Title: "ttyDB – Talk to Your Database: Natural Language Querying in Jupyter Notebooks"

Introduction

ttyDB is an interactive, Jupyter-notebook style tool that enables users to retrieve and explore data using plain English commands. By bridging natural-language processing with database connectivity, ttyDB removes the need for manual SQL writing, accelerating data analysis for both technical and non-technical stakeholders. The platform combines conversational AI, a robust query generation engine, and seamless result visualization within the familiar notebook environment.

Tech Stack

Python 3.11 Jupyter Notebook / JupyterLab LangChain framework for LLM orchestration OpenAI GPT-4 (or compatible open-source LLM) SQLAlchemy for database abstraction Pandas for data manipulation FastAPI for backend services Docker for containerized deployment PostgreSQL, MySQL, SQLite (supported databases) Plotly / Matplotlib for visual output

Methodology

User inputs a natural-language query in a notebook cell. The input is sent to the LLM via LangChain, which performs intent detection and extracts relevant entities. LangChain builds a structured prompt that instructs the LLM to generate a syntactically correct SQL statement targeting the selected database. Generated SQL is validated using SQLAlchemy's safe execution layer to prevent injection attacks. The validated query runs against the target database; results are fetched into a Pandas DataFrame. Results are displayed inline, optionally transformed into charts using Plotly or Matplotlib. Feedback loop: users can refine the query or ask follow-up questions, which are contextualized by the LLM to maintain conversation state.

Future Scope

Support for NoSQL databases (MongoDB, Elasticsearch) through natural-language translation layers.

Multi-user collaboration within shared notebooks, enabling real-time query co-authoring.

Domain-specific fine-tuned LLMs for industries such as finance, healthcare, and retail to improve accuracy of generated queries.

Integration with BI tools (Power BI, Tableau) for seamless handoff of visual dashboards.

Advanced security features: role-based access control, query auditing, and encrypted communication channels.

Offline mode with locally hosted LLMs to reduce reliance on external APIs and improve data privacy.