DATASTAX

Unleashing the power of GenAI

Preventing Hallucinations, Securing AI and Controlling Costs

July 2024

Meet Michel and Dieter

Michel de Ru
Solution Engineer

Dieter Flick
Solution Engineer

DATASTAX ©2024 DataStax. – All rights reserved

Abstract

Join Dieter and Michel in this session as they demonstrate how leveraging Astra DB's vector store and Langflow can significantly expedite the development of applications powered by LLMs. This session will provide a detailed look into how these technologies streamline the creation and deployment of LLM-driven solutions, significantly speeding up development processes.

The session will start with an introduction to the vector capabilities of Astra DB, which are essential for managing the high-dimensional data demands of generative AI applications. We will then focus on how Langflow, a pioneering low-code platform, accelerates the development lifecycle of LLM-powered applications. By facilitating rapid prototyping and iteration, Langflow enables developers to reduce development time dramatically.

Agenda

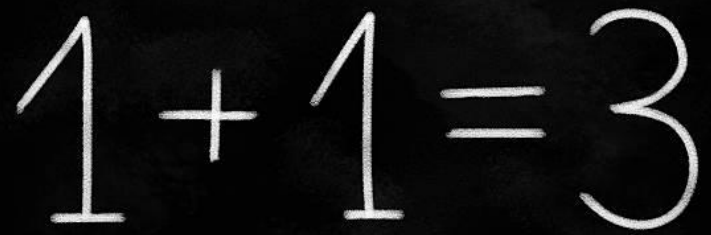
- Setting the Context
- Astra DB and its Vector Capabilities
- Vectorize the Easy Path
- Coding vs No-Coding
- Takeaways

DATASTAX ©2024 DataStax. – All rights reserved

Setting the Context

“AI can be dangerous if it doesn't know which data is correct”

A concern among developers, users and Michel’s neighbour

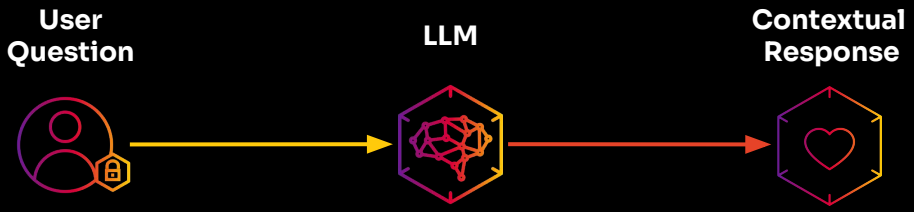


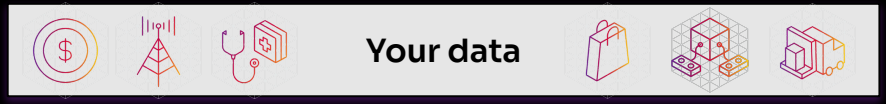
DATASTAX ©2024 DataStax. – All rights reserved

Just a Large Language Model

DATASTAX ©2024 DataStax – All rights reserved.

