MongoDB Delivers Accurate AI Retrieval Wherever Enterprise Data Lives

June 30, 2026

New Voyage AI capabilities and Search for on-premises and private cloud let enterprises build accurate, compliant AI applications to run anywhere without rewriting their applications and relying on bolt-on tools

BENGALURU, India, June 30, 2026 /PRNewswire/ -- MongoDB, Inc. (NASDAQ: MDB) today announced new capabilities at MongoDB.local Bengaluru that address the two reasons enterprise AI projects routinely stall before production: retrieval that isn't accurate enough to trust and infrastructure that can't meet compliance requirements. voyage-context-4, Hybrid Search, and Native Reranking work together to improve retrieval accuracy, with Native Reranking alone improving retrieval quality by up to 30%*. The capabilities are powered by Voyage AI models that outperform Google and Cohere on the public Retrieval Embedding Benchmark leaderboard. Search and Vector Search are now generally available for MongoDB Enterprise Advanced and Community Edition, bringing the same retrieval capabilities Atlas customers rely on to on-premises, private cloud, and local environments where regulated enterprises and startups operate. Together, these capabilities give enterprises and builders a production-ready retrieval stack that is accurate, compliant, and deployable wherever their data lives.

"The biggest barrier to enterprise AI in production and at scale isn't the LLM. It's memory, retrieval, accuracy, and compliance. Most enterprises aren't blocked by ambition. They're held back by infrastructure that wasn't designed to provide AI with trusted access to enterprise data. Bolting on more systems to solve those problems only creates more vendors, more latency, and more points of failure," said Ben Cefalo, Chief Product Officer, Core Products, MongoDB. "Whether you're running in the cloud, private cloud, or behind a firewall, MongoDB gives you the same production-grade retrieval capabilities wherever your data lives."

Voyage AI: Accuracy begins with top-ranked embedding models
Accuracy is the first bar AI has to clear for production. The second is ensuring AI works from current data, not outdated data sitting in a separate search system. Today, MongoDB launched three new capabilities, built into the database, that deliver more accurate retrieval and keep applications working from current data.

- Native Reranking in MongoDB Atlas, now in public preview, is powered by Voyage AI and delivers up to a 30% boost in retrieval quality directly inside the database, eliminating a leading cause of AI project failure. It works on top of existing search results, with no external APIs, keys, or round-trips to manage.
- Voyage Context 4, now generally available, is a new embedding model built for long documents. It processes long documents in full context rather than isolated chunks, preserving meaning across complex enterprise content for better retrieval accuracy. It drops into existing RAG pipelines without re-architecting.
- Hybrid Search in MongoDB, now generally available, combines full-text and vector search in a single query inside the operational database, delivering precision retrieval without separate systems or complex query logic. Because embeddings stay up to date automatically, agents retrieve from the current state of the data rather than a stale copy.

Emergent Labs is an AI-native app development platform and one of the fastest growing startups in the world. The company first tested its platform on PostgreSQL, where agents repeatedly got stuck in schema migration loops every time users refined their ideas. On MongoDB Atlas, agents create and modify data structures freely as applications evolve, and because search and embeddings live in the same database as that constantly changing data, retrieval keeps up with it.

"Our agents write code, modify data structures, and act on what they read back millions of times a day. If retrieval returns something stale or wrong, the agent builds on it, and the error compounds. MongoDB gives us the retrieval accuracy to keep agents working from the current state of the data, and that's what lets us run two million applications at scale," said Mukund Jha, CEO of Emergent Labs.

Run AI anywhere without compromising on accuracy or increasing risk
Retrieval accuracy is only half the problem enterprises face. The other half is whether they're allowed to run it where their data must reside, and for enterprises in regulated industries, the answer is rarely the public cloud. Data residency mandates, sovereignty rules, and compliance frameworks don't bend for innovation timelines, yet the most capable AI tooling has been built cloud-first, leaving regulated enterprises to choose between compliance and capability.

Today, MongoDB Search and Vector Search are now generally available as an add-on for MongoDB Enterprise Advanced, bringing the same retrieval capabilities MongoDB Atlas customers have been building in on-premises, private cloud, and hybrid environments, with the same platform, API, and technical skills regardless of where the workload runs. Ahead of this release, more than 20 of the world's largest banks and financial institutions have

been evaluating Search for Enterprise Advanced, drawn by the same thing: AI-ready retrieval that runs inside the infrastructure they control.

Search and Vector Search are now generally available for MongoDB Community Edition, enabling builders to implement AI retrieval locally at no cost. A startup can prototype on a laptop with full-text search, vector search, and hybrid search in one single system, then move to Atlas or Enterprise Advanced when it's ready to scale, without re-architecting or switching databases.

Investing in India for the long term

As part of MongoDB.local Bengaluru, the company also announced plans to upskill two million Indian builders by 2030. MongoDB is expanding its MongoDB for Academia program through partnerships with the All India Council for Technical Education, HCL GUVI, and the ICT Academy of Kerala. Since 2023, the program has reached more than 650,000 students.

MongoDB also launched Bengaluru to the Bay, a startup challenge that gives early-stage AI founders a path from India's builder ecosystem to San Francisco's AI community during SF Tech Week experience. $50,000 in MongoDB Atlas credits, travel, and go-to-market opportunities included.

