OpenStack, Postgres Cloud Database on AWS and other cloud deployment options with EDB Postgres.

Favorable pricing and value: The most highly scored client reference survey results were for the pricing model and value (for the price) of EDB Postgres. One of the few open-source DBMS products with a core pricing model, it is a relatively low-cost option that also addresses virtualization subcapacity pricing issues well. All the client reference survey responses for value were either outstanding or just below.

CAUTIONS

Lack of market presence: PostgreSQL and EDB Postgres Advanced Server lack DMSA market presence because they are used primarily for operational DBMS use cases. EnterpriseDB will need to have more targeted marketing — specifically for the DMSA market — to grow this in the future.

Smaller database sizes: Most reference customers for EnterpriseDB have database sizes of less than 20TB. Complex workload management and large database support are missing in the basic PostgreSQL open-source product.

Missing functionality: Half of the reference survey customers for EnterpriseDB called out absent or weak functionality as an issue. This is to be expected, because this is an emerging market for EnterpriseDB. PostgreSQL has been primarily used for transactions, and DMSA functionality such as in-memory, built-in analytics and robust partitioning are only now beginning to be introduced .

Google

Google (<https://cloud.google.com/>) offers BigQuery on its Google Cloud Platform as a managed, in-memory query execution engine — along with its adjacent products (Dataproc and Dataflow) in order to provide a cloud-based data management solution for analytics.

STRENGTHS

Cloud brand recognition: Google BigQuery is a cloud-based solution for data management for analytics and this has broad market appeal. It also has a significant following in the cloud for multiple markets and therefore benefits from strong recognition of the Google brand. Importantly, Google has overtaken many of the "veteran" providers in this market during its first year of inclusion in the Magic Quadrant.

Ease of use and pricing: In terms of ease of use, implementation and pricing, Google's customer references rank BigQuery as one of the most effective offerings of all the suppliers evaluated. In addition to their current assessments, a new enterprise data warehouse (EDW) pricing model was launched at Google Horizon in September 2016, allowing easier price comparisons with other vendors that have a longer history in this market.

Roadmap: Google will be enhancing BigQuery's ease of use to better compete with existing suppliers in this market. The technology roadmap is focused on adding query sharing (sending a link to a previously built query), federated query processing (to add non-Google sources), utilizing machine learning (to enhance platform stability and performance), as well as leveraging fully enabled SQL capabilities to enable big data workloads (mostly on Hadoop) to be converted into DMSA approaches.

CAUTIONS

Growing maturity: The security model for BigQuery is dependent on Google Cloud Identity and Access Management services, which are still evolving. Customers report SQL functionality as a weaknesses; however, as of June 2016, BigQuery is compliant with ANSI SQL:2011. Additionally, Google released a new Java Database Connectivity connector and updated its Open Database Connectivity capability.

Support and professional services: Customer references for Google report that it is performing poorly (in comparison with the other vendors in this Magic Quadrant) for support and for any offering of professional services. Google's delivery model does not focus on

professional services and this market is evolving toward self-service and self-implemented approaches. For the current market delivery, Google has deployed a new Professional Services Organization (PSO) that came online in late 2016, with plans to expand in 2017.

DMSA market awareness: While Google is widely recognized, market awareness for BigQuery itself remains low. The reference customer base for this Magic Quadrant reports Google to be a competitor in approximately 6% of all competitive situations. In the Gartner inquiry base, 5% of clients mentioned Google in the context of data management and the percentage is even lower when focused on questions specific to the data warehouse.

Hewlett Packard Enterprise

HPE (<http://www8.hp.com/us/en/software-solutions/advanced-sql-big-data-analytics/>) offers Vertica Enterprise Edition, a columnar relational DBMS that is delivered as a software-only solution, certified configuration (HPE Converged System Reference Architectures), and cloud (HPE Vertica Machine Images for the Cloud) for AWS and Microsoft Azure with a bring-your-own-license model.

STRENGTHS

Renewed focus on Vertica: HPE has been refocusing its technology portfolio around Vertica as the delivery platform for analytical use cases. This has resulted in a product vision that articulates the major trends in the market and applies them to Vertica. For example, its cloud vision, integrating with Spark, or in-database execution of algorithms.

Sales execution and value for money: Vertica has benefited from strong growth that was well above the overall DBMS market forecast for 2016. This is the result of improved sales execution and pricing. Reference customers continue to praise Vertica for its value for money.

Performance and scalability: Reference customers indicate performance and scalability as major strengths for Vertica. Moreover, this is combined with implementation sizes that are above 100TB for more than one-third of the client references.

CAUTIONS

Portfolio uncertainty due to merger activity: HPE's spinoff/merger deal with Micro Focus (announced in September 2016) will bring yet another major organizational change. It is as yet unclear what the impact will be on Vertica from an innovation and technology roadmap perspective. HPE and Micro Focus have indicated their strategic commitment to Vertica's ongoing and future growth.

Strong competition in the cloud: HPE's cloud roadmap is moving in alignment with the direction of the market and has delivered support for both AWS and Azure; however, market demand for the cloud is already strong and all the major players have also delivered solutions to the market. HPE will therefore need to execute quickly on its cloud roadmap to maintain a presence. To address this need, HPE has formally adopted a model with a higher frequency of product releases during the course of the year.

Administrative concerns: Reference customers for HPE report some challenges with its administration capabilities, such as monitoring, backup and DR. In its recent Vertica release, HPE has made enhancements that should help to address these challenges.

Hortonworks

Hortonworks (<http://hortonworks.com/>) offers the Hortonworks Data Platform (HDP) on Linux and Windows. It also offers Hortonworks DataFlow (HDF) for streaming analytics on an on-premises basis and through various cloud providers. Hortonworks partners with Microsoft (for its Azure HDInsight service) for hybrid on-premises/cloud deployments, and offers a version of HDP on AWS. A free, laptop-capable sandbox version of HDP is available.

STRENGTHS

Popularity of Hadoop: Hadoop is the most visible offering in the nonrelational DBMS space, and Hortonworks is one of the two most prominent Hadoop-based vendors. Hortonworks' adherence to pure open source is seen as an advantage by some of the overall market.

Strong alliances with major players: Hortonworks has OEM relationships with, or partners with, many major players, including Dell, HPE, Microsoft (which based its HDInsight service on HDP), SAS and Teradata, among others. Hortonworks HDP is often the default choice for customers looking for a Hadoop distribution through their more traditional vendor.

Customer purchase intentions: Our reference customer survey shows that 80% of Hortonworks customers intend to purchase more product in the next 12 months, which places it high in relation to the overall vendor landscape.

CAUTIONS

Hadoop momentum slowing: Three years ago, the market for stand-alone Hadoop products was rapidly growing, with these products offering capabilities that were not available from major DMSA vendors. Today, virtually all major DMSA vendors have at least open connectivity with external storage, including Hadoop Distributed File System (HDFS); this capability allows for easy integration from those vendors, so the competitive advantage of pure Hadoop products has been significantly reduced. Hortonworks has added the ability to access data stored in Amazon S3 to its product.

Thin layer of value-added intellectual property (IP): Hortonworks has a vision for the market that is based on broad adoption of open source as a preferred delivery model. Hortonworks has contributed significantly to the open-source code base, as evidenced by its work on Apache Hive improvements. The market is currently seeking incremental IP to manage and administer open source. The market has not yet established a preferred position.

Weakness in traditional data warehouse use case: Pure open-source Hadoop is appropriate for context-independent use cases, but much less so for traditional data warehouse scenarios. This limitation makes it more difficult for pure Hadoop products, such as those from Hortonworks, to expand their footprint in organizations with a mixture of use-case requirements.

Huawei

Huawei (<http://e.huawei.com/en/products/cloud-computing-dc/cloud-computing/bigdata/fusioninsight>) offers FusionInsight, a data management platform combining components of Apache Hadoop, Spark and Storm, and a proprietary massively parallel processing (MPP) DBMS. Huawei has added industry-specific domain models, as well as proprietary extensions to the Hadoop platform for event-stream processing, graph and machine-learning capabilities, and a unified SQL engine that is compatible with its MPP DB and runs on Hadoop. Additional enhancements have been made to the Hadoop scheduler and to the HDFS file system.

STRENGTHS

International presence: Huawei is a global organization that is a worldwide leader in the server, storage, telecom and networking equipment markets. While these strengths have not yet fully translated to the DMSA market, Huawei is a trusted and well-recognized name — especially in its core market in China.

