AI Data Analyst

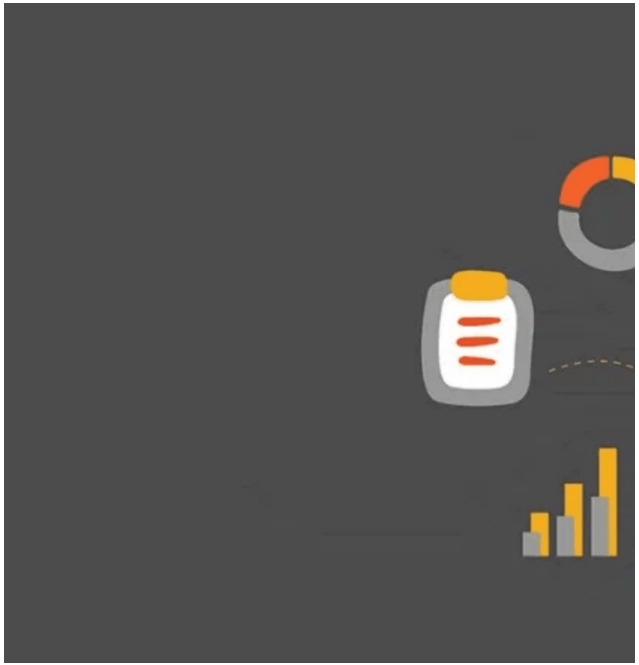
Empowering Data-Driven Decisions through Natural Language Processing

MIRACLE
SOFTWARE SYSTEMS

Traditional Data Analysis Faces Significant Technical Barriers

- **SQL Expertise Gap:** Business users often lack the technical skills to query complex databases directly.
- **Operational Bottlenecks:** High dependency on data teams leads to significant delays in decision-making.
- **Manual Errors:** Hand-written queries are prone to syntax errors and schema misunderstandings.

The Missing Link: A critical need exists for a bridge between human language and structured data.



MIRACLE SOFTWARE SYSTEMS

Seamless Natural Language to SQL Conversion

Intent Interpretation

Leverages advanced Large Language Models (LLMs) to accurately interpret user intent from natural language questions.

Context Awareness

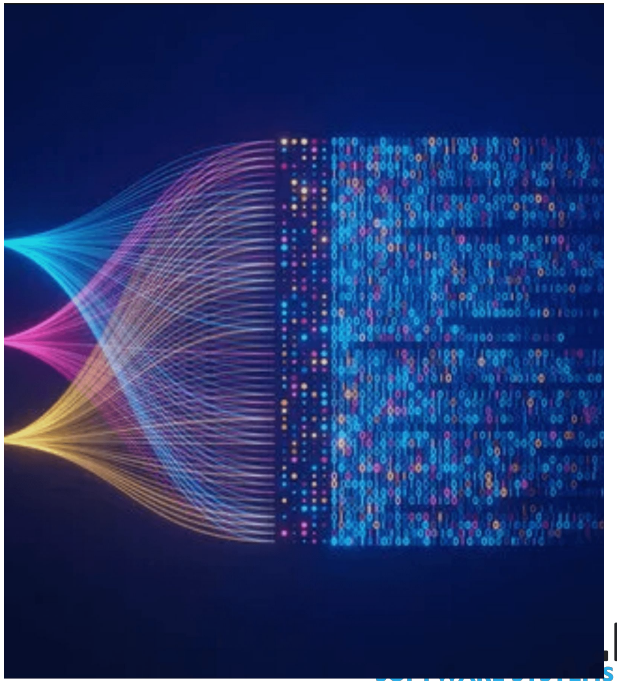
Automates the extraction of database schemas to provide the LLM with the necessary context for precise query generation.

Security First

Ensures data integrity by strictly executing only read-only SELECT queries, preventing unauthorized modifications.

Local Deployment

Designed for local execution, ensuring maximum data privacy, zero latency, and complete control over the environment.



Built on Robust and Modern Open-Source Technologies

Python 3.11+

The core foundation for **high-performance** data processing, offering extensive library support and seamless integration capabilities.

