Sutradhara

Generic LLM-Driven Data Agent Platform – Architecture & Requirements

1. Introduction

Sutradhara is a generic, domain-agnostic backend platform that enables secure natural language interaction with relational databases. It allows users to ask questions in plain language while enforcing identity-based access control, policy governance, and schema-aware data retrieval.

2. Vision & Objectives

The objective of Sutradhara is to act as an intelligent orchestrator between users, enterprise databases, and large language models. The platform ensures that only authorized, scoped, and sanitized data is exposed to AI systems.

3. Core Design Principles

Zero-trust LLM interaction, configuration-driven domain onboarding, strict separation of responsibilities, explainability, and auditability.

4. High-Level Architecture

Sutradhara consists of an API Gateway, Identity Resolver, Intent Resolution Agent, Semantic Mapping Engine, Policy Engine, Data Retrieval Engine, and LLM Summarization Agent.

5. Identity & Context Management

JWT tokens establish user identity, roles, organization, and linked entities. This context drives row-level and column-level security.

6. Intent Resolution

An LLM-based intent agent converts user questions into structured intents, entities, filters, and time constraints using strict JSON output.

7. Schema & Domain Configuration

Domains are onboarded using metadata definitions including database connection details, schemas, logical entities, and semantic mappings.

8. Policy & Access Control

Declarative policies define which roles can access which entities, columns, and rows. Policies are enforced before query execution.

9. Data Retrieval Engine

The engine dynamically builds parameterized SQL queries using validated intent and policy outputs while enforcing limits and masking.

10. LLM Summarization

Only filtered and sanitized datasets are provided to the LLM for summarization. No schema names or identifiers are exposed.

11. Multi-Domain Applicability

Sutradhara supports school, insurance, ecommerce, inventory, HR, and finance systems without code changes.

12. Open Source Tools

Relevant open-source tools include Open Data QnA, MindsDB Open Source, LangGraph, Open Policy Agent, Apache Superset metadata models, and community NLP-to-SQL projects.

13. Non-Functional Requirements

Security, scalability, performance, observability, compliance readiness, and extensibility.

14. Recommended Implementation Workflow

Register domain, configure schema metadata, define policies, deploy intent agent, connect LLM provider, and onboard users.

15. Conclusion

Sutradhara provides a safe, scalable, and explainable foundation for AI-driven access to enterprise data across domains.