

DATA VOICE AI: A Conversational Business Intelligence Assistant Using RAG

Vihar K S Gowda

Department of Computer Science & Engineering (Data Science)
Sri Siddhartha Institute of Technology

Nandan U

Department of Computer Science & Engineering (Data Science)
Sri Siddhartha Institute of Technology

Ramesh E

Department of Computer Science & Engineering (Data Science)
Sri Siddhartha Institute of Technology

N Varshith

Department of Computer Science & Engineering (Data Science)
Sri Siddhartha Institute of Technology

Abstract - This paper presents an intelligent conversational Business Intelligence platform for analytical interaction with organizational databases using Natural Language Processing and Retrieval-Augmented Generation. The proposed system converts natural language business queries into executable SQL statements using GPT-4o-mini and schema-aware prompting techniques. FAISS-based semantic retrieval improves contextual understanding during query processing, while FastAPI manages backend analytical operations. The platform supports multi-database querying across Sales, Finance, Human Resources, and Inventory domains through a unified dashboard developed using HTML, CSS, and JavaScript. Dynamic visualization modules generate KPI metrics, analytical tables, bar charts, line charts, pie charts, and doughnut charts based on retrieved results. The system also supports voice-enabled analytical querying using browser-based speech recognition and provides export functionality for CSV and JSON reports. The developed platform simplifies database interaction for non-technical users and demonstrates the practical integration of conversational AI and Business Intelligence technologies for modern analytical systems.

Keywords: *Conversational Business Intelligence, Retrieval-Augmented Generation (RAG), Natural Language Processing (NLP), Natural Language to SQL, Data Visualization, FastAPI, FAISS, Conversational AI.*

1. Introduction

The rapid growth of digital transformation has significantly increased the volume of organizational data generated across domains such as finance, sales, inventory management, and human resources. Modern enterprises depend heavily on Business Intelligence systems to analyze operational data, monitor organizational performance, and support strategic decision-making processes. Traditional analytical platforms provide dashboards, reports, and SQL-based querying mechanisms for retrieving business insights. However, such systems often require technical knowledge of database querying, making analytical interaction difficult for non-technical users.

Conventional Business Intelligence systems rely heavily on manually written SQL queries and predefined dashboards for analytical operations. Although these approaches provide flexibility, they also create dependency on technical personnel for report generation and database querying. This limitation reduces accessibility and slows analytical decision-making in organizational environments.

Recent advancements in Artificial Intelligence, Large Language Models, and Natural Language Processing have enabled conversational analytical systems capable of understanding user queries and dynamically generating structured responses. Natural Language to SQL systems

simplifies database interaction by allowing users to communicate with databases using conversational language instead of manually writing SQL statements. However, many existing systems suffer from limitations such as poor contextual understanding, limited visualization support, lack of voice-enabled interaction, and insufficient multi-database analytical integration.

To address these challenges, this paper presents DATA VOICE AI, a conversational Business Intelligence assistant that integrates Retrieval-Augmented Generation, Natural Language Processing, dynamic SQL generation, and analytical visualization within a unified web-based platform. The proposed system supports conversational querying, voice-enabled interaction, dynamic visualization, and multi-database analytical operations using contextual retrieval and schema-aware prompting techniques.

2. Literature Survey

The rapid development of Artificial Intelligence, Natural Language Processing, and conversational analytical systems has significantly improved the capability of intelligent Business Intelligence platforms. Several researchers have contributed toward Natural Language to SQL conversion, contextual retrieval systems, conversational analytics, and semantic search frameworks. The following studies provide the theoretical foundation for the proposed conversational Business Intelligence system.

Victor Zhong, Caiming Xiong, and Richard Socher introduced Seq2SQL, a Natural Language to SQL framework capable of converting conversational user queries into executable SQL statements using deep learning and reinforcement learning techniques. Their work simplified database interaction for non-technical users and improved SQL generation efficiency. However, the framework mainly focused on SQL query generation and did not include contextual retrieval, voice interaction, or dynamic dashboard visualization.

Bailin Wang and other researchers proposed RatSQL, a relation-aware semantic parsing framework designed for complex relational databases. The system improved Natural Language to SQL conversion by analyzing relationships between tables, schemas, and user queries. The research demonstrated improved SQL generation performance for multi-table analytical operations. However, the framework primarily concentrated on

semantic parsing and did not support conversational dashboards or Retrieval-Augmented Generation.

Patrick Lewis, Ethan Perez, Aleksandra Piktus, and other researchers introduced Retrieval-Augmented Generation (RAG), which combines information retrieval mechanisms with language generation models to improve contextual understanding during conversational processing. The framework retrieved semantically relevant contextual information from external knowledge sources before response generation. Their work significantly improved contextual accuracy in Natural Language Processing tasks. However, the proposed approach mainly focused on general conversational systems rather than Business Intelligence applications.