LangChain & SQLAlchemy

Providing the essential framework for **LLM orchestration** and standardized database interaction across various SQL dialects.

Ollama & DeepSeek-R1

Utilizing state-of-the-art **local LLMs** to ensure precise SQL generation while maintaining complete data sovereignty.

Streamlit

A sleek, **interactive frontend** that delivers a professional user experience with minimal development overhead.

MIRACLE SOFTWARE SYSTEMS

A Streamlined Four-Step Process from Question to Answer

STEP 01

Schema Extraction

Automatically inspects the SQLite database to map tables and columns for context.

STEP 02

Prompt Engineering

Combines the extracted schema with the user's question into a structured prompt.

STEP 03

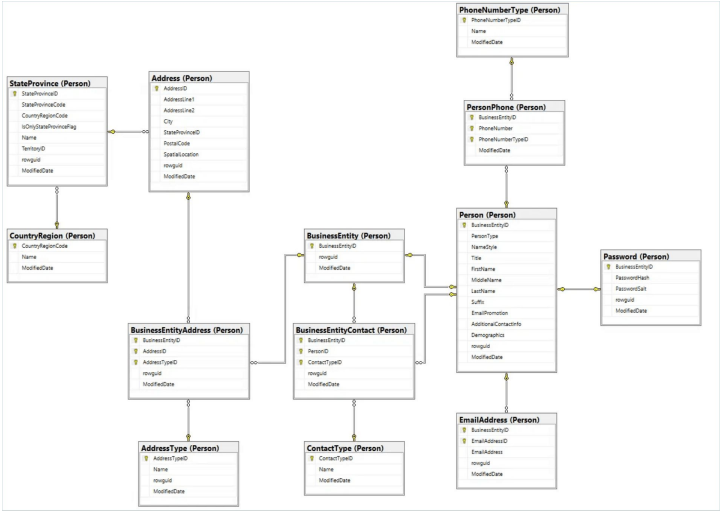
LLM Processing

The DeepSeek model generates a valid SQL query based on the provided context.

STEP 04

Execution & Display

The query runs against the database, and results are presented in the Streamlit UI.



MIRACLE SOFTWARE SYSTEMS

Precision-Engineered Prompts for Accurate Querying

01 / SYSTEM ROLE

Expert SQL Generator

System prompts explicitly define the LLM as an **Expert SQL Generator**, ensuring the model prioritizes technical accuracy and schema adherence.

02 / CONSTRAINTS

Strict Syntax Rules

Rigid constraints prevent the model from including preambles, explanations, or **unsupported SQL features**, resulting in clean, executable code.

03 / POST-PROCESSING

Automated Cleanup

Advanced regex-based processing removes **"Think" tags** and non-SQL artifacts, sanitizing the LLM response for direct database execution.