Extension of the Hadoop Core: Huawei has worked to extend Hadoop to include solutions for vertical industries, as well as the addition of a SQL-on-Hadoop query engine, a proprietary MPP DB, streaming and graph analytics capabilities, and enhancements to the scheduler and distributed storage layer.

Customer loyalty and technical support: All of Huawei's reference customers indicated their intent to purchase additional licenses, products or features within the next 12 months. Additionally, 90% of Huawei's reference customers consider its DMSA technology to be a standard within their organization. Notably, 70% of reference customers also cited technical support to be a key strength.

CAUTIONS

Crowded Hadoop market: Huawei may find it difficult to expand and differentiate itself in a global market that is already crowded with Hadoop vendors. Huawei FusionInsight is still mainly China-based and only starting to gain visibility in the DMSA market outside of China.

Mixed performance results: While a number of reference customers cited performance as a strength, most customers indicating that they reviewed and rejected Huawei's solution did so because of performance in a proof of concept.

Product plus consulting-based approach: While this may be reflective of the nature of the DMSA market as a whole, Huawei appears to be following a product plus consulting-based approach to project development — where FusionInsight serves as a baseline product offering that requires customization to meet customer needs. Notably, 50% of surveyed reference customers cited some concerns around product requirements and Huawei's roadmap responsiveness.

IBM

IBM's (<http://www.ibm.com/us-en/>) current offerings include the traditional solutions of stand-alone DBMS, DB2 LUW, DB2 for z/OS appliances (IBM PureData System for Analytics, IBM PureData System for Operational Analytics, and IBM DB2 Analytics Accelerator); Hadoop,

through IBM BigInsights; and managed data warehouse services. The dbPaaS IBM dashDB offering adds a private cloud capability. IBM DataFirst Method and IBM Watson Data Platform support further evolving hybrid cloud and on-premises deployment and management.

STRENGTHS

Performance, support and account management: Customer references report performance, a good balance of platform capability and open-source combined, and technical account management as strengths for IBM in 2016. The support, in particular, is a reversal of up to five years of customer frustration, and could possibly be attributed to IBM's new "design once and deploy anywhere" approach to its offerings.

Logical data warehouse capabilities: IBM's Fluid Query management was introduced several years ago and allows interoperability across diverse query engines. Tightly linked to the dashDB approach, the Fluid Query solution is designed to allow for the deployment of analytic databases and analytics processing across all of IBM's platforms. This approach includes leveraging various platforms in an on-premises and cloud hybrid environment, as well as in-database analytics support for Apache Spark.

Focus on analytics capabilities: In addition to the current breadth of offerings, IBM Watson Data Platform — which IBM indicates is an artificial intelligence (AI)-powered decision-making platform for developing advanced analytics — is seeing wider adoption. The primary focus is for adding in-depth analytics capabilities that span multiple data types across increasingly large datasets.

CAUTIONS

Complex portfolio: IBM still seeks to be "everything to everybody." Customers and prospects are advised, when engaging with IBM, to keep the focus on their immediate needs; separate IBM's additional capabilities in other components from your current needs. IBM promotes the concept of "design once and deploy everywhere," which should mean you don't need to buy everything at once.

Pricing and cost of operation: IBM's customer references still report dissatisfaction with its pricing and the cost to operate, as well as comments about missing functions. Customers also raise questions about inconsistency regarding the use of open-source components in the platform, which can cause some upgrade issues.

Challenging deployments: Customer references for IBM reveal mixed opinions on the ease of implementation (some indicate it as a strength, but in relative ranking IBM is in the bottom quarter compared with other vendors in this Magic Quadrant). References also report that IBM is a difficult business partner, and IBM has engaged in significant investment and program building to address customer partnering efforts in 2017. Overall, we interpret this to mean that getting IBM through the door and up and running is difficult, but once in place the technical support and ease of use create a better experience on balance.

MapR Technologies (<https://www.mapr.com/>) offers its Converged Data Platform (including both open-source and commercial software) with performance and storage optimizations using Network File System (NFS), nonrelational DBMS (Key-Value and Document models), streaming, HA improvements, and administrative and management tools.

STRENGTHS

Multiple use-case support: The Converged Data Platform supports a wide range of use cases — including streaming, operational and analytical — all running on the same platform. It has multimodel support (with native JSON, time-series and graph), an Apache Drill SQL query engine, SQL supported by MapR-DB, and also includes standard APIs for the support of products from multiple vendors (such as SAP and SAS).

Partnerships, cloud and regional expansion: MapR is now in North America, Europe and Asia/Pacific and has expanded its vertical coverage to include customers in most major industries. It continues to grow its partnerships, which include Cisco, HPE, Microsoft, SAP and Teradata. MapR offers cloud subscriptions with AWS, CenturyLink, Google and Microsoft.

Enterprise robustness: Reference clients continue to praise MapR for its enterprise-readiness, HA and cluster management (more than 80% of those surveyed). Also, the reference customers gave MapR high scores for ease of implementation and use — addressing a challenge from the previous year .

CAUTIONS

Lack of market visibility: MapR continues to lack visibility in the DMSA market. More effort needs to be put into market awareness of its vision and technology, so that the differentiation and uniqueness stand out. Although it has made progress, especially with its growing partnerships, Gartner's advisory service continues to receive fewer inquiries about MapR than about the other vendors in this Magic Quadrant.

Pricing issues: Although rated highly for ease of doing business, pricing can become an issue with MapR. Its server pricing is higher than for other Hadoop vendors and this often leaves it out of the decision process, especially for test and development implementations.

Lack of rapid ecosystem adoption: MapR's reference customers identified challenges with new functionality adoption (for example, lagging in minor open-source Hadoop project support). We believe the functionality issues stem from a strategy of providing a production-ready platform that only supports projects/versions that are ready for real deployments.

MarkLogic

Based in San Carlos, California, U.S., MarkLogic (<http://www.marklogic.com/>) offers an ACID NoSQL document store DBMS in Essential Enterprise, Global Enterprise and Mobile editions, and a free, fully featured developer version. Its solutions can be deployed via leading cloud and virtualization platforms, including those of AWS, Microsoft Azure and VMware.

STRENGTHS

Focus on unified access to data: MarkLogic positions itself as being the best way to bring together silos of data, allowing unified access to data stored in its DBMS and other DBMS systems. As organizations strive to extract greater value from different data sources, this focus is appealing and differentiating.

Customer satisfaction: Reference survey results indicate a high degree of customer satisfaction with MarkLogic. It also reports expanded footprints at some of its key customers, based on successful project implementations.

Revenue and geographic growth: MarkLogic has continued its multiyear growth, in both revenue and an expanding presence in global markets.

CAUTIONS

Change in focus: The new focus on data unification is a change from MarkLogic's traditional positioning as a document DBMS. It may take a while for this vendor to gain traction and customers from this message.

Mind share low: MarkLogic does not have as much market awareness as some other vendors in this Magic Quadrant. The number of inquiries that mention MarkLogic continues to be fewer than for its competitors, but it has increased year over year.

Implementation maturity: Survey respondents indicate that MarkLogic suffers from a number of issues common to vendors with a smaller market base, including a difficulty in finding people with the appropriate skills. This shortage is due, in part, to the difference in the way the MarkLogic product is architected, which leads to different skill set requirements for successful implementation.

MemSQL

MemSQL (<http://www.memsql.com/>) offers a distributed SQL scale-out DBMS with an in-memory row store along with a memory and disk-based column store that supports transaction and analytic use cases. MemSQL extends the DBMS platform to include real-time analytics with streaming data via Apache Spark or Apache Kafka.

STRENGTHS

Broad use-case support: MemSQL's platform supports operational, analytic, and real-time streaming use cases, making it ideal for a wide range of DMSA applications. Half of MemSQL's reference survey respondents also report using the platform for transactional use cases. MemSQL's unified approach serves to productize the Lambda/Kappa architectures.

Flexible deployment options: MemSQL is deployed on commodity hardware and supports on-premises and cloud approaches. Customer references also praised the availability of a community edition that makes it easy to adopt the technology before committing to the enterprise edition. Compatibility with the MySQL wire protocol is also frequently cited as a positive aspect that makes integration easier.

Real-time processing and customer engagement: MemSQL received very high scores from its reference customers for its support of real-time loading of data, with nearly 40% using the product in this capacity. Additionally, all MemSQL reference customers reported that

they were either running the latest version, or would be within 12 months.

CAUTIONS

Limited market awareness: MemSQL was rarely considered by those not selecting it, which indicates a lack of awareness in the wider market. Additionally, less than half of its reference customers considered MemSQL to be the standard within their organization.

