

# Vector Store: Uber's Embedding Platform

Divya Nagar, Staff Software Engineer, Uber Technologies Inc

Xiyuan Feng, Sr Software Engineer, Uber Technologies Inc

# Agenda

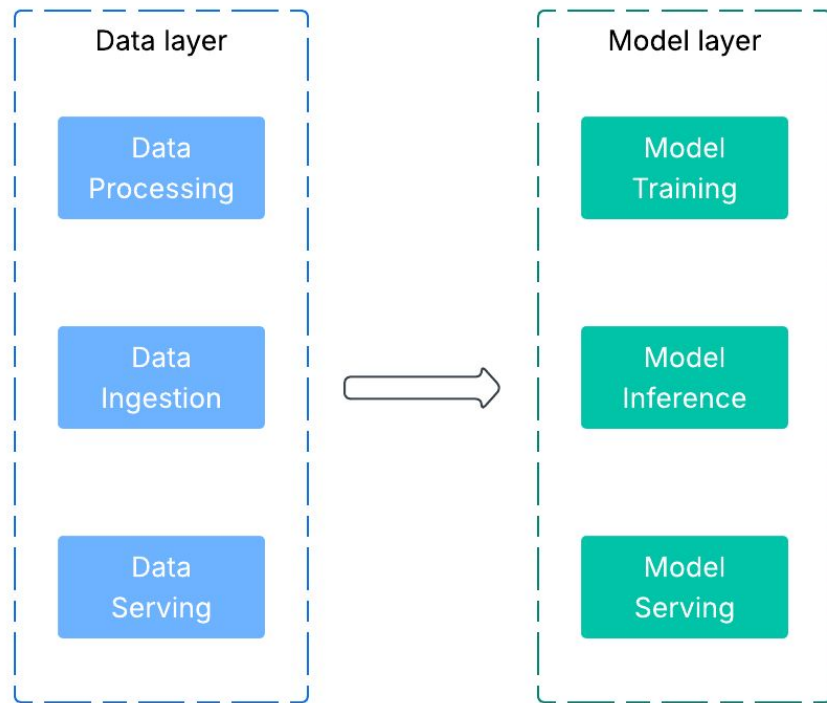
- ❏ What is Vector Store?
- ❏ Vision
- ❏ Data Sources
- ❏ Data Ingestion
- ❏ Data Retrieval
- ❏ Data Versioning

## What is Vector Store?

A vector store is a centralized system for managing embedding datasets, providing ingestion, retrieval, and metadata management to support generative AI and predictive ML use cases.

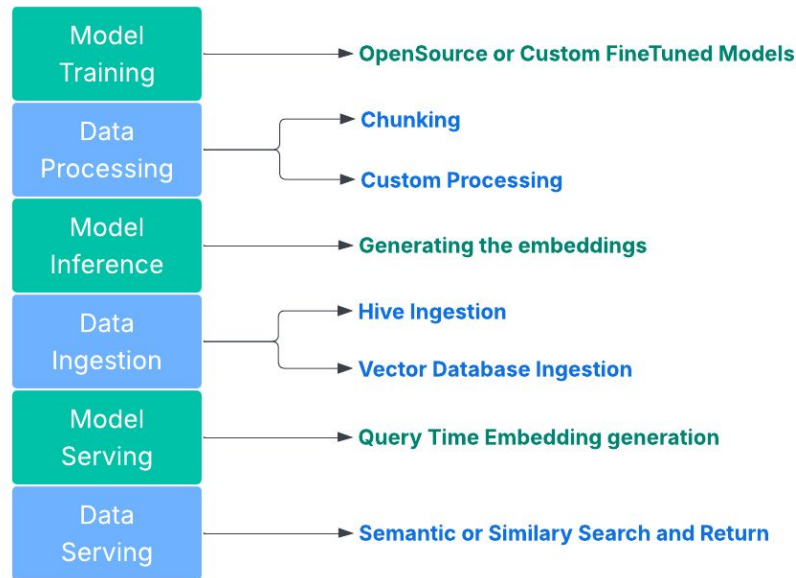
# Feature Store vs Vector Store: Data & Models