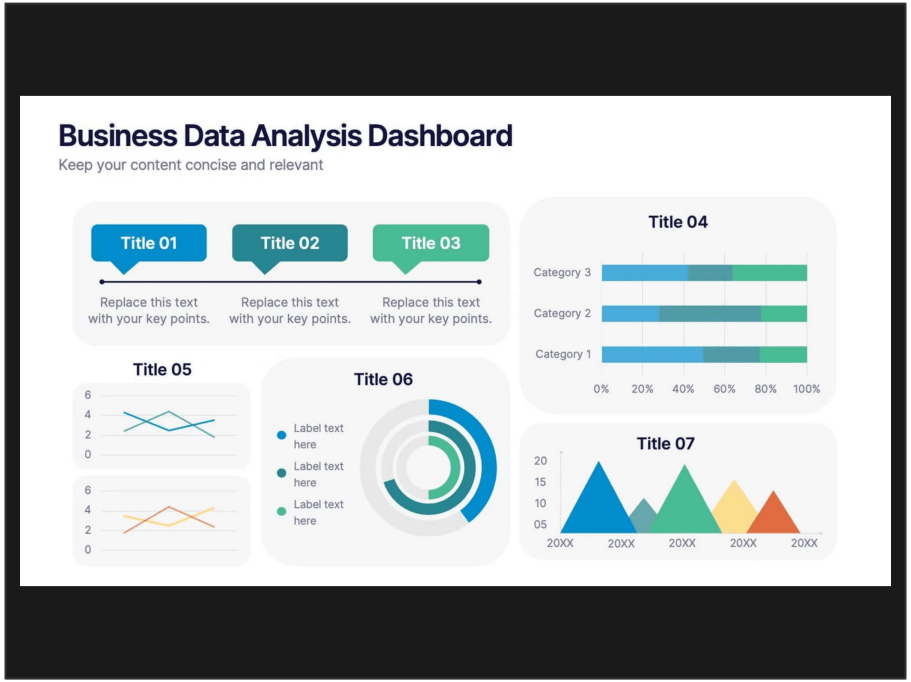
04 / OPTIMIZATION

Zero Temperature

Model temperature is locked at **0.0** to eliminate randomness, ensuring maximum consistency and reproducibility for every data query.

MIRACLE SOFTWARE SYSTEMS

Empowering Users with a Simple and Intuitive Frontend



Natural Language Input

Streamlit-based interface allows users to type questions as if they were talking to a colleague, removing technical friction.

Real-Time Feedback

Interactive elements like spinners and success notifications keep users engaged and informed during the analysis process.

Clear Data Presentation

Results are formatted with high legibility, making complex database outputs easy to digest and act upon instantly.

MIRACLE SOFTWARE SYSTEMS

Prioritizing Enterprise-Grade Security in Local Environments

01
Local Execution

All data processing and LLM inference occur on-premise. No sensitive data ever leaves your local network, eliminating external exposure risks.

02
Read-Only Access

The system is architected to execute only SELECT queries. Destructive operations like DROP, DELETE, or UPDATE are strictly prohibited at the code level.

03
Model Privacy

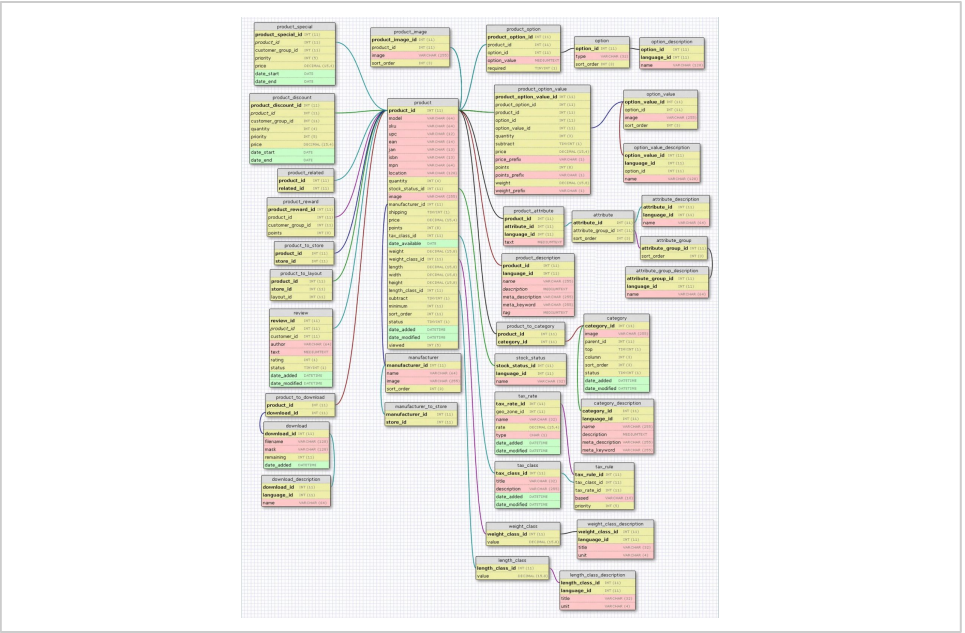
By utilizing Ollama for local model hosting, proprietary data is never used for training external models, ensuring complete intellectual property protection.

