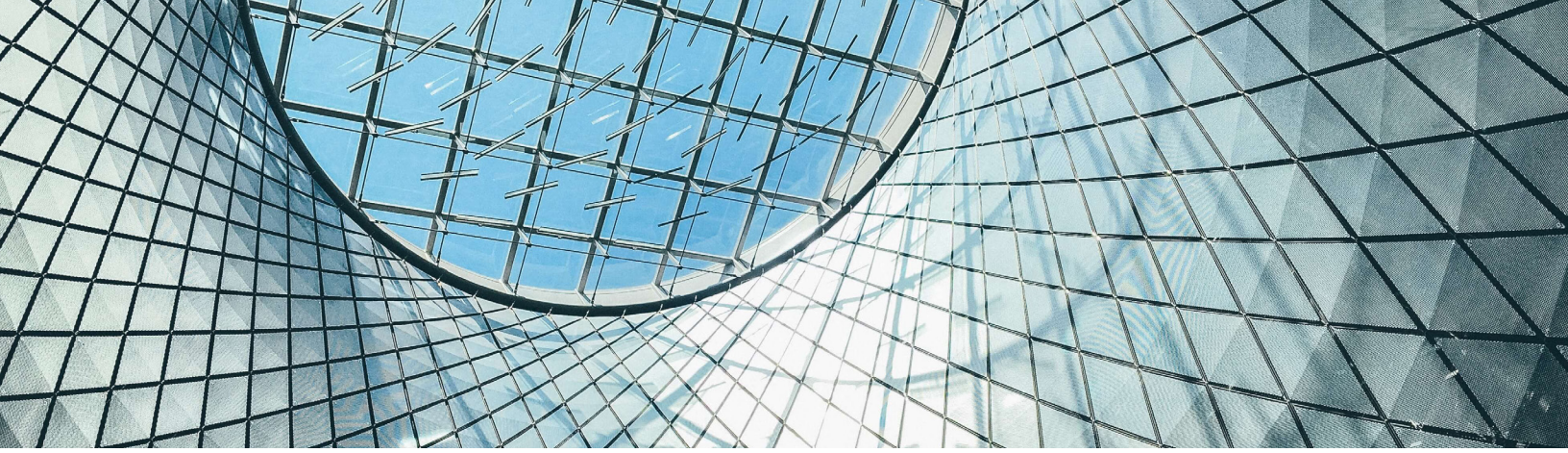

Market Insight Report Reprint

Coverage Initiation: Acceldata looks to accelerate observability of data, processing and pipelines

January 31 2022

by Paige Bartley

Data observability, borrowed from engineering methodology, has helped define a relatively nascent – and active – corner of the data reliability and quality market. Acceldata, applying these principles, is looking to expand the scope of monitoring capabilities to additional layers of the data and compute stack to provide 'multidimensional' visibility.

451 Research

S&P Global
Market Intelligence

Introduction

The engineering concept of observability is flourishing at the IT system interaction and cloud ecosystem level, as well as the data layer itself. While some may associate observability approaches with traditional system monitoring, the fact remains that engineering involvement in data pipelines and workflows is putting pressure on organizations to apply observability principles specifically to the data stack for use cases such as data quality.

Based on 451 Research's Voice of the Enterprise: Data & Analytics, Data Management & Analytics 2021 survey, 73% of respondents report their organization has a dedicated data engineering function. Acceldata is looking to appeal primarily to these enterprise personas with more methodological monitoring of the entire data stack, so that potential problems can be detected and resolved early. Perhaps in an effort to differentiate from technology providers that look to use data observability specifically for data quality use cases, Acceldata is trying to take a broader 'whole stack' approach for data: applying observability to track the enterprise data layer, associated compute processing and data pipelines.

THE TAKE

Engineering principles are increasingly being applied to persistent challenges in IT and data management, essentially looking to shift the detection of potential problems 'upstream' where issues can be addressed early by workers with technical skill sets. But the increasing complexity of the modern data stack complicates matters, and it can be difficult to get adequate visibility into associated data processing and pipelines. Acceldata is looking to expand the scope of data observability to address this complexity: providing detailed insight designed to optimize data systems.

