

Original Article

Conversational Analytics: Turning Enterprise Data into Instant Business Insights with LLMs

***Sandeep Rao Udupi**

Lead Data Engineer, GPS IT Solutions.

Abstract:

Conversational analytics is transforming how business users interact with enterprise data by enabling natural language queries instead of complex SQL, reducing reliance on technical expertise and accelerating decision-making. However, traditional approaches often produce inconsistent or unreliable results due to ambiguous schemas, lack of standardized metric definitions, and limited understanding of business context. These challenges increase dependency on data engineering teams to build pipelines, curated views, and reports, slowing down insight generation. This paper proposes a modern approach that combines a governed semantic model with large language models (LLMs) to deliver accurate and reliable conversational analytics. By positioning the semantic model closer to the data layer on top of curated fact and dimension tables, it standardizes metrics, relationships, and business logic, enabling LLMs to generate context-aware SQL with high accuracy. This approach reduces turnaround time, minimizes dependency on data engineering teams, and enables scalable self-service analytics while maintaining governance, consistency, and a single source of truth across the organization.

Keywords:

LLMs, RAGs, Semantic Model, Semantic Views, Snowflake, Context AI, Cortex Analysts, Medallion architecture, Streamlit, Teams.

1. Introduction

In modern data-driven enterprises, the volume and complexity of data continue to grow rapidly, making it increasingly challenging for business users to access timely and accurate insights. Although contemporary data platforms and business intelligence tools have improved data availability, extracting meaningful information still requires expertise in SQL, data schemas, and system structures. This dependency on technical teams introduces bottlenecks, delays decision-making, and limits the adoption of true self-service analytics.

Conversational analytics has emerged as a promising approach to address these limitations by enabling natural language interaction with data. Leveraging advances in large language models (LLMs), users can query data using plain language and receive insights without requiring technical knowledge. However, early implementations have demonstrated significant shortcomings. When LLMs are applied directly to raw schemas, they often generate inaccurate or inconsistent queries due to ambiguous data structures, lack of standardized metric definitions, and insufficient business context. These limitations reduce trust in generated insights and hinder enterprise adoption. To overcome these challenges, a unified architectural approach is required that combines LLM capabilities



with a governed semantic layer. By positioning the semantic model closer to the data platform, built on curated fact and dimension tables, organizations can establish standardized definitions for metrics, relationships, and business logic. This enables LLMs to interpret user intent accurately and generate context-aware, reliable SQL queries aligned with enterprise data standards.

Building on this foundation, conversational analytics systems are further enhanced through agentic AI workflows that orchestrate multiple stages, including intent interpretation, semantic context retrieval, query generation, validation, and execution. These systems incorporate iterative reasoning, feedback loops, and guardrails to improve accuracy, enforce governance, and reduce hallucinations in generated queries. This multi-step approach enables robust handling of complex analytical scenarios beyond simple query generation. The integration of semantic modeling with agentic AI establishes a scalable and production-ready framework for enterprise analytics. It empowers business users with self-service capabilities while maintaining governance through standardized metrics, role-based access controls, and controlled data access. This paper presents a structured approach to conversational analytics that bridges business intent with data execution, enabling accurate, consistent, and timely insights for improved organizational decision-making.

2. Challenges in Traditional Analytical System

Traditional analytics approaches in enterprises rely heavily on SQL expertise and specialized tools, requiring business users to depend on data engineers and analysts to extract insights. This model introduces multiple handoffs between teams, leading to delays, inefficiencies, and slower decision-making. As a result, organizations often experience increased latency in accessing insights and reduced operational efficiency, limiting their ability to respond quickly to changing business needs. Furthermore, as data volumes and complexity grow, these challenges become more pronounced, making traditional analytics models increasingly unsustainable for modern, data-driven organizations.

2.1. Dependency on Technical Teams

Traditional analytics environments are built around role-based workflows, where business users rely on analysts and engineers to fulfill data requests. Each stage involves interpretation, data preparation, validation, and delivery, creating multiple dependency layers. This reliance not only increases operational overhead but also creates a bottleneck where technical teams effectively become gatekeepers of data access. As illustrated in **Figure 1**, the analytics workflow follows a sequential process in which a business user submits a request that is interpreted by a data analyst, implemented by a data engineer, and executed on the data platform. Each transition between roles introduces dependency and delay, making the overall process inefficient and difficult to scale.