- Independent Operation
- Loose Coupling
- Decoupled Versioning



# Feature Store vs Vector Store: Data & Models

- Tight Integration
- Inference Dependency
- Coupled Versioning





## Feature Store vs Vector Store: Data Sources

Structured/tabular data

Unstructured text & content

Aggregated metrics & time-series features

Knowledge bases, Wiki, code repositories

Point-in-time lookups, joins, feature fetch

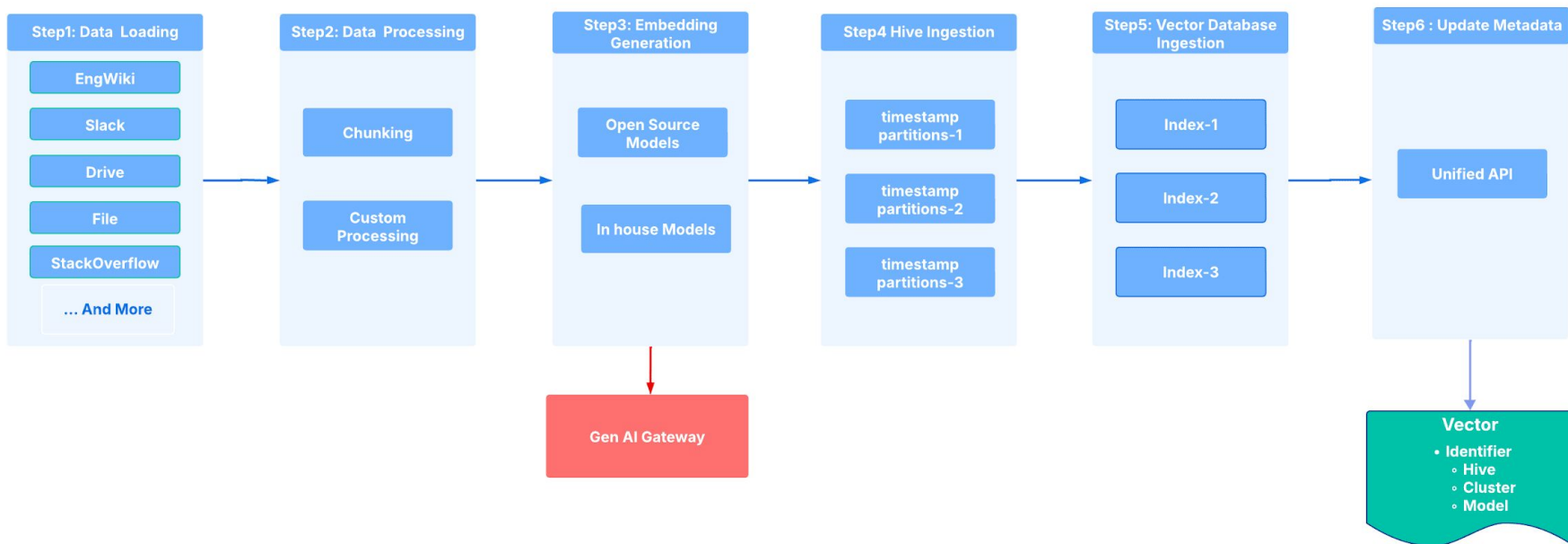
Semantic and Approximate Search



## Vector Store Vision

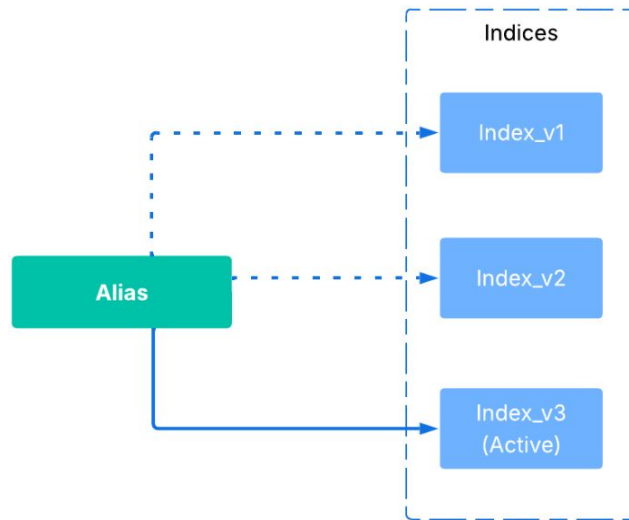
- **Unified Data Platform** – Standardizes ingestion and processing across all data sources.
- **Seamless Embedding Generation** – Runs inference pipelines to produce embeddings at scale.
- **Data–Model Coupling Management** – Handles data/model versioning and compatibility.
- **Unified Serving Layer** – Offers consistent online and offline retrieval APIs.

