Coverage Initiation: Promethium reduces time to find insight with its Data Navigation System

Analysts - Matt Aslett, Matthew Utter

Publication date: Thursday, April 16 2020

Introduction

Driving transformational change at the enterprise level has become synonymous with an organization's ability to leverage data. Unfortunately, accomplishing this is no easy task as companies continue to struggle with data silos, bureaucracy and difficulty in establishing agreed-upon goals. To solve this issue, organizations have started to turn toward the DataOps methodology, improving data pipelines by bridging the gap between data consumers (e.g., data analysts, developers and senior decision-makers) and data operators (e.g., data management and IT professionals).

Promethium's take on DataOps delivers self-service data discovery, natural language processing and data virtualization, as well as automation capabilities that can help eliminate manual tasks encountered during data discovery and assembly. These efforts are aimed at helping enterprises quickly find answers to critical business questions, allowing them to respond to rapid market changes.

The 451 Take

Nearly half (47%) of respondents to 451 Research's Voice of the Enterprise: Data & Analytics 2H 2019 survey indicated that data analysts spend 50% or more of their working hours finding and preparing data for analysis. As such, more businesses are recognizing the need to alleviate pain points and speed up data processes. This has resulted in increased interest in faster data discovery and assembly. Promethium has positioned itself well, as its Data Navigation System (DNS) provides data analysts with AI-driven and automated self-service access to data without requiring data to be migrated to a new platform. Additionally, the vendor is promoting collaboration between data teams and business users via its DataDash application. This process, where multiple actors can actively work on a problem together, creates a positive feedback loop that helps organizations discover the answers they need faster than before. Although this is true, it will be interesting to see

how Promethium tackles the organizational and cultural aspects of DataOps, as these factors play a major part in the successful adoption of these technologies.

Context

Promethium was founded in 2018 and is headquartered in Menlo Park, California. The company is led by founder and CEO Kaycee Lai, who has a range of experience stemming from multiple roles at technology firms. Lai previously served as the president of Waterline Data, head of Virsto sales at VMware, and in positions at Microsoft, EMC and Delphix. VP of engineering Shuo Yang was previously the director of cloud architecture and automation at Electronic Arts and held roles at Huawei and Google. Currently, the vendor has about 20 employees.

The idea for Promethium stems from Lai's experience at Waterline Data observing that multiple enterprises were taking months to get the answers they needed for critical business decisions even after investing in data-cataloging software. Specifically, the company identified issues related to data discovery, validation and assembly that stemmed from multiple sources, vendors, locations and versions of data. This coupled with the reliance on data analysts with the skills required to write SQL queries created situations where answers came too late and had lost their importance to the business. Promethium's technology helps bridge these gaps using techniques that fall under the DataOps methodology.

At the beginning of 2020, Promethium raised $6m in a funding round led by .406 Ventures and Zetta Venture Partners, which had also led the firm's $2.5m seed round in late 2018. This latest financing brings the company's total raised to $8.5m, which it plans to invest in go-to-market strategy and product development. Its product is now generally available and being put through the paces via proof-of-concept and pilot deployments.

Technology and products

Promethium's offering is its Data Navigation System. DNS aims to solve different challenges during data discovery, validation and assembly to help enterprises find answers to critical questions faster. The product utilizes artificial intelligence (AI) and machine learning (ML) technologies, including natural language query and inference of metadata and lineage, and is designed to learn and adapt based on previous questions and queries. DNS is available from multiple sources, such as AWS Marketplace, Azure Marketplace or through channel partners, and it can be deployed as SaaS in the public cloud, in a virtual private cloud or on-premises.

For data discovery, DNS uses universal connectors that gather metadata across multiple data sources (all flavors of relational, NoSQL, HDFS and S3-based data sources are supported) or from data catalogs (Apache Atlas, Waterline Data, Collibra, Alation and Informatica). This is done via the deployment of smart bots, API-driven services that initiate connections with data sources. This approach is designed to avoid the need to move data, helping to avoid duplicates and inefficiencies. From here, DNS validates the data collected, usually multiple files and tables, by searching for duplicates and tracing the data's lineage – it also suggests new data that is required to answer the specific question at hand.