04
SQLite Integration

Leverages a lightweight, file-based database system that provides robust local storage without the security overhead of complex server-client architectures.

MIRACLE SOFTWARE SYSTEMS

Transforming Raw Amazon Data into Actionable Insights



THE QUERY

"Top 5 products by sales volume"

A simple natural-language request—no table names or joins required.

THE ACTION

Automated Schema Mapping

AI detects relevant tables and join keys and builds the SQL automatically.

THE RESULT

Instant Visualization

Structured results are returned, ready for analysis and visualization.

MIRACLE SOFTWARE SYSTEMS

Optimized for Speed and Future Growth

Efficient Schema Caching

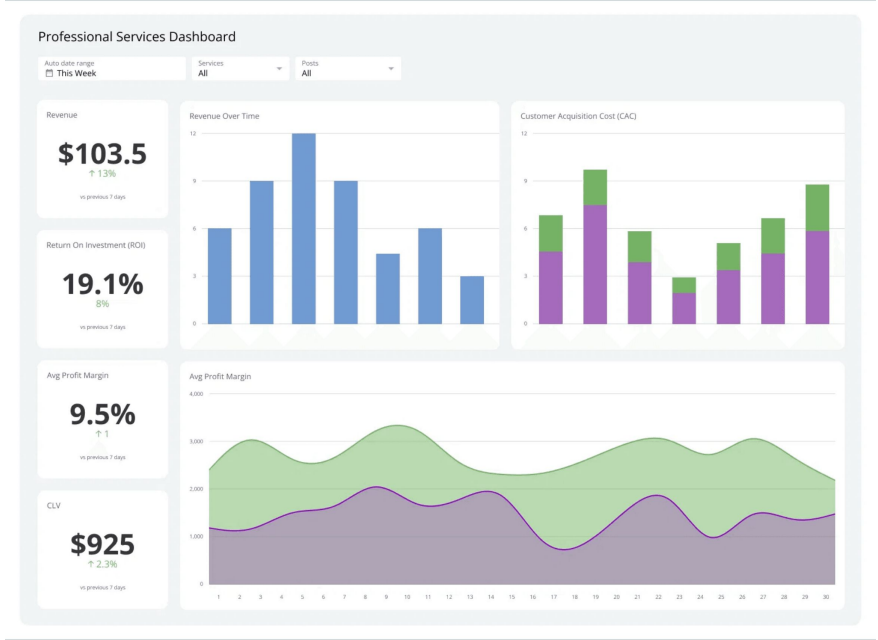
Reduces overhead for repeated queries by caching database metadata, ensuring near-instant response times for frequent users.

Flexible Model Support

Support for various Ollama models allows for scaling based on hardware capabilities, from lightweight to high-performance LLMs.

Modular Architecture

Modular code structure enables easy integration with other database types like PostgreSQL and MySQL for enterprise-wide deployment.



MIRACLE SOFTWARE SYSTEMS

Continuous Evolution of the AI Data Analyst

Phase 01

Data Visualization Integration

Native integration of Plotly and Altair libraries to automatically generate interactive charts and dashboards from query results.

Phase 02

Multi-turn Conversations

Support for contextual dialogue, allowing users to refine and drill down into data through natural, multi-step conversations.

Phase 03

Enhanced Error Handling

Advanced natural language explanations for query failures and automated suggestions for query refinement based on schema constraints.

Phase 04

Hybrid Cloud Expansion

Expansion to support hybrid deployment options, integrating with cloud-based LLM providers while maintaining local data security.

MIRACLE SOFTWARE SYSTEMS

The Future of Data Interaction is Natural

Democratization

AI Data Analyst 2.0 empowers every user to access insights without needing technical SQL expertise.

Efficiency

Reduces the technical burden on data teams, allowing them to focus on high-value strategic initiatives.

Security

A secure, private, and local solution designed to meet the rigorous demands of modern data challenges.

Questions & Answers

MIRACLE SOFTWARE SYSTEMS

THANK YOU

MIRACLE
SOFTWARE SYSTEMS

www.miraclesoft.com