2.1.1. Multi-Step Request Workflow

Each request passes through several stages, including requirement gathering, query development, data validation, and result interpretation. This multi-step workflow often requires back-and-forth communication between business users and technical teams, introducing delays and increasing the likelihood of misinterpretation. The workflow shown in **Figure 1**, business intent must be translated into technical logic by analysts, followed by data preparation and pipeline execution by engineers. This layered translation process increases complexity and reduces the speed at which insights can be delivered.

2.1.2. Impact on Decision-Making

These dependencies often result in turnaround times of 2 to 5 days. In critical industries such as financial services, such delays can have significant consequences, including missed risk signals, delayed fraud detection, and challenges in regulatory compliance. As depicted in **Figure 1**, the cumulative effect of dependency and handoff delays across multiple roles leads to slow insight delivery. Organizations that cannot access insights in a timely manner risk losing their competitive advantage, as decision-making becomes reactive rather than proactive. This limitation highlights the need for more efficient, direct, and scalable approaches to data access.

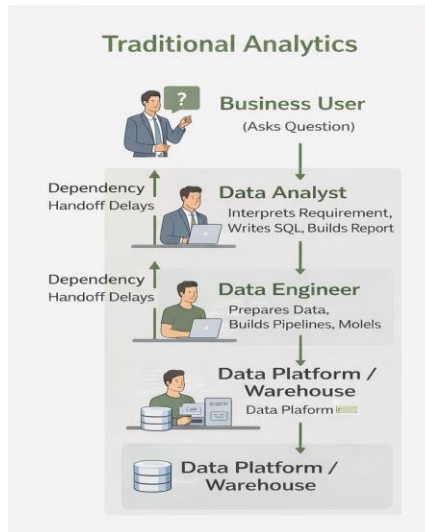


Figure 1. Request Flow in Traditional Analytics

2.2. Fragmented Semantic Layer

The semantic layer plays a critical role in defining business metrics, relationships, and governance rules within an enterprise data ecosystem. However, in traditional architecture, it is often implemented inconsistently across different tools and layers. Instead of being centralized within the data platform, business logic is frequently embedded within individual business intelligence (BI) tools such as Power BI, Tableau, or Looker.

As illustrated in Figure 2, the semantic layer is positioned closer to the Business Intelligence layer rather than the data platform, with multiple downstream applications independently consuming and redefining business logic. Each tool or application maintains its own interpretation of entities, metrics, and queries, leading to fragmented implementations of the same underlying data. This decentralized approach results in duplication of logic and lack of standardization across the organization.

2.2.1. Inconsistency in Business Metrics

When semantic definitions are distributed across multiple tools, different teams often define and calculate key business metrics independently. Indicators such as revenue, customer count, and annual recurring revenue (ARR) may be derived using varying filters, aggregation rules, or assumptions. As shown in Figure 2, multiple consuming layers—including BI tools, dashboards, and custom applications—interact with separate interpretations of the semantic layer. This leads to inconsistencies in reported values across reports and systems, reducing trust in data and increasing the need for reconciliation efforts. Over time, these discrepancies create confusion among stakeholders and hinder effective, data-driven decision-making.

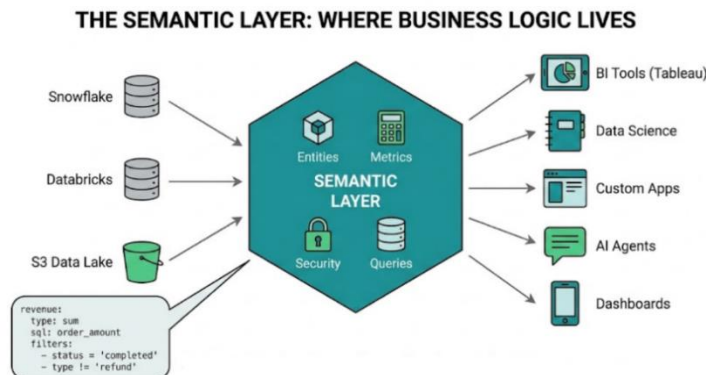


Figure 2. Semantic Layer Closer to BI Tools

3. Conversational Analytics as a Solution

Conversational analytics represents a transformative paradigm in enterprise data access, enabling users to interact with data using natural language rather than complex query languages. By leveraging AI-driven semantic understanding, it addresses key limitations of traditional analytics, including delays caused by sequential team handoffs and inconsistencies arising from fragmented semantic definitions across reporting layers. This approach introduces a centralized semantic model, tightly integrated with the data warehouse, allowing consuming applications to operate on a consistent representation of business metrics. Furthermore, the integration of large language models (LLMs) enhances this paradigm by enabling intuitive query interpretation, context-aware reasoning, and dynamic insight generation. As a result, conversational analytics enables scalable self-service analytics while maintaining governance and consistency across the enterprise.