A focused competitive landscape and associated buzz will be Acceldata's biggest potential challenge. There is no shortage of providers touting data observability capabilities, particularly for data quality use cases. Acceldata is perhaps taking a broader positioning approach than most in this market. It is looking at compute performance monitoring, data reliability and data pipeline visibility. Reaching the correct purchase influencers and stakeholders, of course, will be critical to the company's success, especially given that the platform supports multiple interdependent use cases.

Context

Acceldata was founded in 2018, with significant company leadership experience borrowed from stints at Cloudera and Hortonworks, among other firms such as Adobe, Citrix, Zalando and Zuora. The company's cofounder and CEO, Rohit Choudhary, came from this big-data management heritage, and was personally involved in Hadoop cluster deployments as early as 2009. Acceldata's collective leadership guided its modern hypothesis that if data fundamentally powers the modern enterprise, the underlying data 'power grid' needs to be both reliable and visible.

The company's mission is to offer what Acceldata describes as multidimensional data visibility and data observability, providing a unified data observability platform to track enterprise data, data processing and data pipelines. Of course, observability or monitoring are only as useful as the outcomes and actions that are taken as a result of such visibility. In this sense, Acceldata is largely trying to support business initiatives to improve the reliability, scale and cost of data operations.

To date, Acceldata has raised $45m in VC funding, with participation from firms such as Emergent Ventures, Insight Partners, Lightspeed, March Capital and Sorenson Capital. The company has about 140 employees today, representing approximately 10x team growth since March 2020, at the beginning of the pandemic. Acceldata is headquartered in San Jose, California, but has additional customer success, sales and R&D presence in India. The company also has a presence in Singapore and Thailand.

Acceldata's go-to-market strategy targets Fortune 2000 enterprise accounts with complex data environments. Prospect and customer organizations of Acceldata, as a rule, tend to already have established data engineering functions. The company today has customer accounts in 10 countries; however, it does not disclose exact customer count.

Platform

Acceldata originally went to market with three modular products (Pulse, Torchn and Flow), but the company is in the process of merging these offerings into a single platform, reflecting a trend in the broader data management market to bind capabilities into broader cohesive suites.

This platform will still be aligned with three core areas of observability and visibility: the enterprise data layer itself, data processing and compute, and data pipelines. By shifting data observability as far upstream as possible, the goal of the Acceldata platform is to ensure that low-quality data never enters the data lake – or similar architecture – to begin with. The platform aligns with three key areas:

- **Compute performance monitoring.** The platform has features for compute performance monitoring, with the objective of improving the reliability, scale and cost of data processing. Rather than simply focusing on incident detection and resolution measures, the capabilities are designed to help organizations shift efforts to the prediction and prevention of possible compute performance issues. The platform can provide root cause analysis with event correlation based on historical comparisons, environment health, and resource contention. Workload analysis helps organizations identify problems such as over-provisioned software, and rein in costs so that infrastructure expenses align with business objectives. Analysis via observability methodology, additionally, can help businesses identify processing bottlenecks and further craft execution plans to optimize code and queries.
- **Enterprise data layer reliability.** The platform also offers capabilities explicitly focused on the enterprise data layer, seeking to help ensure data reliability and data quality. Accommodating both data at rest and data in motion, the platform monitors enterprise data across data lakes, data warehouses and other data repositories to detect and resolve potential issues that may impact data reliability, from data ingestion to data consumption. It leverages both embedded AI and ML capabilities for automation, helping build taxonomies for sensitive and related data assets as well as reconcile data in the process of cloud migrations. Anomaly detection, schema detection and data drift detection are included in the platform's capabilities. Data reliability and quality assurance are common, though not exclusive, use cases for the product.
- **Data pipeline observability.** A platform capability, currently in beta offering, is geared toward providing observability specifically for data pipelines. Functionality has an emphasis on monitoring data pipelines across hybrid and interconnected architecture, accounting for the complexity seen today with modern data lake and data warehouse choices, which may be spread across cloud and on-premises environments. Data pipeline observability capabilities enable monitoring to compare to internal/external SLAs, as well as tracking of all transactions and data transformations. Monitoring of data schemas and distributions is also a feature. This aspect of the platform integrates with most common ETL tools out of the box, and helps data teams identify and remediate data pipeline issues quickly, before they spiral into downstream problems. Data pipeline observability is primarily geared toward data engineering personas and users.