Continuing the process, DNS gives users step-by-step instructions to assemble the data and then automatically generates the appropriate SQL statement based on the natural language query. The statement is shown in a graphical format, allowing users to see the relationships that exist within the data and edit parts of the query if necessary. Diving deeper into this process, Promethium employs a federated query approach that leverages Starburst's distribution of Apache Presto to create a virtual view of data in multiple underlying data sources, which appears to BI tools (such as Tableau, Looker, Superset and ThoughtSpot) like a single database table that can be queried with the need for ETL processes.

Promethium strikes at the heart of DataOps not only through its automation of manual data management tasks but also via its ability to connect and streamline processes between different data users. In addition to the discovery, validation and assembly functionality aimed at data operators, DNS includes DataDash, a separate interface for business users that allows them to find, ask and view questions and answers about their company's data. When a user asks a question, DataDash is able to parse keywords and return the likelihood of an answer based on the metadata available. Using AI and ML, the application can discover the relationships between previous questions and use these insights to answer other related questions. Included in the application is a workflow management system where different business users can collaborate on the same questions and provide live feedback to each other.

Competition

As the demands of enterprises change, data management vendors are compiling different functionalities into one service. This phenomenon is occurring in multiple ways, with organizations focusing on internal development as well as finding synergies through acquisition and partnership. Two examples of the latter are Hitachi Vantara and Qlik, which compete with Promethium's main offering. Via its purchase of Waterline Data, Hitachi Vantara now supports the automated tagging of data and metadata to accelerate data discovery. Although Promethium views the Lumada Data Catalog as primarily complementary, Hitachi Vantara also offers capabilities for self-service data preparation and integration of data sources. Qlik similarly provides self-service capabilities to aid data users in their search for the best possible data for analytics input and leverage. The company is focused on machine-learning-driven insights and recommendations on top of its ability to query across multiple data stores.

Companies such as Io-Tahoe, Octopai and Alation will likely be compared with Promethium as they focus on metadata-driven data discovery as well. Although Io-Tahoe focuses mainly on identifying and classifying sensitive data, Octopai and Alation have capabilities that overlap with Promethium's main offering. Octopai can automate the extraction of metadata via the use of machine learning that can map, tag, correlate and identify relationships across multiple metadata sources. On top of this, users can search metadata indexes as well as access visual maps of relationships and dependencies to understand data lineage. Alation is following in the same vein with its focus on self-service enablement, targeting DataOps through the natural feedback loop that is created when those curating data and those leveraging it can easily collaborate.

Although Promethium is aiming to complement data catalog products, some aspects of its offering do share certain qualities with data catalog providers. For example, Promethium's DataDash is comparable to Alation's UI, where users can also find data employing simple language, create queries with the application's assistance, and collaborate with different data users.

Promethium's main offering utilizes Starburst's distribution of Apache Presto, which provides users with a query federation layer for querying data in multiple sources. As such, it's possible that some potential customers may be looking to create something similar deploying either Starburst Presto or Apache. However, employing Presto alone would require users to find data themselves and write their own SQL statements, and they would also be without Promethium's discovery, validation and natural language query capabilities. For the latter, ThoughtSpot could also be considered a rival but only in its visualization and natural language capabilities.

SWOT Analysis

Strengths	Weaknesses
Promethium has identified one of the major pain points when it comes to using data to drive business decisions. Due to its	The vendor is targeting functionality that falls under the DataOps methodology. To achieve all of the promised

smaller size, the company can adjust to the changing market landscape and adapt to shifting customer preferences.	benefits, Promethium may also have to help customers address the cultural and organizational aspects that help deliver the wider advantages that the technology promises.
Opportunities	**Threats**
More organizations are looking to accelerate the deployment of data-driven business decision-making. As these companies begin to leverage more of the data that is available to them, data access and preparation headaches become more obvious, potentially expanding the market for Promethium's offering.	When markets grow fast, competitors enter. The data catalog space is already overcrowded and an increasing number of vendors, both large and small, are recognizing the same market opportunity as Promethium, which could pose significant threats to the company in the future.

Source: 451 Research, LLC