What's new at MongoDB.local Bengaluru 2026

- **voyage-context-4** (Generally available): Next-generation contextualized embeddings with document-level context and auto-chunking; a drop-in upgrade for existing retrieval-augmented generation (RAG) pipelines.
- **Native Reranking in MongoDB Atlas** (Public Preview): Reranking runs inside the aggregation pipeline - no external APIs, no round-trips - and delivers up to a 30% boost in retrieval quality directly inside the database.
- **Hybrid Search** (Generally available): More accurate retrieval by combining full-text precision and vector-based semantic understanding in a single query on live operational data.
- **Search and Vector Search for MongoDB Enterprise Advanced** (Generally available): Production AI behind your firewall, under your compliance framework, with full parity to MongoDB Atlas capabilities.
- **Search and Vector Search in MongoDB Community Edition** (Generally available): Full-text, vector, and hybrid retrieval in self-managed environments, at zero cost to start.
- **MongoDB Atlas Stream Processing: Apache Iceberg Support** (Generally available): MongoDB Atlas now supports Apache Iceberg via the new $iceberg aggregation stage in Atlas Stream Processing, enabling any Atlas collection to be continuously synchronized to Iceberg tables on AWS object storage.
- **Gen2 MongoDB Atlas M30+ Dedicated Clusters on AWS** (Generally available): Next-generation infrastructure for high-scale production workloads.
- **Asymmetric Search Node deployment for multi-region Atlas clusters:** (Generally available): Set Search Node capacity to match each region's actual search traffic and lower total Search Node cost on multi-region clusters by 25–40%+
- **MongoDB for Academia Expansion**: Targeting 2 million builders trained by 2030 through HCL GUVI, ICT Academy of Kerala, and AICTE partnerships.
- **Bengaluru Meets the Bay—startup contest**: $50K in MongoDB credits plus travel and VIP access to MongoDB.local San Francisco for winning founders.

**Based on Voyage instruction-following rerankers on the MAIR benchmark; improvement measured over first-stage retrieval.*

About MongoDB
Headquartered in New York, MongoDB's mission is to empower innovators to create, transform, and disrupt industries with software. MongoDB's unified database platform was built to power the next generation of applications, and MongoDB is the most widely available, globally distributed database on the market. With integrated capabilities for operational data, search, real-time analytics, and AI-powered data retrieval, MongoDB helps organizations everywhere move faster, innovate more efficiently, and simplify complex architectures. Millions of developers and more than 65,200+ customers across industries—including ~75% of the Fortune 100—rely on MongoDB for their most important applications. To learn more, visit mongodb.com.

Forward-Looking Statements
This press release includes certain "forward-looking statements" within the meaning of Section 27A of the Securities Act of 1933, as amended, or the Securities Act, and Section 21E of the Securities Exchange Act of 1934, as amended, including new capabilities announced at MongoDB .local Bengaluru 2026. These forward-looking statements include, but are not limited to, plans, objectives, expectations and intentions and other statements contained in this press release that are not historical facts and statements identified by words such as "anticipate," "believe," "continue," "could," "estimate," "expect," "intend," "may," "plan," "project," "will," "would" or the negative or plural of these words or similar expressions or variations. These forward-looking statements reflect our current views about our plans, intentions, expectations, strategies and prospects, which are based on the information currently available to us and on assumptions we have made. Although we believe that our plans, intentions, expectations, strategies and prospects as reflected in or suggested by those forward-looking statements are reasonable, we can give no assurance that the plans, intentions, expectations or strategies will be attained or achieved. Furthermore, actual results may differ materially from those described in the forward-looking statements and are subject to a variety of assumptions, uncertainties, risks and factors that are beyond our control including, without limitation: our customers renewing their subscriptions with us and expanding their usage of software and related services; global political changes; the effects of the ongoing military conflicts between Russia and Ukraine and Israel and Hamas and recent events in Venezuela on our business and future operating results; economic downturns and/or the effects of rising interest rates, inflation and volatility in the global economy and financial markets on our business and future operating results; our potential failure to meet publicly announced guidance or other expectations about our business and future operating results; reputational harm or other adverse consequences resulting from use of AI and ML in our product offerings and internal operations if they don't produce the desired benefits; our limited operating history; our history of losses; our potential failure to repurchase shares of our common stock at favorable prices, if at all; failure of our platform to satisfy customer demands; the effects of increased competition; our investments in new products and our ability to introduce new features, services or enhancements, including AI and ML; social, ethical and security issues relating to the use of new and evolving technologies, such as artificial intelligence, in our offerings or partnerships; our ability to effectively expand our sales and

marketing organization; our ability to continue to build and maintain credibility with the developer community; our ability to add new customers or increase sales to our existing customers; our ability to maintain, protect, enforce and enhance our intellectual property; our ability to continue to increase revenue from our Atlas platform; the effects of social, ethical and regulatory issues relating to the use of new and evolving technologies, such as AI and ML, in our offerings or partnerships; the growth and expansion of the market for database products and our ability to penetrate that market; our ability to maintain the security of our software and adequately address privacy concerns; our ability to manage our growth effectively and successfully recruit and retain additional highly-qualified personnel; our ability to integrate acquisitions and work with our strategic partners effectively; and the price volatility of our common stock. These and other risks and uncertainties are more fully described in our filings with the Securities and Exchange Commission ("SEC"), including under the caption "Risk Factors" in our Annual Report on Form 10-Q for the quarter ended April 30, 2026, filed with the SEC on May 29, 2026. Additional information will be made available in other filings and reports that we may file from time to time with the SEC. Except as required by law, we undertake no duty or obligation to update any forward-looking statements contained in this release as a result of new information, future events, changes in expectations or otherwise.

Contacts
Investors
ir@mongodb.com

Media
press@mongodb.com

View original content to download multimedia:https://www.prnewswire.com/news-releases/mongodb-delivers-accurate-ai-retrieval-wherever-enterprise-data-lives-302813983.html

SOURCE MongoDB, Inc.