

IBM to Acquire DataStax, Deepening watsonx Capabilities and Addressing Generative AI Data Needs for the Enterprise

Planned acquisition furthers IBM's commitment to open-source innovation; helps clients access untapped, unstructured enterprise data to maximize the impact of generative AI



ARMONK, N.Y., Feb. 25, 2025 /PRNewswire/ -- IBM (NYSE: IBM) today is announcing its intent to acquire DataStax, an AI and data solution provider. DataStax's technology will enhance IBM's watsonx portfolio of products accelerating the use of generative AI, helping companies unlock value from vast amounts of unstructured data.

The acquisition also builds on IBM's commitment to open-source AI. [DataStax](#) is the creator of AstraDB and DataStax Enterprise, NoSQL and vector database capabilities powered by Apache Cassandra®; and Langflow, the open-source tool and community for low-code AI application development.

IBM will continue to support, engage, and innovate with the open-source Apache Cassandra®, Langflow, Apache Pulsar™, and OpenSearch communities in which DataStax participates. IBM's long-standing commitment to open-source AI includes the open-source IBM Granite foundation models and Instruct Lab, a revolutionary approach to advancing true open-source innovation around LLMs.

Harnessing unstructured data for the enterprise

Businesses are struggling to harness valuable, unstructured data from across their business, which is critical for powering generative AI. Without tools to properly ingest and manage this untapped data, costly generative AI projects will fail to deliver their full potential. Even 70% of companies with high-performing generative AI initiatives experience data-related difficulties, [according](#) to McKinsey, and it is estimated that just one percent of enterprise data is represented in today's AI models.

IBM is a leader in helping clients scale generative AI and transform their business using enterprise data. The acquisition of DataStax enhances these efforts. Their vector database excels at harnessing unstructured enterprise data and accelerating its time to value, and Langflow provides a graphical, low-code design environment and component orchestration for generative AI apps that facilitates collaboration across diverse skillsets.

- **AstraDB** and **DataStax Enterprise** provide NoSQL and vector database capabilities powered by Apache Cassandra®, enabling production-ready generative AI applications for the enterprise. AstraDB will enhance the existing vector capabilities of IBM [watsonx.data](#), IBM's hybrid, open data lakehouse for AI and analytics. Thousands of organizations use Apache Cassandra®, including some of the biggest names in software, retail, finance, and ecommerce. Apache Cassandra® provides scalability, availability, fault tolerance, high performance, and multi-data-center and hybrid cloud support. And increasingly, Apache Cassandra® users are [leveraging the database for AI workloads](#). In this context, DataStax brings together a mature datastore with vector and graphRAG capabilities – a critical combination for harnessing unstructured data for generative AI
- **Langflow** is a low-code, open-source app builder for RAG and multi-agent AI applications. It is Python-based and model-, API-, and database-agnostic. Langflow adds additional flexible middleware capabilities to IBM [watsonx.ai](#), the integrated, end-to-end AI development studio for building generative AI applications

"Businesses cannot realize the full potential of generative AI without the right infrastructure – open-source tools and technologies that empower developers, harness unstructured data, and provide a strong foundation for AI applications," says Dinesh Nirral, Senior Vice President, IBM Software. "DataStax possesses deep competency in this area and shares IBM's relentless commitment to simplifying and scaling generative AI for the enterprise."

"Enterprises want to deliver production AI fast, but are still struggling to unlock the value in their data to power AI applications and agents," says Chet Kapoor, Chairman and CEO of DataStax. "DataStax's products solve this problem, accelerating AI's promise with the scalability, security, and accuracy developers and enterprises need. We've long said that there is no AI without data, and are excited to execute this vision with IBM."

DataStax's hundreds of customers include FedEx, Capital One, The Home Depot and Verizon. DataStax was founded in 2010 and is headquartered in Santa Clara, CA.

Financial details of the transaction were not disclosed. The acquisition is expected to close in the second quarter of 2025, subject to customary closing conditions and regulatory approvals.

About IBM

IBM is a leading provider of global hybrid cloud and AI, and consulting expertise. We help clients in more than 175 countries capitalize on insights from their data, streamline business processes, reduce costs, and gain a competitive edge in their industries. Thousands of governments and corporate entities in critical infrastructure areas such as financial services, telecommunications and healthcare rely on IBM's hybrid cloud platform and Red Hat OpenShift to affect their digital transformations quickly, efficiently, and securely. IBM's breakthrough innovations in AI, quantum computing, industry-specific cloud solutions and consulting deliver open and flexible options to our clients. All of this is backed by IBM's long-standing commitment to trust, transparency, responsibility, inclusivity, and service. Visit www.ibm.com for more information.

Media contact:

Kevin Zawacki
Communications, IBM Software
kevin.zawacki@ibm.com

Additional assets available online:  Photos 


<https://stage.mediaroom.com/ibmnewsroom/2025-02-25-ibm-to-acquire-datastax,-deepening-watsonx-capabilities-and-addressing-generative-ai-data-needs-for-the-enterprise>