3.1. Transformation of Analytics Workflow

By integrating natural language interfaces with a centralized semantic layer and large language models (LLMs), the analytics workflow becomes significantly more efficient, scalable, and responsive. As illustrated in Figure 3, the modern approach replaces the fragmented, multi-step workflow with a unified pipeline where user queries are directly interpreted, processed, and executed within a governed data environment. This transformation reduces operational complexity and enables organizations to shift from delayed, reactive analytics to real-time, proactive decision-making.

3.1.1. Direct User-to-Data Interaction

Unlike traditional systems that rely heavily on analysts and engineers, conversational analytics enables business users to directly interact with the data platform using natural language. This eliminates the need for technical translation of business requirements into SQL queries. In Figure 3, the business user initiates the interaction by asking a question in natural language, which is directly processed by the LLM-driven system. The integration of a centralized and governed semantic layer ensures that user queries are accurately mapped to business definitions, enabling seamless interaction with the underlying data warehouse.

3.1.2. Elimination of Handoffs

The conversational interface, powered by LLMs and supported by the semantic layer, acts as an intelligent intermediary that bridges business intent with data execution. This direct interaction eliminates delays caused by role-based handoffs and reduces the need for iterative communication cycles. Additionally, the centralized semantic layer ensures that all queries are interpreted using consistent business definitions, further improving accuracy and trust in the generated insights. By removing these dependencies, organizations can significantly reduce the time required to generate insights from multiple days to near real-time while also improving scalability and enabling true self-service analytics across business functions.

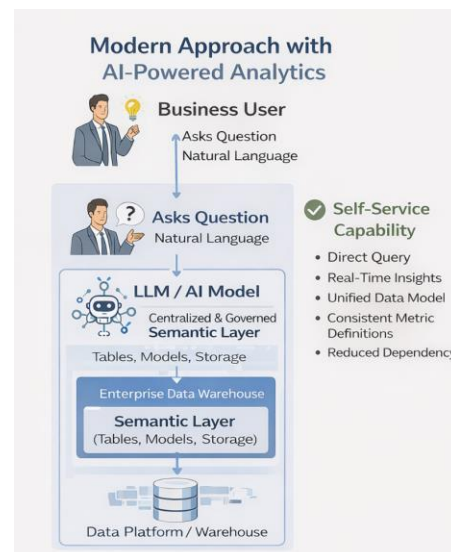


Figure 3. Request Flow in Conversational Analytics

4. System Design of Conversational Analytics

Building a conversational analytics system requires a fundamental shift from traditional, tool-centric BI reporting toward an integrated, AI-driven data access paradigm. Instead of relying on sequential workflows and fragmented semantic models closer to BI tools, the proposed architecture is centered around a centralized semantic model tightly coupled with the enterprise data warehouse and enhanced by large language models (LLMs). This architecture enables seamless interaction between business users and enterprise data through a series of logically connected layers, ensuring accurate interpretation of user intent, governed query execution, and meaningful insight generation. As illustrated in Figure 4, the system establishes a direct and intelligent data access pipeline that bridges business intent with data execution.

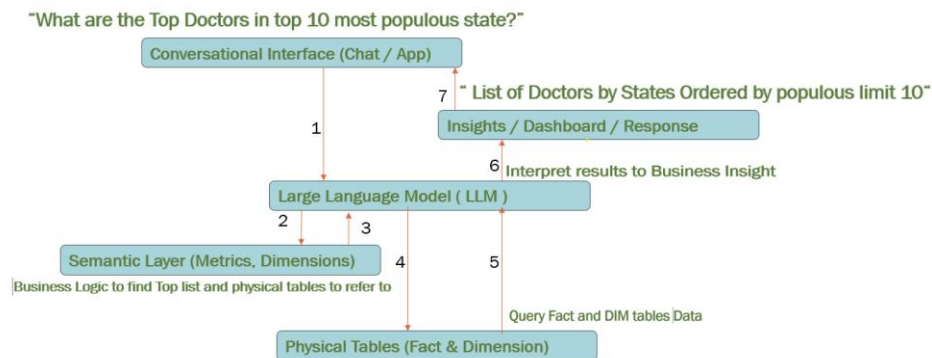


Figure 4. Data Flow between User & Layers

4.1. User Interaction Layer

The user interaction layer serves as the entry point for the conversational analytics system, enabling business users to interact with data using natural language.