# Data Processing & Ingestion



# Dataset Versioning

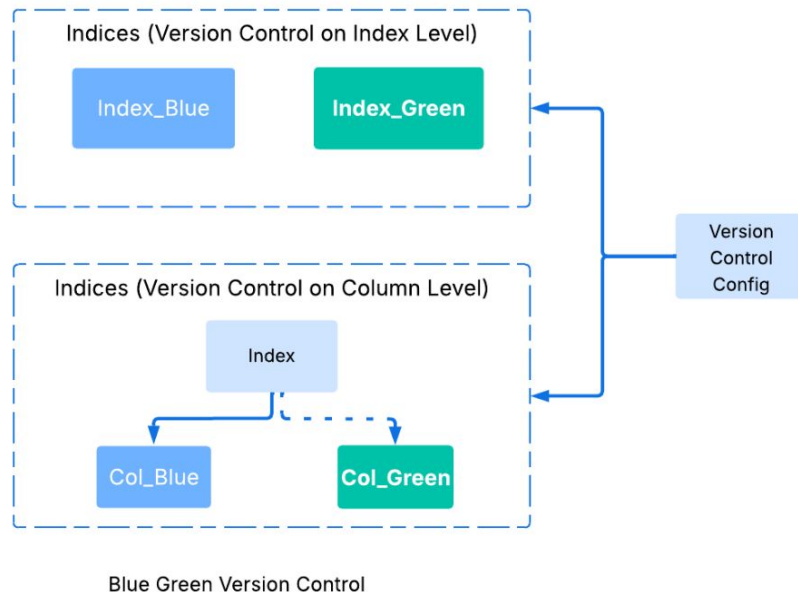
- Snapshot Versions
- Rollback Support
- Lower Reliability Overhead
- Higher Cost



