NAND RESEARCH

RESEARCH NOTE

SAP CLOUD CLOUD VECTOR ENGINE

STEVE MCDOWELL, CHIEF ANALYST
APRIL 11, 2024

CONTEXT

SAP recently introduced its new SAP HANA Cloud Vector Engine, designed to bridge the gap between LLMs' advanced capabilities and enterprise-specific business needs.

NEW: SAP HANA CLOUD VECTOR ENGINE

SAP HANA's new vector engine is integrated directly into SAP HANA Cloud's multi-model database, enhancing the database's ability to process and understand vector data—a fundamental component of AI and machine learning applications, especially those involving natural language processing and semantic understanding.

Key Features and Capabilities of the new capability include:

- **Multi-Model Database Integration:** The vector engine extends SAP HANA Cloud's multi-model capabilities, allowing it to store and process vector data alongside traditional data types such as relational, graph, spatial, and JSON.
- **Enhanced AI and ML Processing:** By enabling the similarity search between vectors, the engine significantly improves the database's ability to support AI-driven tasks such as semantic search, document analysis, and personalized content recommendations.
- **Retrieval-Augmented Generation:** The vector engine supports retrieval-augmented generation (RAG), enhancing the functionality of LLMs by supplementing them with up-to-date, relevant organizational data. This approach addresses one of LLMs' key limitations—their reliance on the static data they were trained on—by dynamically providing them with the context necessary for more accurate and relevant outputs.
- **Optimization for Business Applications:** Integrating the vector engine within SAP HANA Cloud optimizes large language models for business-specific

scenarios. It achieves this by facilitating the connection between the abstract capabilities of LLMs and the concrete, data-driven needs of business processes.

ANALYSIS

The new SAP HANA Cloud vector engine is a forward-thinking response to the increasing demand for more integration of enterprise data with the power of LLMs and generative AI.

By facilitating the combination of LLMs with real-time, company-specific data and business know-how, SAP not only addresses LLMs' current limitations, such as their reliance on outdated training data, but also significantly enhances AI's applicability within the business context and unlocks new insights for businesses.

The ability to process and analyze vector data alongside traditional relational, graph, spatial, and JSON data within a single multi-model database environment is a significant leap forward. This enables the development of innovative applications that are more intuitive and responsive to user needs, leveraging the full spectrum of data types and AI algorithms.

The vector engine's support for RAG addresses one of the critical challenges in deploying LLMs effectively in the enterprise: the need for more access to current, specific data. SAP HANA Cloud delivers more accurate, context-aware AI functionalities by providing LLMs with relevant organizational data.

The introduction of the vector engine is a cornerstone of SAP's broader generative AI strategy, indicating a deep investment in AI as a driver for future growth and innovation. It underscores SAP's vision of creating a seamless ecosystem where AI and data technologies complement and enhance each other, offering businesses a comprehensive suite of tools to leverage AI effectively.

For customers, the enhanced SAP HANA Cloud promises more robust, AI-driven applications that can push the boundaries of what's possible in business operations and customer engagement.

The new SAP HANA Cloud vector engine capability is a significant milestone in fusing AI with enterprise data services, significantly benefiting businesses that leverage technology to drive innovation, efficiency, and growth.