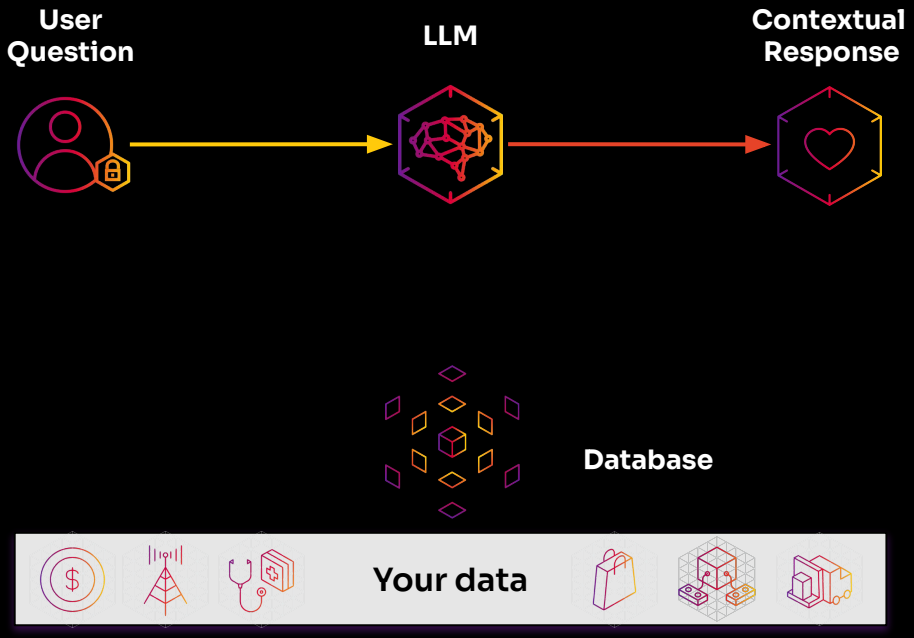
Tapping into your data





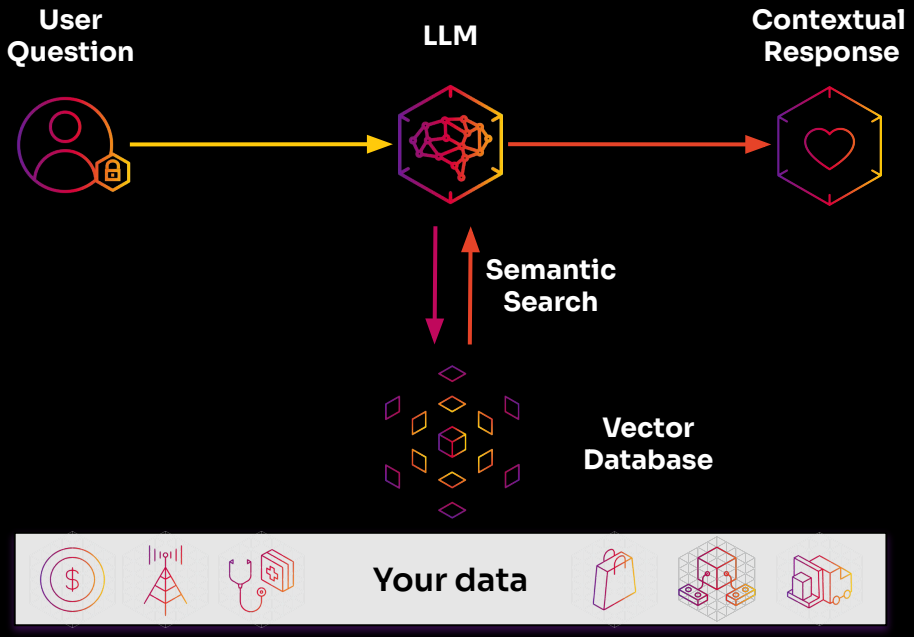
DATASTAX ©2024 DataStax – All rights reserved.

We need a database



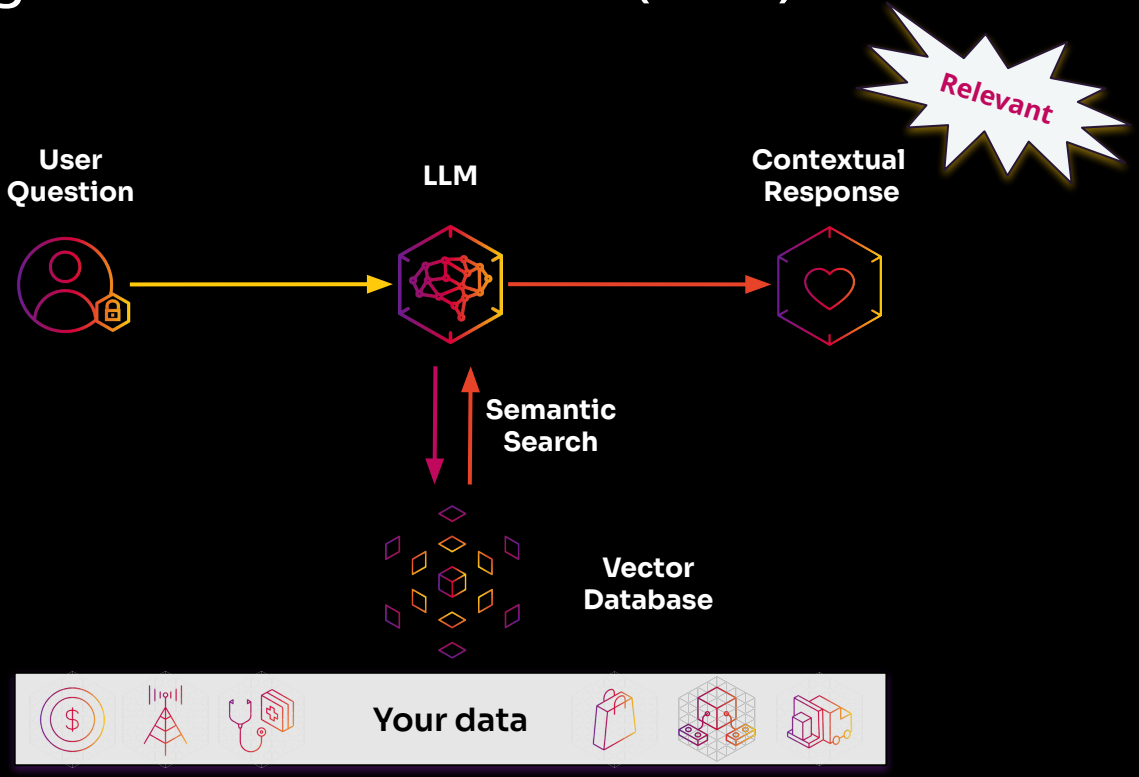
©2024 DataStax – All rights reserved.