Acceldata has cloud-native ambitions for its platform, with a product roadmap to release pure SaaS versions of products during the 2022 calendar year. Compatibility with various data processing environments and data repositories (HDFS, Kafka, Hive, Spark, AWS Lambda, Amazon EMR) is an ongoing effort for Acceldata. This is based not only on strict technology product or open source project popularity, but also on the prevalence of actual production systems in enterprise environments, which are Acceldata's target audience.

Exact pricing strategy for the company's platform approach was not revealed.

Competition

Perhaps unusual for the highly overlapping data management market landscape, data observability technology specialists carve out a relatively tight subsegment, often associated with data quality functionality. As we detailed, Acceldata's capabilities are broadly applicable across the data stack beyond basic data quality use cases, but the company's competitive reality puts it up against a cohort of vendors that often use data quality as a use case.

Notable examples would include Anomalo, Bigeye (formerly Toro), Databand, Datafold, Lightup, Monte Carlo and Soda. These providers are generally underscored by a common denominator of leveraging data observability technology, primarily to ensure data quality downstream for both technical and business users that depend on data access and leverage for insight.

Although associated more with systems and compute observability than granular data observability capabilities, Pepperdata has a similar 'full stack' positioning regarding its observability capabilities.

Other technology products that explicitly target enterprise data engineering functions may offer potential competitive overlap. While perhaps not a direct competitor, Trifacta (recently announced as acquired by Alteryx) offers the Trifacta Data Engineering Cloud. Pending completion of the Alteryx acquisition, the Trifacta Data Engineering Cloud will help provide a data engineering back end to the Alteryx data prep and analytics platform, helping support the data engineering function. Oracle, Qlik and Tableau (now owned by Salesforce) also have varying degrees of data engineering and data management support features for their analytics/visualization platforms.

Data observability's role in ensuring data quality use cases can't be ignored, and there is no lack of providers that overlap the data quality market. Examples of vendors with robust data quality functionality in the traditional sense include Atacamma, Atlan, Datactics, Experian Data Quality, Global IDs, Hitachi Vantara, IBM, Informatica, Precisely (formerly Syncsort), SAS Institute, Semarchy, Syniti, Talend, TIBCO and Zaloni.

SWOT Analysis

STRENGTHS

The company goes 'above and beyond' in terms of data observability, extending beyond simple data quality use cases to help ensure efficiency of the entire supporting data stack and compute resources. Many competitors are primarily concerned with detecting potential data problems early, whereas Acceldata is more actively trying to apply predictive and automated capabilities to ensure that anomalies never have the chance to spiral into problems at all.

WEAKNESSES

Acceldata might be trying to toss a wide net in terms of functionality, especially considering its primary target audience: the enterprise data engineering function. This function is still relatively nascent in many organizations, and may have varying purchasing influence. Acceldata, in advancing its platform approach, may need a stronger business narrative to communicate value to high-level decision makers in addition to more technical manager roles.

OPPORTUNITIES

The data engineering function, particularly within large and data-intensive organizations, is starting to flourish and is becoming well established in many cases. By appealing to the needs and technology preferences of these skilled personas, Acceldata has the chance to root itself as a preferred tool, as the data engineering function continues to gain clout.

THREATS

Many businesses strongly associate the concept of data observability with data quality, and Acceldata may have an uphill battle to educate and prove to its prospects that the technology is helpful for a much broader array of use cases. Since the data observability market landscape today is pretty tightly knit, pricing pressure for products may be an issue if prospects erroneously believe they can make apples-to-apples comparisons.

CONTACTS

The Americas
+1 877 863 1306
market.intelligence@spglobal.com

Europe, Middle East & Africa
+44 20 7176 1234
market.intelligence@spglobal.com

Asia-Pacific
+852 2533 3565
market.intelligence@spglobal.com

www.spglobal.com/marketintelligence