Cambridge Semantics positions Anzo as the semantic layer for the enterprise data fabric

MATT ASLETT
22 OCT 2018

The company has rebranded its core offering with a view to avoiding confusion about the value it provides as a semantic layer for automating the management and analysis of data stored in cloud or on-premises data lakes.

451 Research®

Cambridge Semantics has rebranded its Anzo data management and analytics software, emphasizing its role in providing a semantic model-based data fabric layer for data management and analysis of data stored in cloud or on-premises data lakes.

THE 451 TAKE

While we hadn't encountered confusion related to the previous Anzo Smart Data Lake brand ourselves, we believe the rebranding as simply Anzo is a smart move following the previous consolidation of multiple products into a single offering, particularly when combined with positioning Anzo as a semantic model-based data fabric layer for data management and enterprise knowledge graphs. Understanding of these concepts is still fairly niche, but is growing, and we anticipate growing adoption of both the concepts and their associated technologies as part of the larger industry push toward being more data-driven.

CONTEXT

Cambridge Semantics was founded in 2007 and established its expertise in data management initiatives driven by semantic web technologies and standards. The company launched its first product in 2012 for integrating, preparing, and analyzing structured, unstructured and semi-structured data from multiple sources. It has expanded that platform over the years through internal development, as well as the 2016 acquisition of MPP in-memory graph analytics engine provider SPARQL City, which added the ability to handle large-scale high-performance analytics of multi-structured data. The company has approximately 100 employees, compared with 80 in November 2017. Having traditionally been strong in financial services and life sciences, in particular, the company reports seeing good growth in other sectors, including government, healthcare and retail.

PRODUCTS

The last time we caught up with Cambridge Semantics, it had consolidated a variety of products into a single offering covering data ingestion, data cataloging, in-memory graph processing and semantic modeling, data preparation, and interactive analytics. Then known as Anzo Smart Data Lake, that product is now known simply as Anzo – the company has moved to avoid any potential for the product being confused as a direct alternative to data lake environments based on Apache Hadoop or cloud storage.

The product was always designed to provide a semantic layer for data management and analysis that is installed on top of a Hadoop- or cloud-storage-based data lake environment. Shortening the branding to Anzo is therefore designed to avoid any such confusion, as is the description of Anzo as a semantic layer that connects data from across the enterprise into a single data fabric.

The functional elements that make up Anzo remain the same (albeit subject to ongoing development). They are: automated data ingestion and mapping; a semantic layer catalog that provides graph-based metadata management, governance and lineage; the AnzoGraph in-memory query engine; GraphMarts for creating layers of data preparation tasks, including data cleansing, transformation, linking and access control; and Hi-Res Analytics for interactive self-service analytics.

While Cambridge Semantics is not alone in providing data cataloging capabilities, it is worth noting that its semantic layer catalog approach is based on the company's long track record in semantic web technologies – specifically the W3C Web Ontology Language, Resource Description Framework and the RDF Query Language (SPARQL) standards – and that the semantic layer catalog can also optionally interoperate with an existing data catalog. Anzo is also designed to integrate with analytics products from the likes of Tableau, Qlik and TIBCO.

COMPETITION

Anzo could potentially be compared with a number of different products and services. We believe the most directly comparable products come from vendors that are also looking to enable enterprises to combine data from multiple data sources in order to generate value from the relationships exposed via semantic models as part of knowledge graph projects. Examples include Siren and Stardog, as well as the likes of MarkLogic, Ontotext and Franz, which are building on RDF graph database products.

Native graph database providers such as Neo4j, TigerGraph and Amazon Web Services are also promoting enterprise knowledge graphs (EKGs) as a potential use case for their products. We also see some potential for competition with Maana and its Maana Knowledge Platform, as well as Io-Tahoe and its data-discovery software platform.

While these companies are all focused on surfacing knowledge and insight based on the relationships in underlying data sources, Cambridge Semantics also sees competition from more traditional data management vendors, such as Informatica and IBM, as well as systems integration and consulting firms developing custom environments on-premises and in the cloud.

There is also some potential for comparison with catalog and data lake management providers such as Alation, Qlik's Podium, Unifi Software, Waterline Data and Zaloni, although the latest positioning around Anzo illustrates how it could also be used alongside an existing investment in data lake cataloging functionality by providing a complementary semantic-model-based data fabric layer.

SWOT ANALYSIS

STRENGTHS
Cambridge Semantics has long-standing expertise with semantic web technologies and standards that it combines with high-performance in-memory graph analytics.

WEAKNESSES
Semantic web approaches have seen niche adoption in the past, although we are seeing (and anticipate) greater interest driven by the enterprise knowledge graph concept.

OPPORTUNITIES
The new branding and positioning should help to avoid confusion about Anzo's relationship with existing data lake environments and data catalog deployments.

THREATS
Cambridge Semantics already faces a range of competitors, and we expect established data management vendors to muscle in on the enterprise knowledge graph concept as adoption grows.