With Semantic Search



©2024 DataStax – All rights reserved.

Retrieval Augmented Generation (RAG)



©2024 DataStax – All rights reserved.

Demo:
Without vs. with RAG

Takeaway no. 1
Leverage your competitive advantage

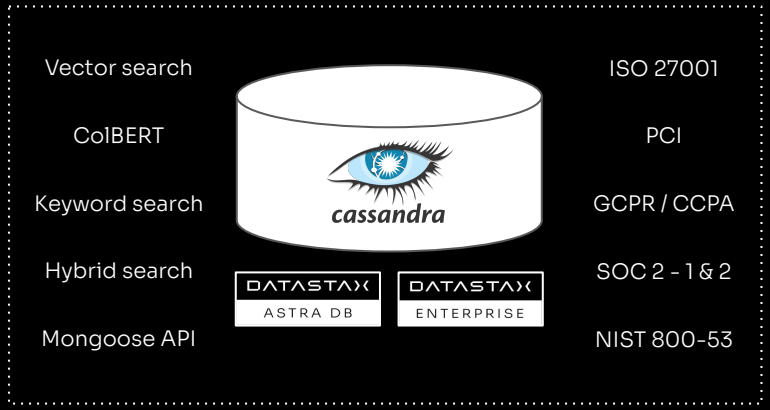
Your data and content are your crown jewels.
Keep them close so you're in full control of innovation and application.
Prevent hallucinations, securing AI, controlling cost