Product maturity: Customer references indicated that the primary reasons for their not selecting MemSQL included missing or lesser functionality, general unfamiliarity or lack of comfort with the product, and poor performance during proof of concept (though this last could be related to skills and ease of implementation). The lack of features such as user-defined functions (planned for 2017) and the maturity of HA, backup and recovery, and administration tools were all cited as areas needing improvement. Customer references also noted the relative difficulty of finding relevant skills in the market.

Small vendor: As noted in last year's Magic Quadrant, MemSQL is a small vendor that is trying to address two very different markets simultaneously: the operational DBMS market, and the DMSA market. While the broad focus is appealing, and the vision is compelling, it could dilute MemSQL's product and marketing efforts and hamper its Ability to Execute.

Microsoft

Microsoft (<https://www.microsoft.com/>) offers SQL Server as a software-only solution, with certified configuration and a data warehouse appliance (Analytics Platform System, an MPP data warehouse appliance). It also offers Azure SQL Data Warehouse, Azure HDInsight (Hadoop distribution based on Hortonworks) and Azure Data Lake as cloud services.

STRENGTHS

Cloud leader: Microsoft demonstrates strong market understanding with regards to the cloud. With Azure SQL Data Warehouse it addresses the growing interest in cloud data warehousing, but also hybrid on-premises and cloud use cases, and begins to demonstrate hybrid capabilities with stretch tables.

Logical data warehouse: With Azure Data Lake supporting Apache Spark and U-SQL, it forms the basis of the LDW implementation in the cloud and supports data of multiple formats and various types of processing.

Cloud pricing flexibility: Azure SQL Data Warehouse pricing, which allows independent scaling of storage and compute, meets the market demand for flexible pricing and adjustment to changing compute capacity according to workload or time period.

CAUTIONS

New cloud solutions require a proof of concept: More than half of the Azure SQL Data Warehouse reference customers had deployments below 5TB. This is not surprising, because the technology was only released earlier in 2016, but suggests that customers looking at running large and complex mixed-workload data warehouses in the cloud should run a proof of concept.

Limited guidance to navigate LDW choices: Microsoft has a confusing approach to LDW implementation, with multiple competing accesses and processing engines including Spark, U-SQL and PolyBase. This requires clients to make the right choices if considering using one of these options as the logical query layer on top of their multiple-format data.

Cloud lock-in: Microsoft offers a very appealing set of products for many use cases, but the ability to integrate or combine with other technologies in a best-fit engineering manner appears to be limited. In the cloud, in particular, its portfolio is designed to encourage clients to combine multiple Microsoft cloud services — with a potential lock-in challenge for clients. This concern is being addressed by Microsoft through cloud partnerships offering a wide variety of options running on Microsoft Azure. Microsoft is also working on expanding its access capabilities to other sources, including Amazon S3, by leveraging the Metanautix technology (from its acquisition).

MongoDB

MongoDB (<https://www.mongodb.com/>) offers both an open-source and commercial nonrelational document DBMS. The offering supports automatic sharding, failover, secondary indexes (including arrays), geospatial data and text search, as well as management tools (cloud-based and on-premises). MongoDB is offered as on-premises software and MongoDB Atlas, a dbPaaS solution.

STRENGTHS

Enhanced capabilities and cloud: MongoDB has been busy enhancing the capabilities of the DBMS with stronger HA governance and management functionality, including the introduction of pluggable storage engines (and the addition of WiredTiger). This has led to a wider use for data warehousing, especially as part of the LDW and for operational analytics. In June 2016, MongoDB added MongoDB Atlas, a dbPaaS solution.

Growing DMSA use cases: Although MongoDB is used more for operational DBMS use cases, it is growing in use cases for DMSA. This year, its reference customers stated their use for traditional and logical data warehousing in addition to the expected use as an operational data warehouse.

Ease of use and flexibility: This year, MongoDB's reference customers again called out the schema flexibility as a strength. They also called out the ease, speed and flexibility of programming with MongoDB. Although this is not surprising with a large developer-based community, this flexibility and ease of use is now spreading to the end-user side of the customer experience.

CAUTIONS

Market perception: MongoDB is primarily an operational DBMS and, as such, the market's perception of it is not as a DMSA system. If MongoDB desires momentum in DMSA it must spend more time and budget on marketing programs directed at the DMSA market.

Limited DMSA functionality: The reference survey called out a high number of responses citing inadequate performance or scalability when using MongoDB as a DMSA. Aggregation query, text search, and performance of the BI Connector were specifically listed.

Level of expertise required: Reference customers for MongoDB reported that the majority of their users were in the "expert users" rather than the "business analyst" category, implying that the use of MongoDB for DMSA requires a high level of technical expertise.

Oracle

Oracle (<https://www.oracle.com/>) provides the Oracle Database 12c, the Oracle Exadata Database Machine, the Oracle Big Data Appliance, the Oracle Big Data Management System, Oracle Big Data SQL and Oracle Big Data Connectors. Oracle Cloud provides Oracle Database as a Service, Oracle Exadata Cloud Service, and Oracle Big Data Cloud Service

STRENGTHS

Technical vision and capabilities: Oracle has been a leader in DBMS technologies for decades, and its capabilities make it one of the two most prominent vendors in the DMSA market.

Integration with Hadoop distributions: With its Big Data SQL, Oracle extends its reach to a number of Hadoop distributions; providing not only virtual access, but also advanced features such as predicate push down to the target platforms.

Market presence: Gartner surveys show that more than 70% of DMSA purchase decisions stay with the established DBMS vendor for the organization. This factor, coupled with Oracle's strong technical capabilities in the DMSA area, makes it a popular choice for the use cases covered in this document.

CAUTIONS

Customer sentiment: As a premier technology vendor, Oracle charges premium prices and negotiates tough contracts, which can result in customer dissatisfaction. A significant portion of Oracle's installed base is less than happy with its business practices in this area.

Slow to the cloud: Oracle has been promising full cloud support for its DBMS products for more than five years. While the Oracle Database Cloud Service offering has now been available for more than two years, and dbPaaS offerings are now in production, other cloud vendors have been able to offer alternatives for years and have offerings in that market space.

Operational concerns: Although complaints about Oracle's capabilities are extremely rare, more than a quarter of survey respondents mentioned support, or the difficulty of finding skilled people, as one of the main weaknesses of Oracle, which is on the higher side of all vendors covered.

Pivotal

Pivotal (<https://pivotal.io/>) offers its Pivotal Greenplum Database as an open-source MPP database based on PostgreSQL. It also offers Pivotal HDB based on the open-source Apache Hawq project for SQL processing on top of Hadoop. Both solutions are offered as a software product or a fully managed service through Pivotal Data Operations Services, and can run either on-premises or in the cloud. Pivotal also offers an appliance configuration of Pivotal Greenplum in a joint configuration with Dell EMC.

STRENGTHS

Appeal of the open-source model: Pivotal has been able to attract new customers to its solution, thanks to its open-source model, while at the same time retaining long-term customers. Moreover, 60% of its reference customers indicate planning to purchase more from Pivotal in the next 12 months.

Renewed focus on Greenplum DB: In 2016, Pivotal began refocusing on Greenplum Database as a critical asset of its DMSA product portfolio, with new features for workload management and query optimization. Reference customers praised the query optimization performance gains.

Suitability for data science: Pivotal's solutions appear to be particularly suitable for context-independent use cases, leveraging Pivotal Greenplum and in-database analytics with Apache MADlib. Reference customers demonstrated an above-average number of data scientist and expert users.

CAUTIONS

Transition to open source: Pivotal has made its solutions open source and shifted from a license-based to a subscription-based revenue model. These radical changes in its business model constitute a risk during the transition phase, but may end up being what the market demands.

Lack of some expected functionality: Pivotal is late to deliver on some major market demands, such as in-memory capabilities or cloud in a dbPaaS model. In 2016, Pivotal focused on delivering to its core customers and new capabilities are on the roadmap.

Maturity of management and administration: Reference customers for Pivotal indicate gaps in management capabilities such as backup, restore and HA, which could be improved.

SAP

SAP (<http://go.sap.com/index.html>) offers both SAP IQ and SAP Hana. SAP IQ, is a column-store DBMS available as a stand-alone DBMS. SAP Hana is an in-memory column-store DBMS that supports operational and analytical use cases; it is also offered as an appliance (from more than a dozen hardware vendors), a cloud solution (public and private and SAP Cloud Platform, of which SAP Hana is one component) and a reference architecture (SAP Hana tailored data center integration [TDI]).

