SQLQueryAI: A Neural Interface for Structured Query Synthesis

(A Generative AI Text-to-SQL Framework for Real-Time Data Access)

Sharmila Devi Chandariah

Senior Technical Lead, Senior Member IEEE , Policy Administration System, GenAI ML, NLP Research

Abstract

SQLQueryAI is a domain-agnostic Generative AI model developed to convert natural language into SQL queries, enabling rapid, accurate, and democratized access to structured data. Initially customized for the insurance domain as *InsureQueryAI* and adopted by Fortune 100 company as *IntelliQuery*, the solution has demonstrated transformative impact achieving 93.33% accuracy and reducing manual effort by 94%, resulting in $600K annual savings for a single team. SQLQueryAI supports both basic and Retrieval Augmented Generation (RAG) architectures, allowing it to scale across industries such as healthcare, finance, and government. By eliminating the need for technical SQL expertise, it empowers non-technical users to retrieve data efficiently, reducing bottlenecks in analytics and decision-making. Its contextual awareness, schema-based customization, and dynamic query generation set it apart from existing tools. SQLQueryAI contributes to national technological advancement by enhancing productivity, reducing costs, and promoting inclusive data access making it a strategic asset for U.S. enterprises and public institutions.

Problem Statement

Across industries in the United States, organizations face a persistent challenge in accessing structured data from complex databases. Writing SQL queries a prerequisite for data retrieval is a manual, time-consuming, and error-prone process that demands technical expertise. This creates a bottleneck for non-technical users such as business analysts, product managers, and domain experts, who often possess critical contextual knowledge but lack SQL proficiency. As a result, data access becomes centralized within IT teams, slowing down decision-making, increasing operational costs, and reducing agility.

The problem intensifies during large-scale data migration projects, regulatory reporting, and analytics initiatives, where thousands of queries must be written to validate, extract, and transform data. In domains like insurance, healthcare, and finance, where data complexity is high and compliance is critical, manual query writing introduces risks of inaccuracies, delays, and rework. Moreover, existing tools in the market fail to address contextual nuances, struggle with complex joins, and often produce incorrect or incomplete SQL statements.

This lack of accessible, accurate, and scalable query generation solutions limits innovation, increases dependency on technical resources, and hinders organizations from fully leveraging their data assets. The absence of a domain-agnostic, intelligent system that can translate natural language into precise SQL queries represents a significant gap in enterprise data management. Addressing this problem is essential not only for operational efficiency but also for enabling inclusive, data-driven decision-making across sectors vital to the U.S. economy.

Background and Method of Approach

In today's data-driven economy, organizations across sector insurance, healthcare, finance, and government rely on structured data to make decisions, ensure compliance, and drive innovation. However, accessing this data often requires writing SQL queries, a task that demands technical expertise and familiarity with complex database schemas. This creates a barrier for non-technical users and slows down analytics, reporting, and operational workflows.

The challenge intensifies during data migration, regulatory audits, and analytics initiatives. For example, in the insurance sector, migrating legacy policy data to modern platforms involves thousands of SQL queries tailored to specific business scenarios. These queries must account for hierarchical relationships, multi-layered coverage structures, and domain-specific terminology. Manual query writing is time-consuming and error-prone, leading to inefficiencies and increased costs.