Jeff Johnson, Matthijs Douze, and Hervé Jégou developed FAISS, a vector similarity search framework designed for efficient semantic retrieval from large embedding collections. FAISS improved similarity search performance and enabled scalable contextual retrieval for conversational AI systems. The framework became an important component in Retrieval-Augmented Generation pipelines and semantic search applications.

Recent advancements in Large Language Models have also improved conversational analytical systems and Natural Language to SQL frameworks. Transformer-based language models demonstrated the ability to understand user intent, generate analytical queries, and support conversational interaction across complex database environments. Prompt engineering and schema-aware contextual injection techniques further improved SQL generation quality and reduced ambiguity during analytical processing.

Although existing systems improve conversational querying, semantic retrieval, and analytical processing, several limitations still exist. Most traditional Business Intelligence systems depend heavily on predefined dashboards and manual SQL querying methods, making analytical interaction difficult for non-technical users. Many existing conversational analytical systems also lack secure query validation, contextual retrieval integration, voice-enabled interaction, and dynamic visualization support.

The proposed system addresses these limitations by integrating Retrieval-Augmented Generation, Natural Language to SQL conversion, FAISS-based contextual retrieval, dynamic visualization, voice-enabled interaction,

and multi-database analytical support within a unified conversational Business Intelligence platform.

3. Proposed Methodology

The proposed system was developed as a conversational Business Intelligence platform capable of converting natural language business queries into executable analytical operations. The platform combines Retrieval-Augmented Generation, Natural Language Processing, semantic retrieval, and dynamic visualization within a unified web-based environment. The system simplifies analytical interaction with organizational databases for users without SQL knowledge.

The frontend interface was developed using HTML, CSS, and JavaScript to provide an interactive dashboard for analytical querying and visualization. Users can submit text-based or voice-based queries, select databases, view generated SQL statements, analyze graphical outputs, and export analytical reports. FastAPI was used as the backend framework for handling API communication, query processing, SQL execution, and response management.

The Retrieval-Augmented Generation pipeline improves contextual understanding during conversational processing. User queries are converted into vector embeddings, and FAISS similarity search retrieves semantically relevant contextual information from indexed schema descriptions and documents. The retrieved contextual data is integrated into prompt construction before Natural Language to SQL conversion, improving analytical query interpretation and reducing ambiguity.

Schema-aware prompting techniques were implemented to improve SQL generation efficiency by providing only relevant schema information associated with the selected database. GPT-4o-mini dynamically generates executable SQL statements based on user intent, retrieved context, and schema details. Generated queries are validated before execution to improve database security and system reliability.

SQLite databases were used for storing organizational analytical data related to Sales, Finance, Human Resources, and Inventory Management. After successful SQL validation, the generated queries are executed on the selected database, and the retrieved results are forwarded to the visualization module for analytical presentation.

The proposed system also supports voice-enabled analytical interaction using browser-based speech recognition mechanisms. Voice queries provided by users are converted into text and forwarded to the conversational analytical pipeline for SQL generation and database processing. This feature improves accessibility and enables hands-free interaction with the Business Intelligence dashboard during analytical operations.

The visualization module converts retrieved database results into KPI cards, analytical tables, bar charts, line charts, pie charts, and doughnut charts. The system also supports export functionality for CSV and JSON reports directly from the dashboard interface. The modular architecture improves scalability, maintainability, and flexibility for future enhancements.

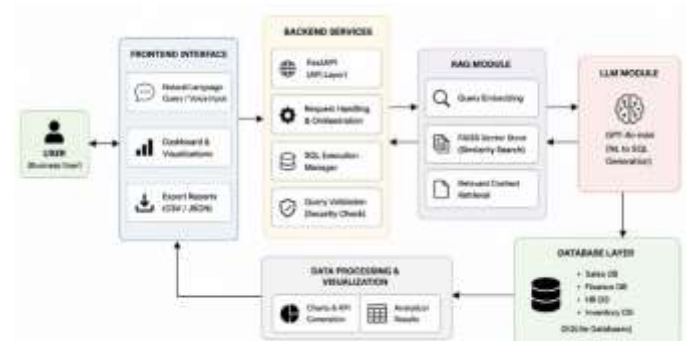


Fig. 1 System Architecture of DATA VOICE AI

4. Results and Discussion

The developed conversational Business Intelligence platform was tested using analytical queries across Sales, Finance, Human Resources, and Inventory databases. The evaluation mainly focused on conversational query handling, SQL generation, contextual retrieval, dashboard visualization, and multi-database analytical interaction. The obtained results demonstrated that the system successfully processed natural language business queries and generated meaningful analytical outputs through an interactive dashboard interface.

