

Coverage Initiation: DataGPT eyes business decision-makers with chat-driven analysis play

Analysts - Krishna Roy

Publication date: Friday, June 14 2024

Introduction

Data-driven decision-making is by no means "wall to wall." Only 27.1% of respondents to 451 Research's Voice of the Enterprise: Data & Analytics, Data Science & Decision Intelligence Platforms 2023 survey reported that nearly all strategic decisions within their organizations were data driven. An inability for business decision-makers to successfully conduct analysis on their own to obtain the insight they need — and to make decisions based on data — is a longtime barrier to making it more commonplace. DataGPT is seeking to enable organizations of all sizes to obtain the insight they need by letting business stakeholders chat with their data, deploying an internally developed, large language model-driven analysis platform.

The Take

The company has hit the market at the right time as conversational data analysis has become increasingly prevalent, validating it as a new approach to self-service analytics for business decision-makers. It is also fair to say that chat-based analytical capabilities vary significantly from vendor to vendor — and it is not immediately obvious which are superior. A major challenge for DataGPT, therefore, is one of market education to communicate the underlying capabilities it has developed in the form of analytics and compute engines — as well as a claimed internally developed LLM — to set its offering apart. While a differentiator, creating an internally developed LLM is expensive. As such, securing additional funding seems wise if the vendor fails to grow its business. Otherwise it could be acquired, causing unwanted disruption to the customer base it has amassed thus far.

Context

DataGPT is led by CEO Arina Curtis, who launched the business alongside co-founders Darren Pegg (CTO) and Sasha MacKinnon (chairman) in 2021 to make data-driven decisions easier for business personnel so they don't have to rely on overloaded data teams to provide them. The company's

initial focus was on data exploration with its internally developed Data Navigator. However, it made a pivot in October 2023 with the general availability of DataGPT, which is primarily deployed to chat with data but has real-time data navigation as a secondary aspect.

The vendor employs a direct sales model to target DataGPT at organizations of all sizes — startups, midsize firms (50-100 employees) and enterprises. It is courting business individuals because they struggle with self-service analysis. That said, data teams buy it. DataGPT has reportedly been purchased by 17 customers so far, and it is on track to land 25 customers by July. Its clients operate in a variety of industries, such as education, shipping, payment processing, healthcare, gaming and aviation.

The company has raised \$10 million in seed funding from undisclosed investors so far, and is currently seeking another seed round. It has 14 employees.

Products and technology

As a chat-based analytics platform, DataGPT allows business individuals to type a question into it, such as "Tell me about online delivery rates," and get a response in the form of an accompanying narrative insight and visualization. It also autocompletes questions and suggests follow-up questions based on the one initially asked.

Moreover, DataGPT uses analytics to give business individuals a detailed breakdown of key performance drivers. Its homegrown analytics engine runs data requests employing SQL, and can also call machine-learning models and external APIs. To ensure that the queries processed by its analytics engine are carried out swiftly, cost-efficiently and at scale, DataGPT is underpinned by a homegrown "lightening compute" engine — so-called because the vendor claims that it runs 90 times faster than a modern data warehouse and can process thousands of queries in milliseconds.

The vendor requires customer data to already be in a warehouse. It cites SQL data warehouses such as Amazon RedShift, Snowflake, Google BigQuery and Microsoft Azure as commonly deployed by customers. The specific LLM involved is configured by DataGPT depending on the data and types of questions asked. The company uses two types of LLM. One is OpenAI's GPT-4 Turbo, and the other it claims to have developed internally. Its own LLM is a continual focus for development, claimed to be currently comparable to GPT-4 Turbo.

Since unleashing DataGPT in October 2023, the vendor has found customers live in chat mode, rather than Data Navigator, which is its other interface, deployed to enable free-form exploration and analysis by dragging and dropping dimensions and measures, for instance. This customer behavior has prompted DataGPT to introduce a new interface that operates in chat-only mode.

Additionally, the vendor has found that users frequently don't know where to begin with an analysis, which has spurred the introduction of question suggestions for starting points and a breakdown of the data queries to understand how DataGPT understood the question, as well as what filters were applied to data extraction queries. The breakdown offered to end users also surfaces the metrics, dimensions and any filters employed to generate the insight in a narrative to assist with comprehension and insight transparency.