STRENGTHS

Broad DMSA enhancements: SAP has been enriching SAP Hana with new enhancement releases about every six months, now including stronger HA/DR capabilities, multitenant in the DBMS, streaming data, built-in analytics, integration with and extension of Spark with the SAP Hana Vora in-memory engine, and a strong data tiering capability based on SAP IQ. There is a clear direction from SAP to position SAP Hana as a DMSA platform for all use cases, both within and outside of the SAP ecosystem.

Maturity: As SAP Hana matures as a DBMS, it is used across all DMSA use cases. SAP now offers a version (SAP Hana 1.0) that is supported for three years as a stable platform, and SAP Hana 2.0 for those that want regular new releases with enhancements. This offers production customers the option of a stable release without major enhancements disrupting the production platform.

Performance: Performance was the most frequently cited strength from SAP's customer references. SAP's scores for overall performance were all above average, with half of its customer references rating it as outstanding .

CAUTIONS

SAP-only market perception: A combination of de-emphasizing SAP IQ as a stand-alone platform, the market's perception of SAP Hana as an expensive product, and the overwhelming use of SAP Hana in SAP application implementations, has held back SAP from being a top choice in DMSA implementations outside the SAP ecosystem. Based on Gartner client interactions, SAP has made some progress toward changing this perception. However, it is almost never considered for a platform outside the SAP ecosystem.

Use outside of SAP applications: SAP's positioning of the applications on SAP Hana has raised questions from some Gartner clients about the need for an SAP-specific data warehouse — SAP business warehouse (BW) and SAP BW/4HANA. SAP is driving the position of operational analytics as part of the applications (Suite on Hana and S/4HANA) leaving customers to ask, "why not use my data warehouse of choice for the historical data?"

Implementation and functionality: In 2016, some inquiry customers and survey references reported software issues. Reported challenges relative to SAP Hana include difficult implementation and weak or missing functionality. SAP's innovation strategy and roadmap do not always align with broader market expectations. Gartner recommends diligent review of SAP's Hana 2.0 release and the future product roadmap to determine if SAP's vision for the market matches an organization's needs.

Snowflake Computing

Snowflake Computing (<https://www.snowflake.net/>) offers a fully managed data warehouse-as-a-service on AWS infrastructure. Snowflake supports ACID-compliant relational processing as well as native support for document store formats such as JSON, Avro and XML. A native Spark connector, R integration, support for user-defined functions (UDFs), dynamic elasticity, and temporal support round out the core capabilities. Snowflake is only available in the AWS cloud.

STRENGTHS

Built for the cloud: Snowflake's architecture is built to support the separation of resources from the ground up, with independently provisioned and scaled storage, compute and services tiers. This approach provides high levels of flexibility, scalability and automation capabilities that resonate with customers. Notably, 90% of Snowflake's customer references report it to be the designated DMSA technology standard within their organization.

Diverse data management capabilities: Snowflake's native support for document store formats alongside relational data (multimodel) provides the capabilities to consolidate different silos and repositories of data in a single system, allowing analytics on both relational and nonrelational data.

Customer experience, performance and availability: Snowflake scored near top marks with its reference customers on the experience of doing business with it. Additionally, more than 90% of its references indicated their intent to purchase more services from Snowflake within the next 12 months. Snowflake also received very high marks for system availability and performance.

CAUTIONS

Cloud-only vendor: As a cloud-only vendor, Snowflake will struggle to play in use-case-specific hybrid cloud architectures that require compatibility between on-premises and cloud-based deployments.

Pricing: Surprisingly, pricing was a significant factor in those cases where Snowflake had been considered, but was not selected. Snowflake scored slightly above average with its reference customers on the value for money it provided, and above average for the suitability of its pricing methods. Recently announced pricing initiatives in October and November 2016 are aimed at reducing the overall cost to customers and include an 80% reduction (overall) in storage costs (following AWS's storage price reductions), which should help to alleviate these concerns in the future.

Traditional and operational use-case focus: Snowflake's core strengths lie in the traditional and operational data warehouse use cases; fewer than half of its reference customers deploy Snowflake in a logical or context-independent data warehouse capacity. While Snowflake's strong multimodel capabilities (particularly around document data) make it well-suited for use cases requiring analysis of both structured and semistructured data, initial customer adoption has focused on more traditional approaches.

Teradata

Teradata (<http://www.teradata.com/>) 's offerings include a DBMS solution, data warehouse appliances and a cloud data warehouse solution (all MPP) — both on its own managed cloud and on public cloud provider infrastructure such as AWS and Azure. Support for the LDW comes with its Unified Data Architecture (UDA). Teradata QueryGrid (part of the UDA) provides multisystem query support via Teradata's own software, as well as open-source Presto. Teradata also offers Aster Analytics and Hadoop support via all three major distributions, as well as analytic consulting services.

STRENGTHS

Technical vision and capabilities: Teradata has built a DMSA platform that addresses all use cases: traditional, operational, context-independent and logical. Teradata's ability to integrate with multiple data sources (including Hadoop and streaming data via Teradata

Listener), to provide a unified query interface (via QueryGrid and Presto), and to deploy in multiple environments and form factors (including appliances, software and the cloud) demonstrate its market leadership in product capabilities and a strong vision for the future.

Market presence: Nearly 80% of Teradata's reference customers consider it to be the standard for DMSA within their organizations. Even when Teradata was not selected, it is almost always considered as part of vendor selection evaluations.

Performance: Teradata's reference customer scores for performance, across a broad range of use cases, were among the best of all the vendors in this Magic Quadrant.

CAUTIONS

Business model transition: Teradata's traditional appliance business continues to be challenged by cloud and software-only approaches. Teradata is addressing this shift by embracing these new approaches, but they represent a fundamental shift in its business model and its recent financial performance is suffering as a result.

Threat from incumbent operational DBMS vendors: While Teradata's technology is some of the most technically advanced and capable in the DMSA market, many potential customers may gravitate toward their existing incumbent DBMS (already in use for operational use cases) rather than diversifying to a heterogeneous environment. The advanced capabilities of hardware and software support the expanding footprint of incumbent vendors for an increasing number of use cases, rendering Teradata's optimized stack approach most appropriate for the most demanding analytical use cases.

Pricing and customer engagement concerns: Teradata is still viewed as an expensive, high-end offering. Those reference customers not selecting it most frequently cited price as the primary reason. Also, less than half of Teradata's reference customers indicated their intent to purchase additional licenses or product within the next 12 months. Teradata's recently announced move from perpetual licenses to subscription-based licenses should help address these concerns.

Transwarp Technology

Transwarp (<http://www.transwarp.cn/?lang=en>) offers the Transwarp Data Hub (TDH), a full suite of Hadoop distribution components that is supplemented by its SQL engine, machine learning, NoSQL search engine and stream processing. TDH is available on Microsoft Azure (China) and Ucloud. Transwarp also offers TDH as an appliance called Transwarp TxData Appliance.

STRENGTHS

Vertical industry focus: Transwarp is establishing a presence in the public sector and in the transportation and logistics and financial services industries. While many competing suppliers also have a presence in these vertical spaces, transportation and logistics is especially relevant in a geographically diverse market area, and even more so when there is a growing emphasis on the technology of smart cities — as is present in China.

Ease of implementation, support and value: Reference customers for Transwarp gave it an above-average rating for ease of implementation, professional services, support and overall value achieved from the product. They also reported a low rate of bugs or functionality gaps

compared with other suppliers in this market.

Solution architecture: Although the presence of a change data capture capability is more attuned to data integration markets, it is a highly applicable concept when utilizing Hadoop for analytics and transactions. A combination of containers, scheduler/workload management (for memory, storage and even a VLAN manager) and system services that manage autoscaling and orchestration (load balancing) round out Transwarp's solution architecture.

CAUTIONS

Broader technology expectations: In 2015, Transwarp was rated highly on its vision for including ACID transactional capabilities and stored procedures to its Hadoop distribution. However, the DMSA market is composed of a broad range of solutions well beyond Hadoop distributions. Most traditional DBMS vendors can now support multiple processing engines, file types, multimodel or polyglot, and more – and include Hadoop capabilities and compatibility. Transwarp continues to add new features in 2016, including an event-driven streaming engine and a search engine with ANSI SQL interfaces.

Competition from established vendors: Transwarp no longer enjoys singular status in its home market. In 2016, Huawei – which has an established presence in China – brought its own Hadoop offering to the market. Although Transwarp has continued to secure funding (in 2016), it now needs to prepare for these larger competitors.