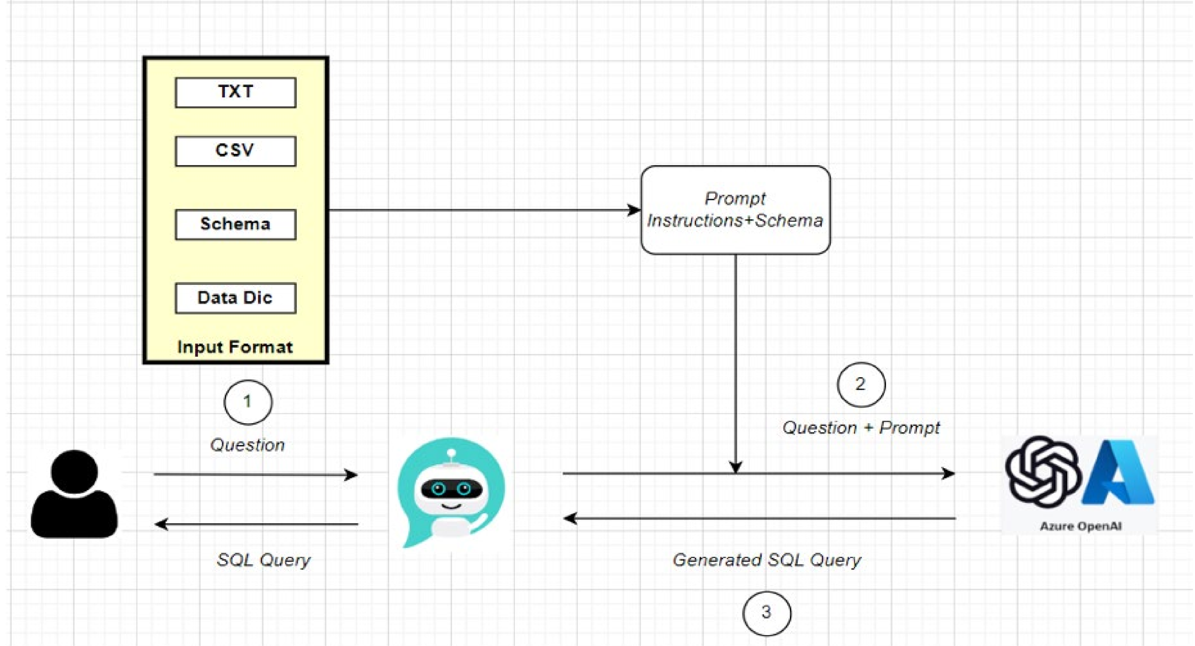
Existing tools like AskYourDatabase and SQL Chat attempt to simplify query generation but fall short due to limited contextual awareness, poor handling of complex joins, and lack of domain customization. To address this gap, Sharmila Devi Chandariah developed **SQLQueryAI**, a Generative AI-powered solution that translates natural language into accurate SQL queries. It is domain-agnostic, scalable, and customizable, enabling organizations to democratize data access and reduce manual effort.

SQLQueryAI supports two architectural approaches:

1. **Schema Based Approach**
   Designed for simple queries, users input natural language into a chatbot UI. The system combines this input with schema and instructions, which are processed by a Large Language Model (LLM) to generate SQL.

   Use Cases: Simple queries, small databases, basic user expertise

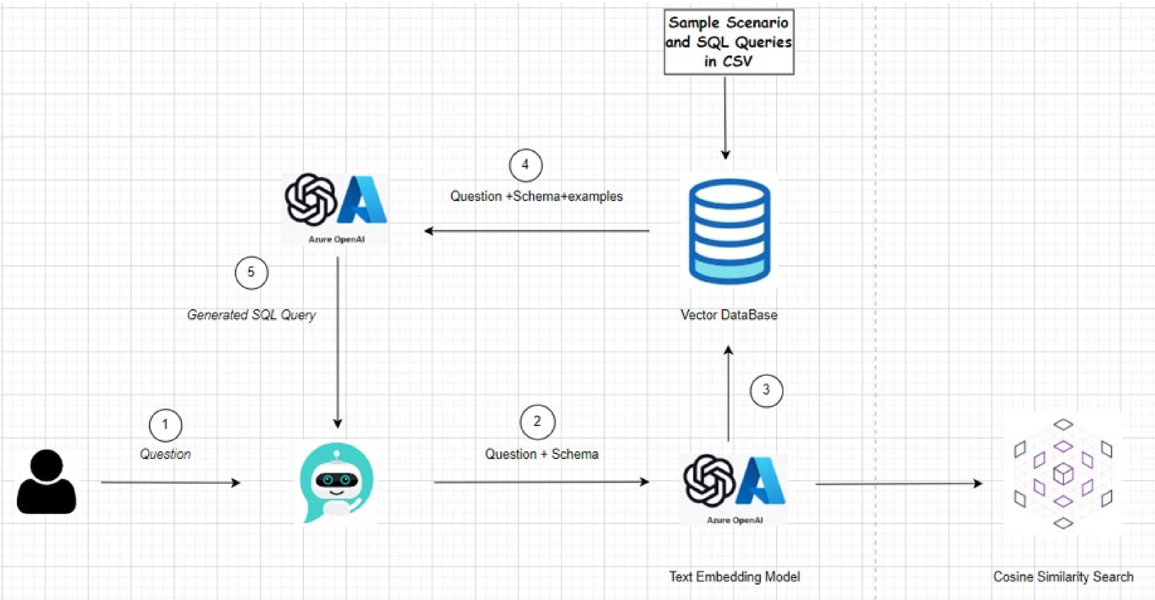
   Advantages: Quick setup, low cost, fast prototyping



