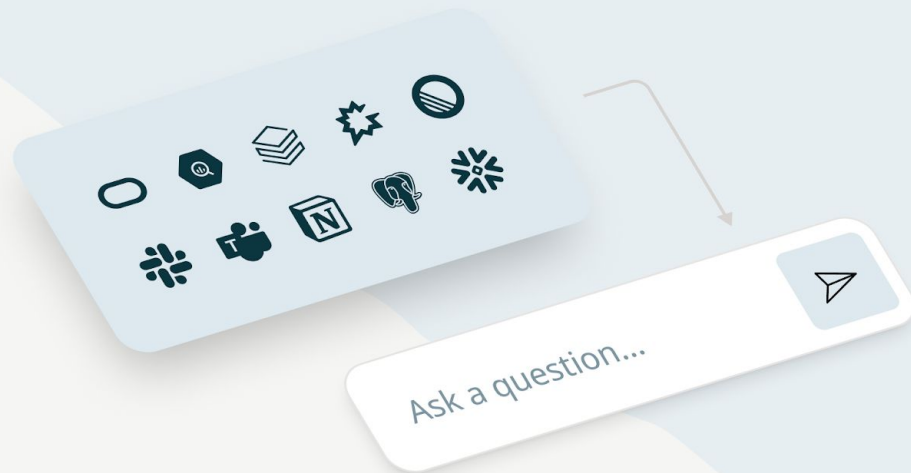


# Connect. Unify. Respond.

Any data, anywhere with  
human level intelligence.