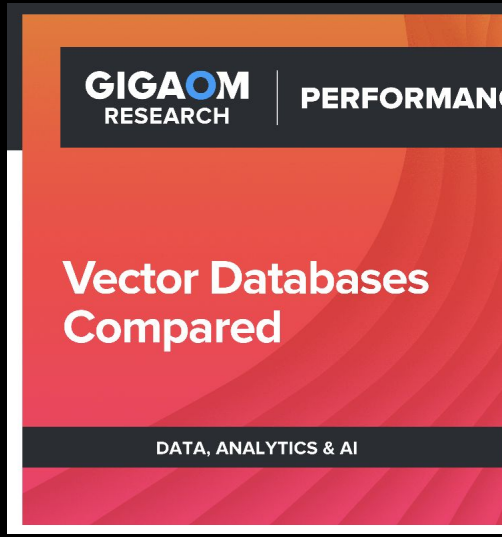
Astra DB and its Vector Capabilities

DATASTAX ASTRA DB

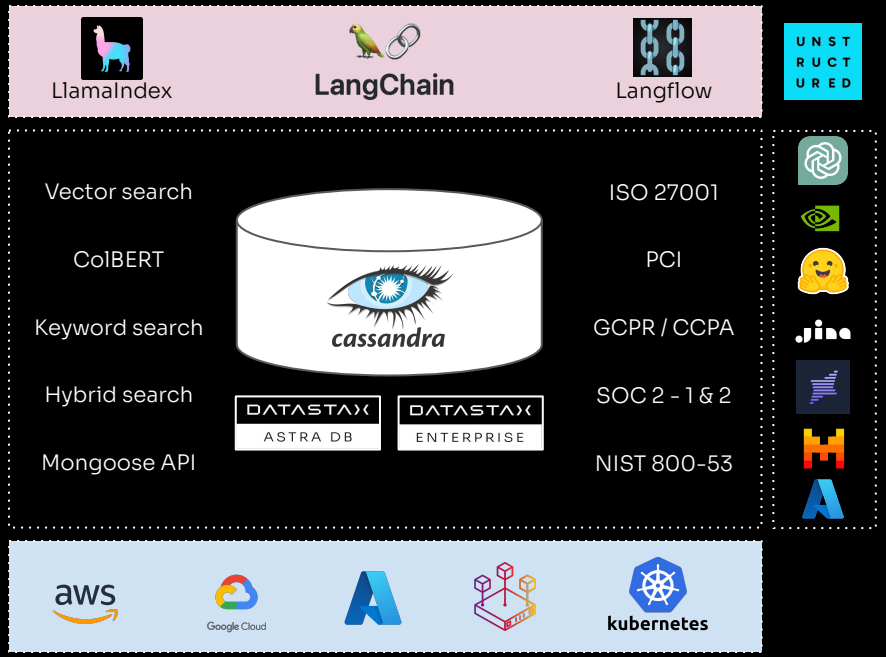
DATASTAX ENTERPRISE

Up to **20% more relevant AI Search** than Pinecone
9x more throughput
Up to **74x faster response times**
Astra DB delivers vector data updates with **ZERO delay**