2. Retrieval Augmented Generation (RAG) Approach
Ideal for complex queries and domain-specific logic, RAG integrates sample queries and schema context into a vector database. Using cosine similarity, it retrieves relevant embeddings to augment the LLM prompt.

Use Cases: Complex joins, healthcare/insurance logic, non-technical users

Advantages: High accuracy, scalable, context-aware



Insurance Domain Customization
At Fortune 100 company, SQLQueryAI was customized as *IntelliQuery* for Property & Casualty insurance. Enhancements included entity recognition, scenario libraries, and contextual understanding of terms like "claim reserves" and "premium calculations."

Evaluation and Impact
Tested across 30 scenarios, SQLQueryAI achieved 93.33% accuracy. It reduced manual effort by 94%, saving $600K annually for one team, and cut SQL errors by 99%. These results demonstrate its reliability, scalability, and national significance.

Results

The implementation of SQLQueryAI at Fortune 100 company demonstrated significant operational and financial impact. The model was tested across 30 real-world scenarios ranging from simple to highly complex SQL queries. It achieved an overall accuracy rate of **93.33%.**The remaining 6.67% were corrected after prompt refinement, showcasing the model's adaptability and learning capability.

SQLQueryAI reduced manual SQL writing effort by **94%**, transforming tasks that previously took 8 hours into just 5 minutes. This automation led to an estimated **$600,000 in annual savings** for a single team, with potential for multi-million-dollar savings across departments. Additionally, the tool achieved a **99% reduction in SQL errors**, improving data quality and minimizing rework.

These results validate SQLQueryAI' s effectiveness in enhancing productivity, accuracy, and cost-efficiency. Its success in the insurance domain highlights its scalability and relevance across other industries, making it a nationally significant innovation in enterprise data access.
.

Conclusion

SQLQueryAI represents a transformative advancement in enterprise data access and automation. By enabling natural language to SQL conversion, it bridges the gap between business users and complex databases, democratizing data retrieval across industries. Its dual architecture such as Basic and RAG offers flexibility for both simple and complex use cases, while domain customization ensures relevance and precision.

The successful deployment at Fortune 100 company, where SQLQueryAI was branded as *IntelliQuery*, highlights its real-world impact: 93.33% accuracy, 94% effort reduction, and $600K in annual savings for a single team. These results underscore the model's scalability, reliability, and cost-effectiveness.

Beyond technical innovation, SQLQueryAI contributes to national productivity by reducing operational bottlenecks, empowering non-technical users, and enhancing data-driven decision-making. Its domain-agnostic design makes it applicable across sectors such as healthcare, finance, and government, positioning it as a strategic asset for U.S. enterprises.

As organizations increasingly adopt AI for digital transformation, SQLQueryAI stands out as a pioneering solution with measurable benefits and broad societal relevance.

Real Time Communications Conference and Expo at Illinois Tech
IEEE International Conference
IEEE RTC 2024 - October 8-10, 2024 - IIT, Chicago, IL, USA