The conversational query processing module was tested using aggregation queries, grouped analytical operations, filtering conditions, and relational database queries. Queries related to sales analysis, expense categorization, employee statistics, and inventory monitoring were processed successfully through the Natural Language to SQL pipeline. The generated SQL statements produced accurate analytical outputs across different organizational databases.

The Retrieval-Augmented Generation mechanism improved contextual understanding during analytical processing. FAISS-based semantic retrieval identified relevant schema descriptions and contextual information before SQL generation. The retrieved contextual data reduced ambiguity during conversational interaction and improved analytical query interpretation for complex business requests.

The dashboard interface dynamically generated KPI cards, analytical tables, and graphical visualizations based on retrieved database results. Time-series analytical queries produced line chart visualizations, while grouped categorical queries generated bar chart and pie chart representations. Dynamic visualization simplified analytical interpretation and improved user interaction with organizational data.



Fig. 2 Conversational Business Intelligence Dashboard

The conversational interface also supported multi-database analytical interaction through a unified dashboard environment. Users were able to switch dynamically between Sales, Finance, Human Resources, and Inventory databases using the database selection module. Contextual prompts, query suggestions, and schema information were updated automatically during database switching operations.

Voice-enabled analytical querying was implemented using browser-based speech recognition. Spoken business queries were converted into text and forwarded to the backend analytical pipeline for SQL generation and database execution. The conversational interface improved accessibility and reduced dependency on manual query typing during analytical interaction.

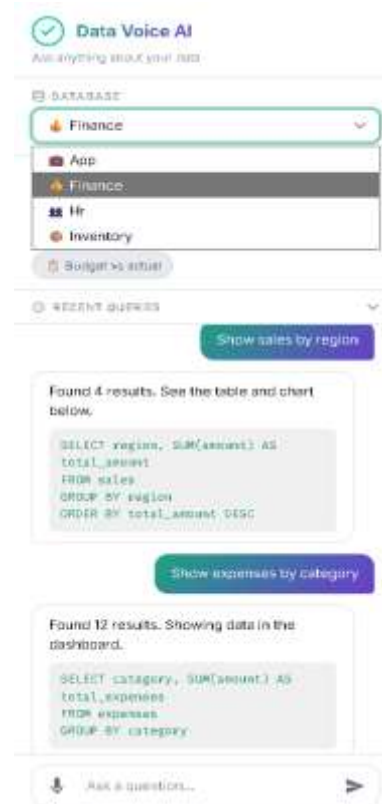


Fig. 3 Conversational Query Interface with Multi-Database Support

The export functionality successfully generated downloadable CSV and JSON reports from retrieved analytical outputs. Users were able to export analytical tables and processed business data directly from the dashboard interface for external reporting and analytical usage.

The implementation results confirmed that the proposed conversational Business Intelligence platform successfully integrates Retrieval-Augmented Generation, contextual retrieval, Natural Language to SQL conversion, voice-enabled interaction, and dynamic analytical visualization within a unified analytical framework. The developed system improves accessibility for non-technical users and simplifies interaction with organizational databases through conversational analytical processing.

5. Advantages

The proposed conversational Business Intelligence platform provides several advantages for analytical interaction and organizational data processing. The system simplifies database querying by allowing users to interact with databases using natural language instead of manually writing SQL statements. This reduces dependency on

technical expertise and improves accessibility for non-technical users.

The integration of Retrieval-Augmented Generation improves contextual understanding during analytical query processing. FAISS-based semantic retrieval helps the system identify relevant schema information and contextual descriptions before SQL generation, improving query interpretation and reducing ambiguity during conversational interaction. The platform supports analytical operations across multiple organizational databases including Sales, Finance, Human Resources, and Inventory Management within a unified dashboard environment. Multi-database support improves flexibility and allows users to perform analytical operations across different business domains without switching applications.

Dynamic visualization modules improve analytical interpretation by generating KPI cards, analytical tables, bar charts, line charts, pie charts, and doughnut charts automatically based on retrieved database results. Interactive visualization improves dashboard usability and simplifies understanding of business insights. The voice-enabled querying mechanism improves accessibility by supporting hands-free conversational interaction through browser-based speech recognition. The export functionality also improves report portability by allowing users to download analytical outputs in CSV and JSON formats. The modular architecture improves scalability, maintainability, and flexibility for future enhancements. Individual modules such as conversational processing, contextual retrieval, SQL generation, database execution, and visualization can be modified independently without affecting the overall workflow of the system.