Documentation and divergence from open source: The combination of a lack of documentation, fast releases and divergence from open source is creating some difficulties for implementers of Transwarp's solutions. Reference customers report that an absence of adequate documentation is a major issue, with neither the documents nor the customers able to adapt to the release cycles (sometimes only two months apart). Transwarp references also report that some of the functional additions are proprietary in nature, which confounds the overall desire to leverage open source.

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

EnterpriseDB

Google

Huawei

Snowflake Computing

Dropped

Action is no longer actively engaged in data management solutions for the analytics market.

Exasol did not meet the revenue inclusion criteria.

Hitachi did not demonstrate production customers from at least two distinct geographic regions.

Kognitio did not meet the revenue inclusion criteria.

Infobright did not meet the revenue inclusion criteria.

Inclusion and Exclusion Criteria

To be included in this Magic Quadrant, vendors had to meet the following criteria:

Vendors must have DMSA software generally available for licensing or supported for download for approximately one year (since 1 December 2015). We do not consider beta releases.

We use the most recent release of the software to evaluate each vendor's current technical capabilities. For existing solutions, and direct vendor customer references and reference survey responses, all versions currently used in production were considered. For older versions, we considered whether later releases may have addressed reported issues, but also the rate at which customers refuse to move to newer versions.

Product evaluations included technical capabilities, features and functionality present in the product or supported for download on 1 December 2016. Capabilities, product features or functionality released after this date could be included at Gartner's discretion and in a manner Gartner deemed appropriate to ensure the quality of our research on behalf of our nonvendor clients. We also considered how such later releases might reasonably impact the end-user experience.

Vendors have to provide 20 verifiable production implementations that will exhibit the revenue generated from 20 distinct organizations with DMSA, indicating they are in production, and:

A minimum of \$10 million in revenue with a 50% growth rate year over year.

Or more than \$40 million in revenue. Revenue can be from licenses, support and/or maintenance.

The production customer base must include customers from three or more vertical industries (see Note 3).

Customers in production must have deployed data management solutions for analytics that integrate data from at least two operational source systems for more than one end-user community (such as separate business lines or differing levels of analytics).

Vendors must demonstrate production customers from at least two distinct geographic regions (see Note 4).

To be included, any acquired product must have been acquired and offered by the acquiring vendor as of 30 June 2016. Acquisitions after 30 June 2016 are considered legacy offerings and are represented by a separate dot until publication of the following year's Magic Quadrant.

Support for the included data management for analytics products had to be available from the vendor. We also considered products from vendors that control or contribute specific technology components to the engineering of open-source DBMSs and their support.

We included in our assessments the capability of vendors to coordinate data management and processing from additional sources beyond the evaluated DMSA. However, vendors in this Magic Quadrant need to offer the ability to manage physical persistence of the data.

Vendors must provide support for at least one of the four major use cases (see Note 2).

Vendors must at least provide relational processing. Depth of processing capabilities and variety of analytical processing options are considered as advantageous in the evaluation criteria.

Vendors participating in the DMSA market had to demonstrate their ability to deliver the necessary services to support a data warehouse through the establishment and delivery of support processes, professional services and/or committed resources and budget.

Products that exclusively support an integrated front-end tool that reads only from the paired data management system did not qualify for assessment in this Magic Quadrant.

We also considered the following factors when deciding whether products were eligible for inclusion:

- Relational DBMSs.

- Nonrelational DBMSs.

- Hadoop distributions.

- No specific rating advantage was given to the type of data store used (for example, relational DBMS, graph DBMS, HDFS, key-value DBMS, document DBMS, wide-column DBMS).

- Multiple solutions used in combination to form a DMSA were considered valid, but each solution must demonstrate maturity and customer adoption.

- Cloud solutions were considered viable alternatives to on-premises solutions. The ability to manage hybrid on-premises and cloud solutions is considered advantageous for inclusion.

- Open-source solutions

Gartner may include, at its discretion, additional vendors in cases of known market adoption for classified, but unspecified, cases

The following technology categories are excluded:

BI and analytical solutions that only offer a DMSA that is embedded or that embed a DMSA from another provider.

BI and analytical solutions that only offer a DMSA that is limited specifically to the vendor's own BI and analytical solution or whose customers only use the solution within the same vendor stack.

In-memory data grids.

Prerelational DBMSs.

Object-oriented DBMSs.

Gartner analysts are the sole arbiters of which vendors and products are included in this Magic Quadrant.

Other Vendors to Consider

Gartner's Magic Quadrant process (see Note 5) involves research on a wider range of vendors than appears in the published document. In addition to the vendors featured in this Magic Quadrant, Gartner clients sometimes consider the following vendors when their specific capabilities match the deployment needs (this list also includes recent market entrants with relevant capabilities). These vendors were not included in the Magic Quadrant, because they either failed to meet the definition or one of the inclusion criteria, but they can be valid alternatives to the featured vendors. Unless otherwise noted, the information provided on these vendors derives from responses to Gartner's initial RFI for this document or from reference survey respondents. The following list is not intended to be comprehensive:

Exasol offers an in-memory column-store MPP DBMS, which is available as a free single-node edition, a clustered solution and an appliance as well as a plug-in for Tableau. It is also offered as a fully managed solution on EXACloud and on third-party cloud service providers such as AWS, Microsoft (Azure) and Rackspace. An Exasol cluster can have nodes on-premises and in the cloud. Known for its strong performance, Exasol has demonstrated traction in the market for the traditional data warehouse use case, but also for the context-independent data warehouse use case. However, while Exasol has demonstrated strong growth and adoption across several industries and geographies, it failed to meet the 2017 inclusion criteria based on revenue, which was higher than in previous years.

Infobright offers a column-vectored, highly compressed DBMS under a MySQL-based or PostgreSQL-based API layer. It markets the commercial Infobright Enterprise Edition (IEE), for which there is a trial download. Infobright targets analytics for machine-generated data and the IoT. Additionally, the newest offering, Infobright Approximate Query (IAQ), uses a statistical approach to generating high-value approximations quickly for complex queries over large datasets. Infobright was not included in this year's Magic Quadrant because it did not meet the revenue inclusion criteria of more than \$40 million in annual revenue or 50% year-over-year growth.

Kognitio offers the Kognitio Analytical Platform as a software data warehouse DBMS engine. The Kognitio Analytical Platform is an in-memory DBMS engine with nonrelational JSON support. In 2016, Kognitio (as a software technology company) reorganized, deciding to concentrate on a software-only product and dropping its hardware integration and appliances. Kognitio on Hadoop provides integration with Apache Hadoop, running the Kognitio Analytical Platform on Hadoop using Apache Yarn for management. Kognitio is primarily a U.K.-based company with offices in Chicago, Illinois, U.S. Kognitio was not included in this year's Magic Quadrant, because it failed to meet the revenue inclusion criteria.

LexisNexis Risk Solutions . High-Performance Computing Cluster (HPCC) Systems is an open-source computing platform for big data processing and analytics. HPCC Systems utilizes LexisNexis Risk Solution's data-centric Enterprise Control Language (ECL) to simultaneously query data and support that query with integrated/embedded analytics. HPCC Systems can scale across very large datasets and grid computing clusters. A two-stage process scale-out as an MPP environment spans across commodity-class hardware/servers. HPCC Systems is offered primarily as a data management analytics solution "as a service." The company is based in Alpharetta, Georgia, U.S. and is owned by the RELX Group. LexisNexis Risk Solutions is not included in this edition of the Magic Quadrant because it did not have the required number of production, on-premises analytics customers.

Qubole offers its Qubole Data Service, a cloud Hadoop processing service that enables distributed processing of queries across data persisted in the cloud or in databases accessible through dedicated connectors. Data of various formats can be processed through Hadoop, Spark, Hive, Presto and Pig engines. Qubole is available on multiple cloud service platforms such as AWS, Microsoft Azure, Oracle Bare Metal Cloud and Google — thereby giving its users the option of several cloud service providers. Data is read by Qubole Data Service for processing. Apache Hadoop Yarn is used for managing all resources and jobs across the cluster. In addition, it offers autoscaling that uses cloud APIs to dynamically increase and decrease processing capacity in order to guarantee availability and meet SLAs. Qubole is designed to enable the use of data through various processing engines such as SQL, MapReduce, Hive, Spark, HBase or Pig. Qubole offers an interesting approach to the LDW — by offering a Hadoop processing tier on top of data that is not under its management, such as native cloud object storage — and as a result did not meet the DMSA definition for this Magic Quadrant.