Dynamic benchmarking is another new customer-driven capability. This enables users to conduct time-series analysis to compare two different segments in a couple of different periods. Multidimensional and percentile-based analytics are other types of analysis capabilities introduced as part of an ongoing strategy to increase DataGPT's purview so it can manage a wider range of analysis types. Anomaly analysis to detect very tiny segments that exhibit anomalous behavior compared with the overall trend is currently in internal testing mode, for instance.

Packaging and pricing

DataGPT is a cloud-native service that currently runs in the AWS cloud and deploys S3 storage. That said, the company does not store customer data in the enterprise version because enterprises are highly conscious of security. Instead, enterprises can buy it as a virtual private cloud deployment.

Its package for startups is hosted by DataGPT as a multi-tenant service because startups are typically price-conscious and less concerned with security. Additionally, the vendor offers a version that uses a customer's own S3 bucket for data storage, which is aimed at medium-sized organizations.

Pricing involves a monthly subscription that starts at \$1,750 per month for 10 users. Organizations that want to evaluate the offering have the option to do so using DataGPT Xpress, which the vendor introduced in May as an alternative to a free trial. DataGPT Xpress houses preconfigured templates and thus does not have the adjustability of the full DataGPT platform.

The offering currently is only available for Google Analytics. The company plans to eventually furnish it with connectivity to other popular business applications from the likes of Shopify Inc., HubSpot Inc. and Salesforce Inc. DataGPT Xpress costs \$99 per team (which is three users) per month. Customers that sign up for it get a two-week free trial.

Finally, the vendor intends to unleash an embedded version with "Powered by DataGPT" branding, citing high customer demand to "integrate it inside with platform" as the reason for doing so. DataGPT Embedded is currently in beta testing with customers.

Competition

Conversational analytics has become increasingly common since OpenAI unleashed ChatGPT, resulting in the adoption of large language models for analytical insight. Purveyors of analytics platforms such as ThoughtSpot, Sisense, Alteryx, Tellius, Pyramid Analytics and Salesforce-owned Tableau have embraced the trend of including LLM-driven insight in their analytics platforms. Microsoft Corp. is the latest big gun to do so with the debut in May of Copilot within Power BI.

However, the major difference between DataGPT and these contenders' LLM-driven offerings is that they provide a chat experience as one component of a broader stack with other analytical and data management capabilities included in the mix, which DataGPT does not. Therefore, DataGPT is more akin to DataChat because both vendors are solely focused on offering a conversational analytical experience underpinned by multiple internally developed technologies — albeit not identical. For example, DataChat has not developed its own LLM and, moreover, provides a spreadsheet-driven interface as an alternative to its chat interface because it also targets analysts.

Finally, it is worth pointing out that SQL wrapper tools could — at face value — appear to be DataGPT rivals because of a shared focus on SQL-to-natural-language translation. However, DataGPT is underpinned by analytics and compute engines, so it can handle and process more complex questions in a quicker way than an SQL wrapper tool. The vendor claims that these capabilities also enhance accuracy and interpretation of questions asked.

SWOT Analysis

Strengths	Weaknesses
Most generative AI-based analysis offerings are designed for and targeted at enterprises, which gives DataGPT an early-mover advantage among smaller organizations. Its homegrown analytics and compute engines allow for a breadth of analysis capabilities in a cost-efficient and fast way.	By using DataGPT branding, the vendor is hitching itself to the success of OpenAI — which is by no means guaranteed. It also risks spreading itself too thin by pursuing enterprises, midmarket firms and startups.

Coverage Initiation: DataGPT eyes business decision-makers with chat-driven analysis play

Opportunities	Threats
Developing an LLM is expensive, so raising additional funding is a sensible strategy as it will provide the investment to do so, as well as serve to accelerate other road map items. The vendor will also have a chance to build an original equipment manufacturer customer base when DataGPT Embedded is available.	Many more established and high-profile analytics platforms than DataGPT's use generative AI to enable business individuals to have a conversational experience, providing a stiff competitive environment to operate in and the risk of being crowded out if it is perceived as a niche specialist.

Source: 451 Research.