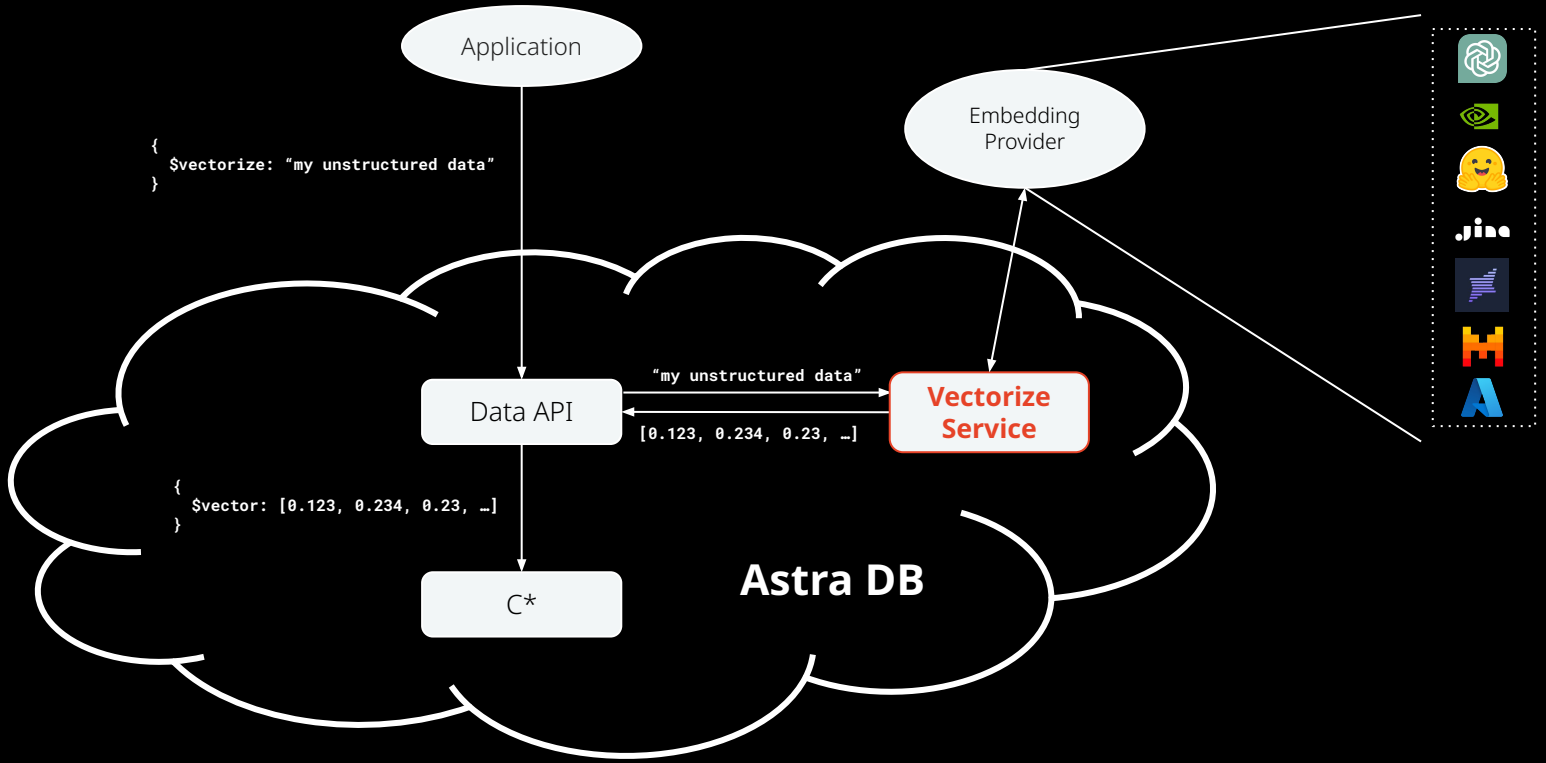


Introducing RAGStack



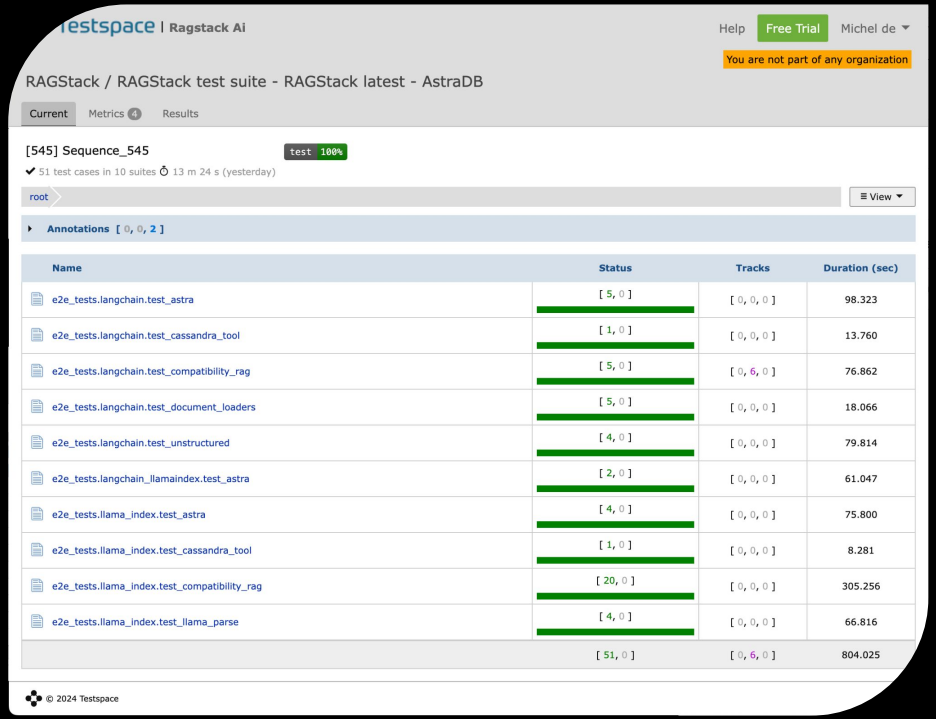
RAG on **your data** in the **public cloud** as well as **private cloud**

Vectorize the Easy Path



Demo:
Vectorize the Easy Path

RAGStack is ready for the enterprise



Curated stack for GenAI
Opinionated stack with LangChain, Langflow, LLamaIndex, Unstructured.io and more.

Dependency management
Brittle dependencies are something of the past with guaranteed to work package mgmt.

Vulnerability tested
Automated tests make sure no vulnerabilities creep in through 100s of dependencies

Enterprise ready
With 24x7x365 support offered through a well trusted partner. Also includes Long Term Support (LTS).

©2024 DataStax – All rights reserved.

“Every company building with generative AI right now is looking for answers about the most effective way to implement RAG within their applications,” said Harrison Chase, CEO, LangChain.

“DataStax has recognized a pain point in the market and is working to remedy that problem with the release of RAGStack. Using top-choice technologies, like LangChain and Astra DB among others, Datastax is providing developers with a tested, reliable solution made to simplify working with LLMs.”

See the joint webinar

Harrison Chase
CEO, LangChain

Coding demo with RAGStack

Langflow

- A UI for **LangChain**
- Low-code approach to build AI and LLM-powered applications.
- Makes it 100x faster for developers to prototype, iterate and ship AI applications to production.

DATASTAX ©2024 DataStax. – All rights reserved

No-Coding demo with Langflow

Takeaway no. 2
It’s not about AI tools and services

Tomorrow a better AI will emerge. But your data and content will always be your crown jewels.
Make sure you stay in control and leverage a database with Vector support.
Leverage enterprise hardened solutions like RAGStack and use rapid prototyping tools.

©2024 DataStax – All rights reserved.

What will you (re)create today?

Try it on Astra DB get $300 of credits and win!
(raffle for new signups)

2x spot at the DataStax Oktoberfest table

DATASTAX

One-click RAG @ Scale

Visit us in Hall 2.2 Booth 2_02

DATASTAX