Evaluation Criteria

Ability to Execute

Product or Service: This criterion represents increasingly divergent market demands — ongoing traditional, logical data warehousing, operational data warehousing and context-independent data management for analytics. The largest and most traditional portion of the analytics and data warehouse market is still dominated by the demand to support relational analytical queries over normalized and dimensional models (including simple trend lines

through complex dimensional models). Data management for analytics' solutions are increasingly expected to include repositories, semantic data access (such as federation/virtualization) and distributed processing, in combination — referred to in the market as LDWs (see Note 1). All traditional demands of the data warehouse remain. Operational data warehouse use cases also exhibit traditional requirements, plus loading streaming data, real-time data loading and real-time analytics support. Users expect solutions to become self-tuning and to reduce staffing required to optimize the data warehouse, especially as mixed workloads increase. Context-independent warehouses (CIWs) do not necessarily support mixed workloads (though they can), nor do they require the same level of mission-critical support. CIWs serve more in the role of data discovery support or "sandboxes." CIWs are expected to meet the demands of ad hoc queries and varied processing options such as Python, machine learning (ML), R or graph.

Overall Viability: This criterion includes corporate aspects, such as the skills of the personnel, financial stability, R&D investment, the overall management of an organization and the expected persistence of a technology during merger and acquisition activity. It also covers the company's ability to survive market difficulties (crucial for long-term survival). Vendors are further evaluated on their ability to establish dominance in meeting one or more discrete market demands.

Sales Execution/Pricing: For this criterion we examine the price/performance and pricing models of the DBMS, and the ability of the sales force to manage accounts (judged by the feedback from our clients and feedback collected through the reference survey). We consider the market share of the DBMS software. Also included is the diversity and innovative nature of a vendor's packaging and pricing models, including the ability to promote, sell and support the product within target markets and around the world. Aspects such as vertical-market sales teams and specific vertical-market solutions are considered for this criterion.

Market Responsiveness/Record: This criterion is based upon the concept that market demands change over time and track records are established over the lifetime of a provider. The availability of new products, services or licensing (in response to more recent market demands), and the ability to recognize meaningful trends early in the adoption cycle, are particularly important. The diversity of delivery models as demanded by the market is also considered an important part of this criterion (for example, its ability to offer dbPaaS, software solutions, data warehouse "as a service" offerings or certified configurations).

Marketing Execution: This criterion includes the ability to generate and develop leads, channel development (through internet-enabled trial software delivery), and partnering agreements (including co-seller, co-marketing and co-lead management arrangements). We also considered the vendor's coordination and delivery of education and marketing events throughout the world and across vertical markets, as well as increasing or decreasing participation in competitive situations. This year, events and education are part of marketing execution.

Customer Experience: Evaluation of this criterion is based on customer reference surveys and discussions with users of Gartner's inquiry service during the previous six quarters. We also considered the vendor's track record on proofs of concept, customers' perceptions of the

product, and customers' loyalty to the vendor (this reflects their tolerance of its practices and can indicate their level of satisfaction). This criterion is sensitive to year-over-year fluctuations, based on customer experience surveys. Customer input regarding the application of products to limited use cases can be significant, depending on the success or failure of the vendor's approach in the market.

Operations: This criterion evaluates the alignment of the vendor's operations, as well as whether and how this enhances its ability to deliver. This criterion considers a vendor's ability to support clients throughout the world, around the clock and in many languages. Anticipation of regional and global economic conditions is also considered.

Table 1. Ability to Execute Evaluation Criteria

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

Source: Gartner (February 2017)

Completeness of Vision

Market Understanding: This criterion evaluates the vendor's ability to understand the market and shape its growth and vision. In addition to examining a vendor's core competencies in this market, we consider its awareness of new trends such as the increased demand from end users for mixed data management and access strategies, which matches the growing variety of skills and roles and the changing concept of the data warehouse and analytics data management; or, the value and position regarding emerging terminology such as data lakes or multimodel. Understanding the different audiences for various categories of data and associated SLAs (compromise, contender and candidate; see Note 6) is crucial, as is a demonstrable track record for altering strategy and tactical delivery in response to both opportunistic segments in the market and the broader market trends.

Marketing Strategy: This criterion evaluates a vendor's marketing messages, product focus, and ability to choose appropriate target markets and third-party software vendor partnerships in order to enhance the marketability of its products. This criterion includes the vendor's responses to the market trends identified above and any offers of alternative solutions in its marketing materials and plans. Investor relations is becoming an important part of marketing strategy; not investor sentiment, which can run contrary to vendor fiscal health, but vendor management and response to that sentiment.

Sales Strategy: Evaluation of this criterion encompasses all plans to develop or expand channels and partnerships that assist with selling, and is especially important for younger organizations because it can enable them to greatly increase their market presence while maintaining lower sales costs (for example, through co-selling or joint advertising). This criterion also covers a vendor's ability to communicate its vision to its field organization and, therefore, to its clients and prospective customers. Pricing innovations and strategies — such as new licensing arrangements, in particular in support of cloud and cloud/on-premises combinations, and the availability of freeware and trial software — are also included in this criterion.

Offering (Product) Strategy: When viewed from a vision perspective, this criterion is clearly distinguished from product execution. We evaluate the roadmap for enhancing capabilities across all four use cases (see Note 2). This also includes expected functionality and a timetable for introducing new market demands that will specifically include, but is not limited to, roadmaps and development plans for:

- Supporting a varied level of data latency with a growing focus on streaming and continuous data ingestion and access for analysis.

- Semantic design tier and metadata management capabilities.

- System and solution auditing and health management to ensure use-case SLA compliance

- Static and dynamic cost-based optimization, with the potential to span processing environments, data structures and storage options.

- Management and orchestration of multiple processing engines

- Elastic workload management and process distribution across cloud, and cloud and on-premises hybrids, including the separation of processing and storage

- Supporting a best-fit-engineering approach for their DMSA implementation. This requires vendors to support an open strategy (that is, allowing easy combining of technologies from heterogeneous vendors) as an alternative to an integrated strategy (that is, demonstrating value from the vendor stack integration), or both in combination.

Business Model: This criterion evaluates how a vendor's model addresses a target market with its products and pricing, and whether the vendor can generate profits with this model (judging by its packaging and offerings). We consider reviews of publicly announced earnings and forward-looking statements relating to an intended market focus. For private companies

(and also to augment publicly available information), we use proxies for earnings and new customer growth — such as the number of Gartner clients indicating interest in, or awareness of, a vendor's products during calls to our inquiry service.

Vertical/Industry Strategy: This criterion evaluates the vendor's ability to understand its clients. A measurable level of influence within end-user communities and certification by vertical industry standards bodies are of importance here. A specific product or solution roadmap to support a targeted vertical is considered to be a successful focus.

Innovation: Vendors demonstrate ability in this criterion by developing new functionality, allocating R&D spending and leading the market in new directions. This criterion also covers a vendor's ability to innovate and develop new functionality for accomplishing data management for analytics. Also addressed is the maturation of alternative delivery methods such as cloud infrastructures, as well as solutions for hybrid on-premises and cloud, and cloud-to-cloud, data management support. The vendor's awareness of new methodologies and delivery trends is also considered. Organizations are increasingly demanding data storage strategies that balance cost with performance optimization, so solutions that offer separation of compute and storage in a cloud environment, or that address aging and temperature of data, will become increasingly important.

Geographic Strategy: This criterion considers the vendor's ability to address customer demands in different global regions using direct/internal resources or in combination with subsidiaries and partners. We also evaluate a vendor's global reach and roadmap for addressing specific geographic regulatory requirements, particularly for cloud deployments. A specific product or solution roadmap to support a targeted geographic region is considered to be a successful focus.

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
Vertical/Industry Strategy	Low
Innovation	High
Geographic Strategy	Medium

Geographic Strategy	Medium
---------------------	--------

Source: Gartner (February 2017)

Quadrant Descriptions

Leaders

The Leaders quadrant includes five traditional large vendors that have had to adapt to this rapidly changing market. Notably, all have decided to pursue all four use cases for data warehousing with at least an average maturity of execution and vision. The span of ratings for Leaders' Completeness of Vision in this Magic Quadrant is quite narrow, although each vendor has a distinctively different vision. This year the push for cloud has affected the relative ratings among the Leaders and has also led to the inclusion of market disruptor AWS.

The data warehouse is usually the largest data management system in most organizations and the market for them is therefore large in terms of revenue, trained professionals and the variety of data management solutions (ranging from simple to complex). However, while data warehousing continues to be a major use case it has declined in importance, forcing the leading vendors to address new trends – such as data lakes and context-independent data warehouses – focusing on data science uses cases.