4.1.1. Conversational Interface

This interface can be implemented through web-based applications such as Streamlit or integrated into enterprise collaboration platforms such as Microsoft Teams. Users can express business queries in plain language without requiring SQL knowledge or familiarity with underlying data schemas. This abstraction significantly improves accessibility and usability across business functions.

4.2. Intelligence and Interpretation Layer

This layer is responsible for understanding and interpreting user queries using advanced LLMs such as GPT or Claude

4.2.1. Intent Recognition and Context Extraction

The system analyzes user input to identify intent, extract relevant entities, and map business terminology to underlying data concepts. The LLM captures key elements such as metrics, dimensions, filters, and timeframes while resolving ambiguities in user queries. To improve accuracy, the system retrieves relevant metadata from the semantic layer. By incorporating semantic definitions, relationships hallucinations and improving reliability, and business rules, the system ensures that query interpretation is aligned with enterprise data standards, reducing

4.3. Semantic Mapping and Query Generation Layer

This layer translates interpreted user intent into structured queries aligned with the centralized semantic model.

4.3.1. Intent Recognition and Context Extraction

User queries are mapped to predefined entities such as fact tables, dimension tables, and standardized metrics. This abstraction eliminates dependency on physical schemas and ensures consistency in how business concepts are interpreted across the organization. The system generates optimized SQL queries using predefined templates and logic. Generated queries incorporate joins, filters, aggregations, and time-based calculations while enforcing governance policies such as role-based access control and data masking.

4.4. Data Execution and Processing Layer

This layer executes generated queries on the enterprise data warehouse.

4.4.1. Intent Recognition and Context Extraction

Modern cloud data platforms provide elastic compute, workload isolation, and query optimization capabilities. Techniques such as cache, partition pruning, and query optimization ensure efficient processing of large-scale datasets.

4.5. Insight Generation and Delivery Layer

This layer converts processed query results into actionable insights and delivers them to users.

4.5.1. Natural Language Insight Generation

LLMs transform structured outputs into human-readable explanations. The system highlights key trends, anomalies, and business drivers aligned with the user's query. Insights are delivered in multiple formats to enhance usability. Outputs include conversational responses, tables, and visualizations, with support for iterative query refinement and deeper exploration within the same session.

5. Conclusion

In conclusion, the proposed conversational analytics architecture represents a significant shift from traditional, tool-driven data access to an intelligent, user-centric analytics paradigm. By integrating a governed semantic layer with large language models and scalable data platforms, the approach eliminates technical barriers while ensuring data consistency, governance, and accuracy. The architecture enables direct interaction between business users and enterprise data, significantly reducing dependency on technical teams and accelerating insight generation from days to near real-time. By centralizing semantic definitions, it establishes a single source of truth, ensuring consistent interpretation of business metrics across all users and applications.

Furthermore, the incorporation of agentic workflows and context-aware query generation enhances reliability, reduces ambiguity, and supports complex analytical use cases. This positions conversational analytics as a scalable and production-ready solution for modern enterprises. Overall, this work demonstrates that conversational analytics is not merely an incremental improvement but a foundational transformation in how organizations access and utilize data. It enables faster, more accurate, and more accessible insights, ultimately bridging the gap between business intent and data-driven decision-making.

Looking ahead, further advancements in adaptive semantic modeling, explainable AI, and hybrid analytics across structured and unstructured data sources can enhance the robustness and applicability of such systems. Additionally, improvements in agent-driven orchestration, real-time data integration, and performance optimization will continue to strengthen scalability, efficiency, and user trust, enabling conversational analytics to evolve into a core capability of next-generation enterprise data platforms.

References

- [1] Liu, S., Xu, J., Tjangnaka, W., Semnani, S., Yu, C. J., & Lam, M. (2023). SUQL: Conversational Search over Structured and Unstructured Data with Large Language Models.
- [2] Zhu, J., Chen, L., Ke, X., Fang, Z., Li, T., Gao, Y., & Jensen, C. S. (2025). Beyond Relational: Semantic-Aware Multi-Modal Analytics with LLM-Native Query Optimization.
- [3] Towards Reliable Conversational Data Analytics. Proceedings of EDBT (Extending Database Technology Conference).
- [4] Martins, M. et al. (2025). Talking to Data: A Systematic Review of Conversational Agents for Visual Analytics.
- [5] Wang, X. et al. (2025). LLM and Agent-Driven Data Analysis.