NebulaByte: Vector Database Guide

NebulaByte Vector Database Technologies

What are Vector Databases?

Vector databases store and retrieve high-dimensional vector embeddings efficiently. They enable similarity search, which is fundamental for:

- Semantic search
- Recommendation systems
- RAG applications
- Image and video search

Popular Vector Databases:

1. FAISS: Facebook's library for efficient similarity search
2. Chroma: Open-source embedding database
3. Pinecone: Managed vector database service
4. Weaviate: Open-source vector search engine
5. Milvus: Cloud-native vector database

Key Concepts:

- Embeddings: Numerical representations of data
- Distance Metrics: Cosine similarity, Euclidean distance
- Indexing: Optimizing search performance
- Dimensionality: Typical ranges from 384 to 1536

FAISS Features:

FAISS provides multiple index types including IndexFlatL2 for exact search, IndexIVFFlat for faster approximate search, and IndexHNSW for hierarchical navigable small world graphs.