Challengers

This year, Challengers have been focusing on execution over vision. Customer satisfaction, value for money and the ability to address unique demands have all affected the execution ratings of these vendors.

For the 2017 Challengers, the key to success will be to continue to invest in great execution and to differentiate based on what has made success possible thus far or has enabled their progress in Completeness of Vision. Progressing on both fronts at the same time can prove challenging for smaller vendors; less so for the large ones.

Visionaries

In 2017, the Magic Quadrant has a single Visionary. This is the result of demonstrating a unique vision for this market; that is, to act as the semantic reconciliation tier between sources and various processing engines. It is too early to determine if this vision for the market will be the right one, but in 2017 it is unique.

Niche Players

In 2017, we see a growing number of Niche Players. This year the Magic Quadrant has seen a "zoom in" effect, resulting from the broadening of the market definition combined with a higher bar for inclusion in terms of revenue and geographic and industry adoption. We are also seeing a fast maturation of the market's interest in Hadoop distributions and new competing approaches for data lakes (such as storage on AWS or Azure). Hadoop distributions are no longer seen as a mandatory piece of an LDW; moreover, the push to the cloud has further

challenged Hadoop distribution vendors that were mainly designed for on-premises deployments. Finally, the remaining vendors in this quadrant are either new to the market this year or address a very narrow use case.

Context

Although the Leaders quadrant this year is largely populated with large traditional vendors that are relatively close to each other, we also see a new entrant in AWS. AWS has continued to focus on market execution, while also progressing its vision in addressing a broader set of use cases by combining and delivering multiple services to the market. The DMSA market has continued to attract new vendors, which appear in the Challengers and Niche Players quadrants.

This market therefore remains in a state of significant flux; disruption is likely to continue throughout 2017 and into 2018, which will erode the installed bases of the large traditional vendors. When such disruption occurs, the entire market usually moves away from a single mature trajectory and splits in two in terms of vision and execution. The result is that during the next two years (until the end of 2018) this market is likely to be much more volatile, with changes in leadership a possibility. For example, in 2017 the market is challenging the vision of Hadoop distribution vendors becoming the platform for DMSA; new and yet unproven visions (such as MarkLogic's) are emerging. In the coming years, changes in mix of use cases, deployment options and DMSA capabilities are likely to continue to disrupt the vision aspect of this Magic Quadrant.

The race is on to deliver cloud solutions and hybrid cloud and on-premises offerings. Specialized cloud vendors and all the traditional vendors have already started working to deliver these offerings. In parallel, we are seeing appliances that have been reliable sellers for many vendors losing some of their appeal and becoming niche offerings.

This Magic Quadrant has a lot of white space in the upper-right corner, indicating that the market continues to demand more innovation and better execution to address the needs of combined cloud and on-premises deployments, as well as cloud and big data combinations. New demands for separation of storage and compute have also emerged.

Vendors are continuing to innovate and support the new use cases demanded by the IoT and digital business. For these, timeliness of access and the analysis of data become more important than the volume and variety aspects of big data.

Market Overview

The DMSA market continues to evolve. Customers now expect solutions that support all types of data for analytics and that take a coordinated approach. This demands different types of integrated solution and an interoperable services tier for managing and delivering data. Data lakes and the ability to manage streaming data are now being pursued by a growing number of organizations.

Data and analytics leaders must be aware of the market's evolution and prepare hybrid technology platforms that expand the data warehouse beyond any current practice. This is especially important because the influence of the LDW has created a situation in which multiple repository strategies are now expected, even from a single vendor. Interest is also growing in cloud solutions, as alternatives to on-premises solutions, although we expect hybrid cloud and on-premises deployments to become the norm.

Gartner notes the following key trends in this market:

The definition of the data warehouse is expanding. The term "data warehouse" does not mean "relational, integrated repository." The data warehouse is what we built to do that, but the new SLAs indicate that sometimes data should be integrated, and sometimes not. This new market demands a much broader data management solution for analytics. This is best explained by comparing two guiding architectural approaches (see "The Data Warehouse DBMS Market's 'Big' Shift").

Enterprise data warehouse (EDW): An integrated, subject-oriented, time-variant and physically centralized data management system mounted on hardware that is optimized for mixed workload management and large-query processing.

Logical data warehouse (LDW): An optimized combination of software and hardware that delivers a logically consistent, subject-oriented integration of time-variant data that is accessed via a centralized data management infrastructure. It uses a combination of repositories, virtualization and distributed processes. The LDW is part of a larger movement to establish a wider market for DMSAs.

The concept of the LDW emerged as the first practical architecture for the newly emerging analytic data management requirements. The LDW will continue to grow in popularity during the next five years. The terminology used about LDWs will become the de facto vocabulary for describing how to evolve a traditional data warehouse into a broader DMSA.

The push for the cloud. More organizations are considering cloud IaaS or PaaS as the means of deploying their analytical environments. Although appealing in terms of the flexibility and agility of deployment and pricing, this approach will demand further support for hybrid on-premises and cloud options. This will set new expectations for the LDW, requiring new ways of managing access and processing needs across these hybrid environments. Cloud solutions are also challenging the traditional positioning of appliances, which is causing further disruption for traditional data warehouse vendors.

The role of data lakes. "Big data" was a term that served as a catalyst for change in the data warehouse environment. Implementers have identified three highly useful patterns for big data in analytics: data exploration/data science "sandboxes"; offloading of history from the warehouse; a staging area for all data prior to loading into the data warehouse as needed. In 2015, we saw the emergence of data lakes as a popular approach for addressing the three use cases and for extending the role of the data warehouse. Successful organizations pursuing the use of data lakes as part of their LDW implementation are taking a best-of-breed (BOB) approach, because no single product is a complete solution. Organizations are

even employing multiple products from a single vendor, which is their interpretation of a BOB implementation within the vendor stack. However, they are now seeking approaches to integrate with these new, very large data management and analytic silos, because they understand that new user populations — such as business analysts, data scientists and data engineers — need to have access to all data (see Note 7).

The emergence of best-fit engineering. The BOB deployment model includes a combination of different software (proprietary license and open-source license), file management systems, communication and semantic middleware, and variable hardware/network components. Generally, BOB has meant acquiring leading technologies in several areas and then hiring expert implementers to accomplish the deployment. However, BOB is being replaced by a concept Gartner calls "best-fit engineering" (see "The Data Warehouse DBMS Market's 'Big' Shift"). The difference between the two is that under BOB, implementers select the best solution for part of their architecture and then reuse it in secondary roles that emerge — even if such secondary functionality is less than optimal. In best-fit engineering, the least-required technology for each function is considered first. For example, it is possible to use a DBMS to facilitate access to external files and tables under BOB, but by adding a different technology to the stack that is specifically focused on data virtualization (and therefore has different, and possibly superior, optimization capabilities), it becomes best-fit engineering. Each technology is used for its most appropriate purpose and is therefore much more likely to exhibit a low cost for a precise need. In 2017, best-fit engineering approaches are pushing for new implementations of data lakes where data is just stored in a cloud object store. Bringing data to the processing tier is driven by the analysis requirements. The concept of the separation of storage and processing is not only a means to optimize cloud spending, but also about allowing multiple processing tiers to run on top of multiple persistency tiers.

Acronym Key and Glossary Terms

AWS	Amazon Web Services
BI	business intelligence
BOB	best-of-breed
DBA	database administrator
dbPaaS	database platform as a service
DMSA	data management solution for analytics
DR	disaster recovery
EDW	enterprise data warehouse
ELT	extraction, loading and transformation

HA	high availability
HDFS	Hadoop Distributed File System
IaaS	infrastructure as a service
IoT	Internet of Things
IP	intellectual property
JDBC	Java Database Connectivity
LDW	logical data warehouse
MPP	massively parallel processing
PaaS	platform as a service
S3	Simple Storage Service (Amazon)
SMP	symmetric multiprocessing
SSED	source-system extracted data

Note 1

Logical Data Warehouse Definition

The LDW is a new data management architecture for analytics that combines the strengths of traditional repository warehouses with alternative data management and access strategies. It has seven major components:

- Repository management
- Data virtualization
- Distributed processes
- SLA management
- Auditing statistics and performance evaluation services
- Taxonomy and ontology resolution
- Metadata management

Note 2

Use Cases

Traditional Data Warehouse: This use case involves managing historical data coming from various structured sources. Data is mainly loaded through bulk and batch loading.