Most widely-adopted AI  
query engine in the world



400K Deployments

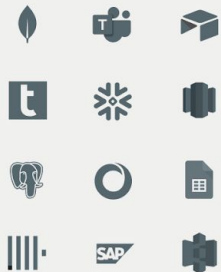


30K GitHub Stars



200+ Enterprise Datasources

Zero-ETL Support for all major  
large-scale enterprise datasources



Industry-leading tech partners



Google



Microsoft

ORACLE



Backed by top investors



BENCHMARK

Mayfield

Y Combinator



**Unify Data**



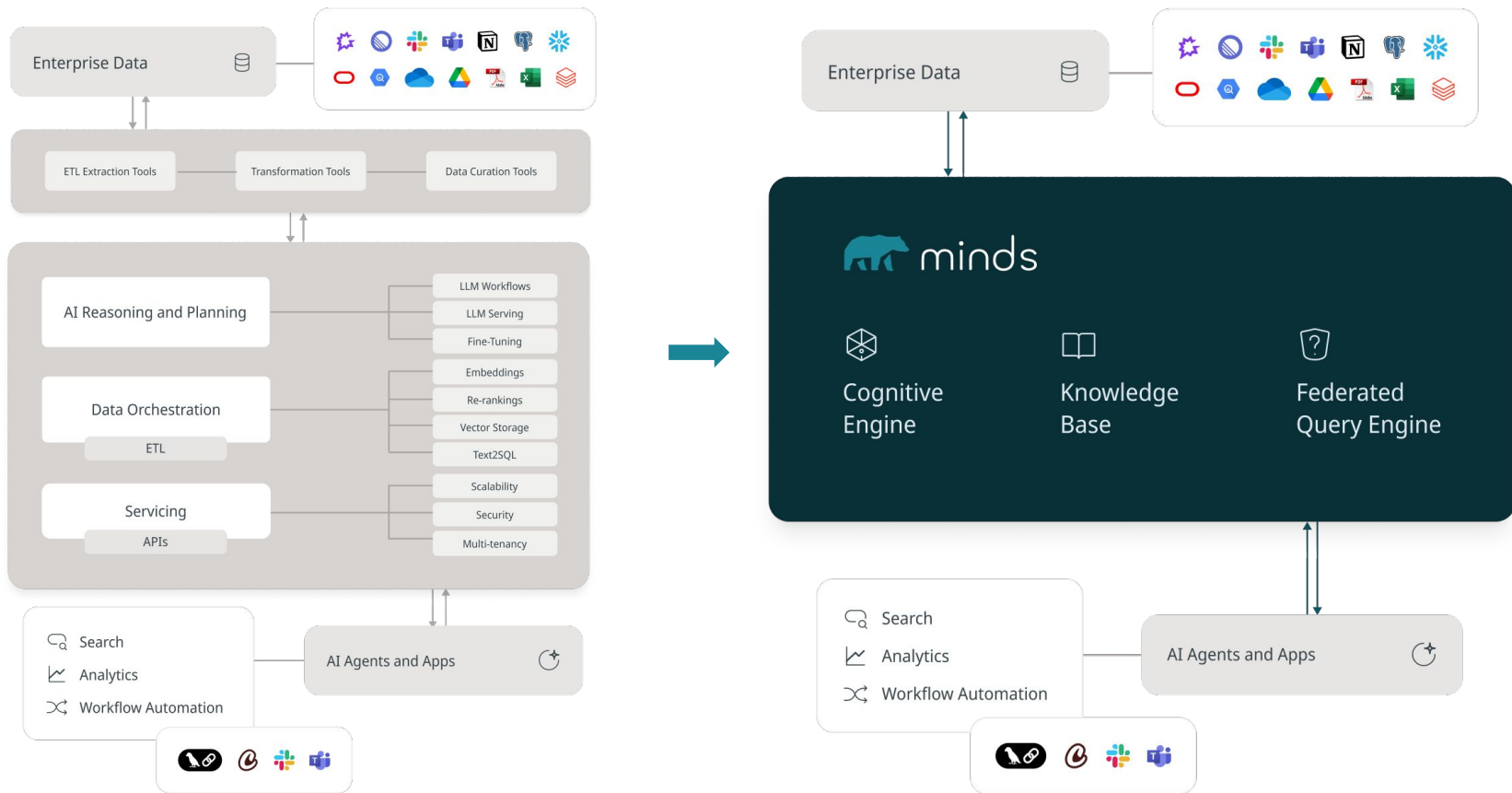
**Accelerate Insights**



**Empower Teams**



# MindsDB Simplifies and Derisks AI Implementations





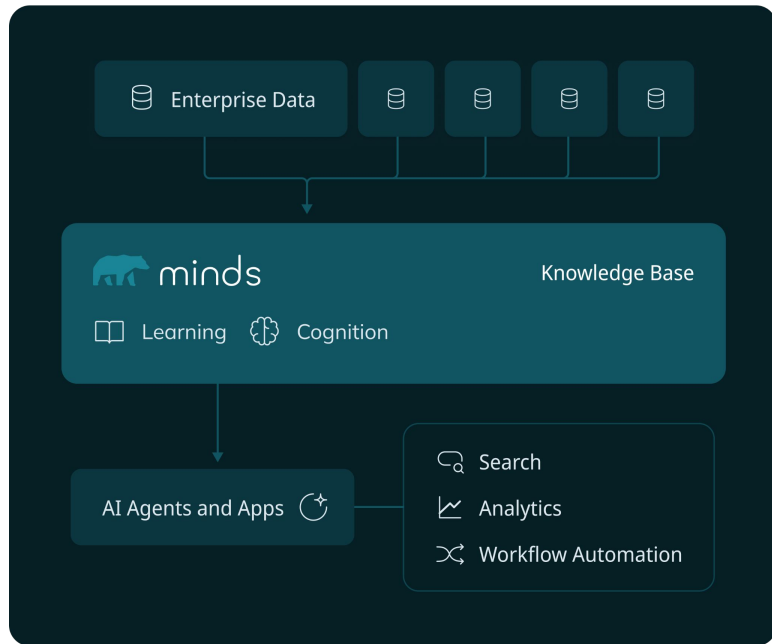
# Why MindsDB

Eliminate the complexity of fragmented enterprise data

## Minds

MindsDB empowers AI and humans to make data-driven decisions instantly.

- Query any data, anywhere
- Secure and scalable to any infra
- No data engineering required





# The MindsDB Stack

A Seamless Opportunity for Maximum AI Value

## Federated Query Engine (Connect)

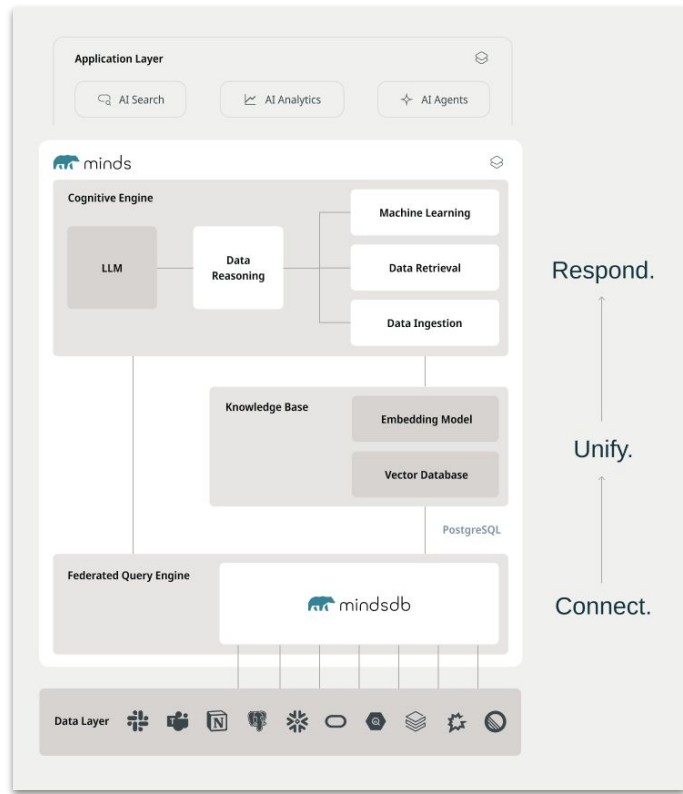
- Retrieves accurate, up-to-date knowledge
- Translates from SQL to any query language

