



REDUCING MANUAL SEARCH EFFORTS BY 50% WITH AI-DRIVEN DATA RETRIEVAL

THE CHALLENGE

Pharmaceutical organizations depend on large volumes of commercial and clinical data to guide market strategy, maintain regulatory compliance, and generate competitive insights, but accessing the right information quickly can be difficult when data is spread across multiple systems. In this case, the client's analytics teams struggled to efficiently locate relevant commercial drug data within a Snowflake environment due to the lack of an intuitive search capability, forcing them to manually navigate databases and metadata. To address this challenge, the client partnered with AE Partners to implement an AI-powered data discovery and retrieval solution that streamlined search, improved access to key datasets, and accelerated insight generation for analytics and business intelligence teams.

THE SOLUTION

Following a comprehensive assessment of the client's data architecture and user needs, the AE Partners team designed and implemented an AI-driven search and data retrieval framework to simplify access to complex datasets.

Key components of the solution included:

- Developing a metadata-driven search framework within the Snowflake environment to accurately identify relevant databases and tables, ensuring precise query targeting.
- Deploying an AI-powered SQL agent to automate data extraction, allowing users to retrieve specific datasets through natural language queries or predefined search conditions.
- Implementing Retrieval-Augmented Generation (RAG) to improve search accuracy for unstructured data by combining AI-generated responses with relevant information from a vector database.
- Leveraging cosine similarity search to match user queries with the most relevant datasets based on intent and context.
- Creating an optimized embedding and indexing system using Azure OpenAI and ChromaDB, enabling faster and more accurate data discovery.

This solution enabled analytics teams to search, discover, and retrieve commercial drug datasets far more efficiently than traditional manual methods.



TECH STACK

- Azure OpenAI Service
- GPT-3.5 Turbo
- Snowflake
- FastAPI
- React
- ChromaDB
- Vector embeddings and similarity search

THE RESULTS

The AI-powered solution significantly improved the client's data discovery and analytics workflows.

Key outcomes included:

- 50-60% reduction in manual effort required for metadata creation and data searching
- Improved search accuracy and result quality across commercial datasets
- Accelerated data discovery, enabling faster analytics and reporting
- Enhanced business intelligence capabilities through AI-powered cognitive search