• The traditional data warehouse use case can manage large volumes of data (see Note 8) and is primarily used for standard reporting and dashboarding. To a lesser extent, it is used for free-form querying and mining, or operational queries. It requires high levels of capability for system availability and administration and management, given the mixed workload capabilities for queries and user skills breakdown.

Operational Data Warehouse: This use case manages structured data that is loaded continuously in support of embedded analytics in applications, real-time data warehousing, and operational data stores.

• This use case primarily supports reporting and automated queries in support of operational needs, and will require HA/DR capabilities to meet those operational needs. Managing different types of users or workloads, such as ad hoc querying and mining, will be of less importance, because the major driver is to meet operational excellence.

Logical Data Warehouse: This use case manages data variety and volume for both structured and other content data types.

• Besides structured data coming from transactional applications, this use case includes other content data types such as machine data, text documents, images and video. Because additional content types can drive large data volumes, managing large volumes is an important criterion. The LDW is also required to meet diverse query capabilities and support diverse user skills. This use case supports queries reaching into other sources than the data warehouse DBMS alone.

Context-Independent Data Warehouse: This declares new data values, variants of data form and new relationships. It supports search, graph and other advanced capabilities for discovering new information models.

• This use case is primarily used for free-form queries to support forecasting, predictive modeling or other mining styles, as well as queries supporting multiple data types and sources. It has no operational requirements and favors advanced users such as data scientists or business analysts, resulting in free-form queries across (potentially) multiple data types.

Note 3

Vertical Industry Sectors

Accommodation and Food Services

Administrative and Support and Waste Management and Remediation Services

Agriculture, Forestry, Fishing and Hunting

Arts, Entertainment, and Recreation

Construction

Educational Services

Finance and Insurance

Healthcare and Social Assistance

Information

Management of Companies and Enterprises

Manufacturing

Mining

Professional, Scientific and Technical Services

Public Administration

Real Estate Rental and Leasing

Retail Trade

Transportation and Warehousing

Utilities

Wholesale Trade

Note 4

Geographic Regions

North America (Canada and the U.S.)

Latin America (including México)

Europe (Western and Eastern Europe)

The Middle East and Africa (including North Africa)

Asia/Pacific (including Japan)

Note 5

Research Methodology for Magic Quadrant

Gartner uses multiple inputs to establish the positions and scoring of vendors in our Magic Quadrants. These are adjusted to account for maturity in a given market, market size and other factors. For this update of the Magic Quadrant, the following sources of information were used:

Original Gartner research, often utilizing our market share forecasts to establish the breadth and size of a market.

Publicly available data, such as earnings statements, partnership announcements, product announcements and published customer case studies.

Gartner inquiry data collected from more than 16,000 inquiries conducted by the authors and the wider analyst community within Gartner during the previous 20 months. These inquiries provide input on, for example, use cases, issues encountered, license and support pricing, and implementation plans.

RFI surveys issued to vendors, in which they were asked to provide specific details about versions, release dates, customer counts and distribution of customers worldwide, among other things. Vendors could refuse to provide any information in this survey, at their discretion.

Surveys of reference customers (with almost 300 new responses added this year to four prior years of survey data). Vendors were asked to identify a minimum number of reference customers. Gartner augmented the vendor-provided population by adding Gartner inquiry client contacts as potential respondents. Responses were voluntary for all participants. These surveys included questions to confirm that customers were current license holders. Additionally (especially in the case of open-source utilization), customers provided information that confirmed the size and scope of their implementations. Customers were also asked to provide information about issues and software bugs, overall and specific sentiments about their experience of a vendor, the use of other software tools in the environment, the types of data involved, and the rate of data refresh or load. They were also asked about their deployment plans. Historical survey responses were used to identify trends only. Current-year survey responses were used for the commentary in this Magic Quadrant.

Gartner customer engagements, in which we provided specific support, were aggregated and anonymized to add perspective to the other, more expansive research approaches.

It is important to note that this was qualitative research that formed a cumulative base on which to form the opinions expressed in this Magic Quadrant.

Note 6

Data Categories and Associated SLAs

Compromise. There is agreement that the model should be persisted for pervasive use by many end users, and that it exhibits a general tolerance for latency (even as low as two minutes). This SLA has two primary objectives: optimized performance and end-user comprehension of the data model. It is a compromise because the model is deployed to satisfy a low common denominator of use cases and specifically ignores exceptions. This is the traditional data warehousing approach, and is generally best thought of as "least common denominator" data for commonly shared analytics.

Contender. There is no agreement that any combination of data will be persisted or is widely applicable to use cases. The result is a sense of transient data combinations being used by a diverse set of users. However, the source of the information is generally agreed to be an adequate representation for exploring how the data can be used. Because of the changing and evolving nature of this SLA, zero-latency is often requested — but this is actually a proxy for seeing the data in as close to its native form as possible. This is data federation and is often supported by data virtualization software and even with multiple data marts. This approach is used when analysts have not reached any agreement about how to combine disparate data, but seek to combine it under multiple models.

Candidate. There is wide access to the asset, but, because the structure is complex and not always consistent, the default is to present multiple schema-on-read scenarios for different types of analyses. These scenarios explore unexpected forms in the data, but also create postulates of multiple alternative forms in parallel analytics from the data. This is big data analytics. A different barrier to entry exists here, in that users must understand how to parse and process data — almost like a DBMS. Candidates are suggested forms of reading the data and potential uses of how that data is read. As such, they are submitted for consideration — but are not even contenders yet.

Note 7

User Categories

Data scientist: Expertise in statistics, abstract mathematics, programming, business processes, communications and leadership.

Data miner: Expertise in subject areas of data; uses statistical software and statistical models; is fully aware of computer processing "traps" or errors.

Business analyst: Uses online analytical processing (OLAP) and dimensional tools to create new objects; has some difficulty with computer languages and computer-processing techniques.

Casual user: Regularly uses portals and prebuilt interfaces; has minimal, if any, ability to design dimensional analytics.

Note 8

Data Warehouse Data Volumes

The data warehouse volume managed in the DBMS can be of any size. For the purpose of measuring the size of a data warehouse database, we define data volume as source-system extracted data (SSED), excluding all data-warehouse-design-specific structures (such as indexes, cubes, stars and summary tables). SSED is the actual row/byte count of data extracted from all sources.

The sizing definitions of traditional warehouses are:

Small data warehouse — less than 5TB

Midsized data warehouse — 5TB to 40TB

Large data warehouse — more than 40TB

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.



(https://www.gartner.com/technology/contact/become-a-client.jsp?cm_sp=bac_-_reprint_-_banner)

© 2017 Gartner, Inc. and/or its affiliates. All rights reserved. Gartner is a registered trademark of Gartner, Inc. or its affiliates. This publication may not be reproduced or distributed in any form without Gartner's prior written permission. If you are authorized to access this publication, your use of it is subject to the Usage Guidelines for Gartner Services (/technology/about/policies/usage_guidelines.jsp) posted on gartner.com. The information contained in this publication has been obtained from sources believed to be reliable. Gartner disclaims all warranties as to the accuracy, completeness or adequacy of such information and shall have no liability for errors, omissions or inadequacies in such information. This publication consists of the opinions of Gartner's research organization and should not be construed as statements of fact. The opinions expressed herein are subject to change without notice. Gartner provides information technology research and advisory services to a wide range of technology consumers, manufacturers and sellers, and may have client relationships with, and derive revenues from, companies discussed herein. Although Gartner research may include a discussion of related legal issues, Gartner does not provide legal advice or services and its research should not be construed or used as such. Gartner is a public company, and its shareholders may include firms and funds that have financial interests in entities covered in Gartner research. Gartner's Board of Directors may include senior managers of these firms or funds. Gartner research is produced independently by its research organization without input or influence from these firms, funds or their managers. For further information on the independence and integrity of Gartner research, see "Guiding Principles on Independence and Objectivity." (/technology/about/ombudsman/omb_guide2.jsp)"

About (<http://www.gartner.com/technology/about.jsp>)

Careers (<http://www.gartner.com/technology/careers/>)

Newsroom (<http://www.gartner.com/newsroom/>)

Policies (http://www.gartner.com/technology/about/policies/guidelines_ov.jsp)

Privacy (<https://www.gartner.com/privacy>)

Site Index (<http://www.gartner.com/technology/site-index.jsp>)

IT Glossary (<http://www.gartner.com/it-glossary/>)

Contact Gartner (http://www.gartner.com/technology/contact/contact_gartner.jsp)