## Knowledge Base (Unify)

- Makes sense of structured & unstructured data
- Avoids cost and complexity of moving data

## Cognitive Engine (Respond)

- Supports leading proprietary and OSS LLMs
- Orchestrates w/reasoning and planning to surface the best knowledge





# Knowledge Base - Features

## Unified SQL Interface

Manage and query AI systems with standard SQL commands like **CREATE**, **INSERT**, and **SELECT**, reducing the learning curve for developers.

## Intelligent Semantic Search

Searches based on conceptual meaning, using embedding and reranking models to find the most relevant results.

## Flexible Model & Storage Integration

Natively connects to major LLM providers (OpenAI, Azure, Bedrock) and vector databases, offering maximum flexibility .



# Knowledge Base - Features

## Powerful Hybrid Queries

Combine semantic search and keyword search with structured metadata filters in a single query to pinpoint the most accurate information.

## Automated Data Management

Automates complex data handling, including text chunking, embedding, and updates, to streamline the entire data ingestion process.

## Built-in Performance Evaluation

Use a dedicated **EVALUATE** command to measure the accuracy and relevance of your knowledge base, ensuring optimal performance.

# Create a Knowledge Base

Example



# CREATE KNOWLEDGE BASE

Open AI

**embedding\_model1:** Specifies the model used to convert your text data into vector representations (embeddings). This allows for semantic searching (finding content based on meaning, not just keywords).

**reranking\_model1:** Defines an optional model to refine the search results. After an initial retrieval using embeddings, this model re-evaluates and re-orders the results for better relevance.

```
CREATE KNOWLEDGE_BASE my_kb
USING
    embedding_model = {
        "provider": "OpenAI",
        "model_name" : "text-embedding-3-small",
        "api_key": "sk-..." -- optional, default from env variable
    },
    reranking_model = {
        "provider": "OpenAI",
        "model_name": 'gpt-4o',
        "api_key": "sk-..." -- optional, default from env variable
    },
    storage = my_vector_store.storage_table, -- optional, default ChromaDB
    metadata_columns = ['date', 'creator', ...], -- optional
    content_columns = ['review', 'content', ...], -- optional, default content
    id_column = 'id';
```

# Query a Knowledge Base

Example



# SELECT \* FROM my\_kb

Search at Document level

**Action:** Performs semantic search on `<kb_name>` using `WHERE content = '<query_text>'`.

**Returns:** Document `id` and `relevance` score.

**Optional Filtering:** Use `relevance` (0-1) to set minimum relevance.

**Optional Limit:** Use `LIMIT <N>` to restrict the number of results.

```
SELECT id, relevance
FROM <kb_name>
WHERE content = '<query_text>'
AND relevance >= 0.6 -- optional
LIMIT <N>;
```



# Hybrid Search

Search By keyword

**Hybrid Search:** Use `hybrid_search = true` to combine keyword BM25 and similarity search.

```
SELECT id, relevance
FROM <kb_name>
WHERE content = '<query_text>'
AND hybrid_search = true -- optional
LIMIT <N>;
```



# SQL Algebra Powered

Search By keyword

Traditional databases excel at structured data queries, while vector databases handle semantic similarity search. **MindsDB Knowledge Bases** combine both worlds, allowing you to run **complex SQL queries** that: that seamlessly blend:

- **Semantic content search**
- **Traditional metadata filtering**
- **Complex boolean logic** with AND, OR, and NOT operations
- **Advanced SQL features** like GROUP BY, ORDER BY, and UNION

```
SELECT * FROM movies_kb
WHERE content LIKE 'technology computer internet cyber'
      AND ((genre = 'Thriller' AND rating > 7.0)
           OR (genre = 'Sci-Fi' AND rating > 6.0)
           OR (expanded_genres LIKE '%Action%'
               AND rating > 6.5));
```

# Thank You