6. Future Enhancements

Although the developed system successfully demonstrates conversational Business Intelligence functionality, several improvements can further enhance the analytical capabilities and scalability of the platform. Future implementations can integrate advanced Large Language Models for improved conversational understanding, contextual reasoning, and SQL generation accuracy. Support for multi-turn conversational interaction can also be added to maintain contextual continuity during follow-up analytical queries and extended user conversations.

The platform can be further extended to support enterprise-level database systems such as MySQL, PostgreSQL, and MongoDB for handling large-scale organizational analytical operations. Additional security mechanisms

including user authentication, role-based access control, and multi-user session management can improve secure enterprise deployment. Real-time analytical processing and live dashboard updates can also enhance the effectiveness of the system for continuously changing business environments.

Future enhancements may also include mobile application integration and cloud-based deployment for scalable and portable analytical access. The reporting module can be improved by supporting PDF and Excel report generation along with automated reporting features. Multilingual conversational querying and enhanced Retrieval-Augmented Generation using larger contextual datasets can further improve user interaction and analytical accuracy within the conversational Business Intelligence platform.

7. Conclusion

The presented work demonstrates the development of a conversational Business Intelligence platform capable of simplifying analytical interaction with organizational databases using natural language queries. The system integrates Retrieval-Augmented Generation, Natural Language Processing, semantic retrieval, SQL generation, and dynamic visualization within a unified web-based environment. The developed platform reduces dependency on manual SQL writing and improves accessibility for users without technical database knowledge.

The implemented architecture successfully supports analytical operations across Sales, Finance, Human Resources, and Inventory databases through conversational interaction. Context-aware retrieval using FAISS similarity search improved query interpretation during SQL generation, while schema-aware prompting enhanced analytical processing efficiency. The generated SQL statements were executed successfully on SQLite databases, and the retrieved results were presented through KPI cards, analytical tables, and graphical visualizations. The platform also demonstrated voice-enabled analytical querying, multi-database interaction, and export functionality for analytical reports. The obtained implementation results confirmed that conversational AI and Retrieval-Augmented Generation can effectively improve Business Intelligence interaction and simplify organizational analytical workflows. The developed system provides a practical foundation for future AI-driven conversational analytical platforms and enterprise Business Intelligence applications.

References

- [1] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser and Illia Polosukhin, "Attention Is All You Need," *Advances in Neural Information Processing Systems (NeurIPS)*, Vol. 30, 2017.
- [2] Tom B. Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah et al., "Language Models are Few-Shot Learners," *Advances in Neural Information Processing Systems (NeurIPS)*, Vol. 33, 2020, pp. 1877–1901.
- [3] Patrick Lewis, Ethan Perez, Aleksandra Piktus, Fabio Petroni et al., "Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks," *Advances in Neural Information Processing Systems (NeurIPS)*, Vol. 33, 2020, pp. 9459–9474.
- [4] Victor Zhong, Caiming Xiong and Richard Socher, "Seq2SQL: Generating Structured Queries from Natural Language Using Reinforcement Learning," *arXiv preprint arXiv:1709.00103*, 2017.
- [5] Bailin Wang, Richard Shin, Xiaodong Liu, Oleksandr Polozov and Matthew Richardson, "RAT-SQL: Relation-Aware Schema Encoding and Linking for Text-to-SQL Parsers," *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 2020, pp. 7567–7578.
- [6] Jeff Johnson, Matthijs Douze and Hervé Jégou, "Billion-scale Similarity Search with GPUs," *IEEE Transactions on Big Data*, Vol. 7, No. 3, 2021, pp. 535–547.
- [7] Jacob Devlin, Ming-Wei Chang, Kenton Lee and Kristina Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," *Proceedings of NAACL-HLT*, 2019, pp. 4171–4186.
- [8] Nils Reimers and Iryna Gurevych, "Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks," *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing*, 2019, pp. 3982–3992.
- [9] Sebastián Ramírez, "FastAPI Framework Documentation," *FastAPI Official Documentation*, 2024.
- [10] SQLite Development Team, "SQLite Documentation," *SQLite Official Documentation*, 2024.
- [11] OpenAI, "GPT-4 Technical Report," *OpenAI Research Publication*, 2023.
- [12] Chart.js Contributors, "Chart.js Documentation for Interactive Data Visualization," *Chart.js Official Documentation*, 2024.
- [13] Christopher D. Manning, Prabhakar Raghavan and Hinrich Schütze, *Introduction to Information Retrieval*, Cambridge University Press, 2008.
- [14] Ian Goodfellow, Yoshua Bengio and Aaron Courville, *Deep Learning*, MIT Press, Cambridge, Massachusetts, 2016.
- [15] Thomas H. Davenport and Jeanne G. Harris, *Competing on Analytics: The New Science of Winning*, Harvard Business Review Press, Boston,