Open Search Version Control

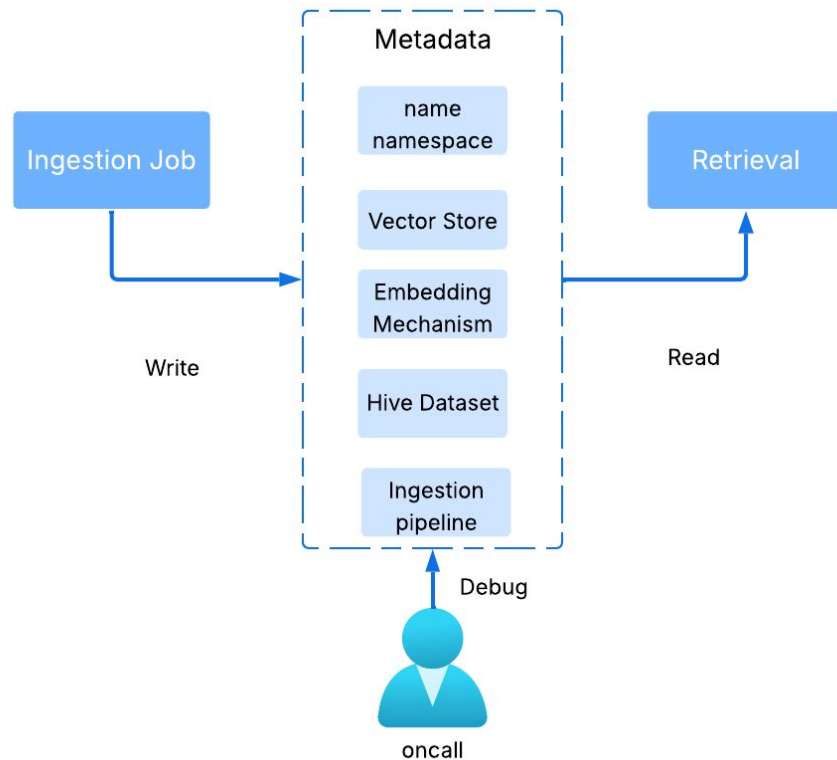
# Dataset Versioning

- Index Override
- Rollback Support
- Higher Reliability Overhead
- Lower Cost

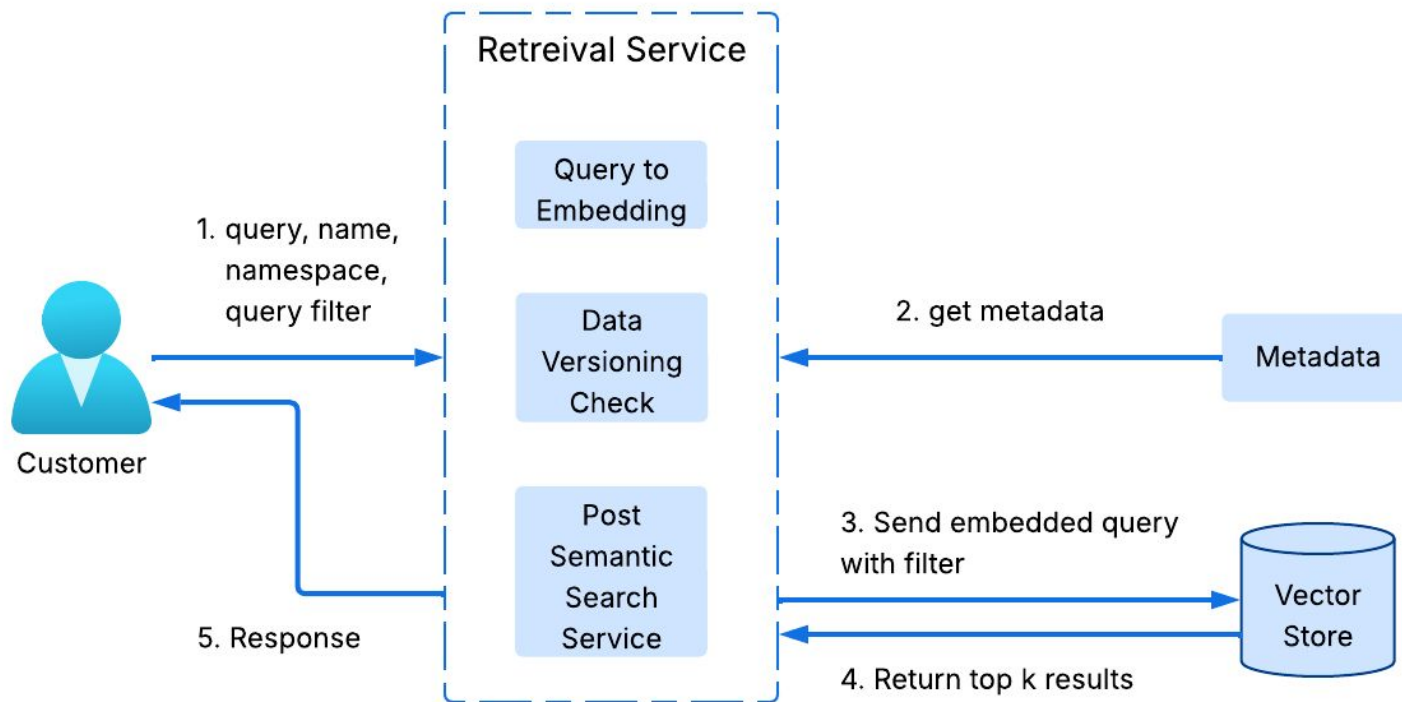


# Dataset Metadata

- Name & Namespace
- Online Serving Info
  - Cluster
  - Alias
- Hive information
- Embedding Model
- Other Metadata



# Dataset Retrieval



## Use cases

- Grocery and Ubereats Global Search
- [Uber gen-ai oncall copilot](#)
- Many other agentic application build within uber

## References

- <https://www.uber.com/blog/palette-meta-store-journey/>
- <https://www.uber.com/blog/genai-gateway/>
- <https://www.uber.com/blog/genie-ubers-gen-ai-on-call-copilot/>