CASE BRIDGE

Unified Legal Retrieval and Generation Across Jurisdictions

SCOPE OF THE PROJECT

The project aims to develop an LLM-powered legal research web tool that enables users to input legal research questions and receive the top 5 most relevant cases along with a concise analysis brief grounded in those cases.

UNDERSTANDING OF PAPER

- The paper introduces CLERC, a large, open dataset built from U.S. case law
- It shows that zero-shot retrievers and LLMs struggle on legal text, but training/fine-tuning significantly improves retrieval and reduces hallucinations in generated analyses.

PROBLEMS/ TYPICAL USERS/ BUSINESS CASE

Lawyers

- Legal professionals often face difficulty identifying the most relevant past cases from massive legal databases.
- This makes high-quality legal research time-consuming and challenging, potentially impacting case preparedness.

Law Students

- Students have limited access to real-world case materials and practical research tools.
- This creates a gap between academic coursework and the demands of modern legal practice.

General Public

- Individuals without legal expertise or access to specialized research databases struggle to find
- This limits their ability to address legal matters that directly affect their lives.

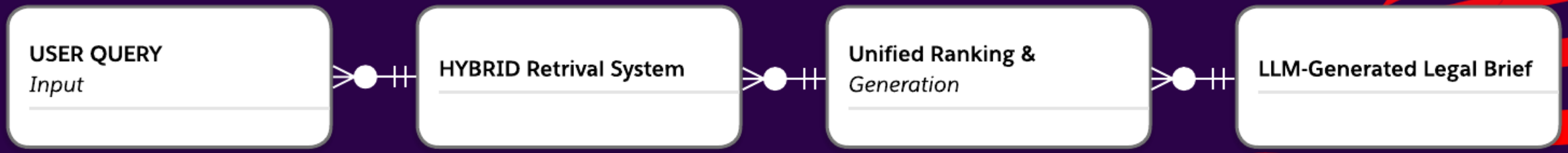
Businesses

- Corporate legal teams spend considerable time reviewing extensive legal records to assess risks or ensure compliance.
- This slows down critical business processes and increases operational

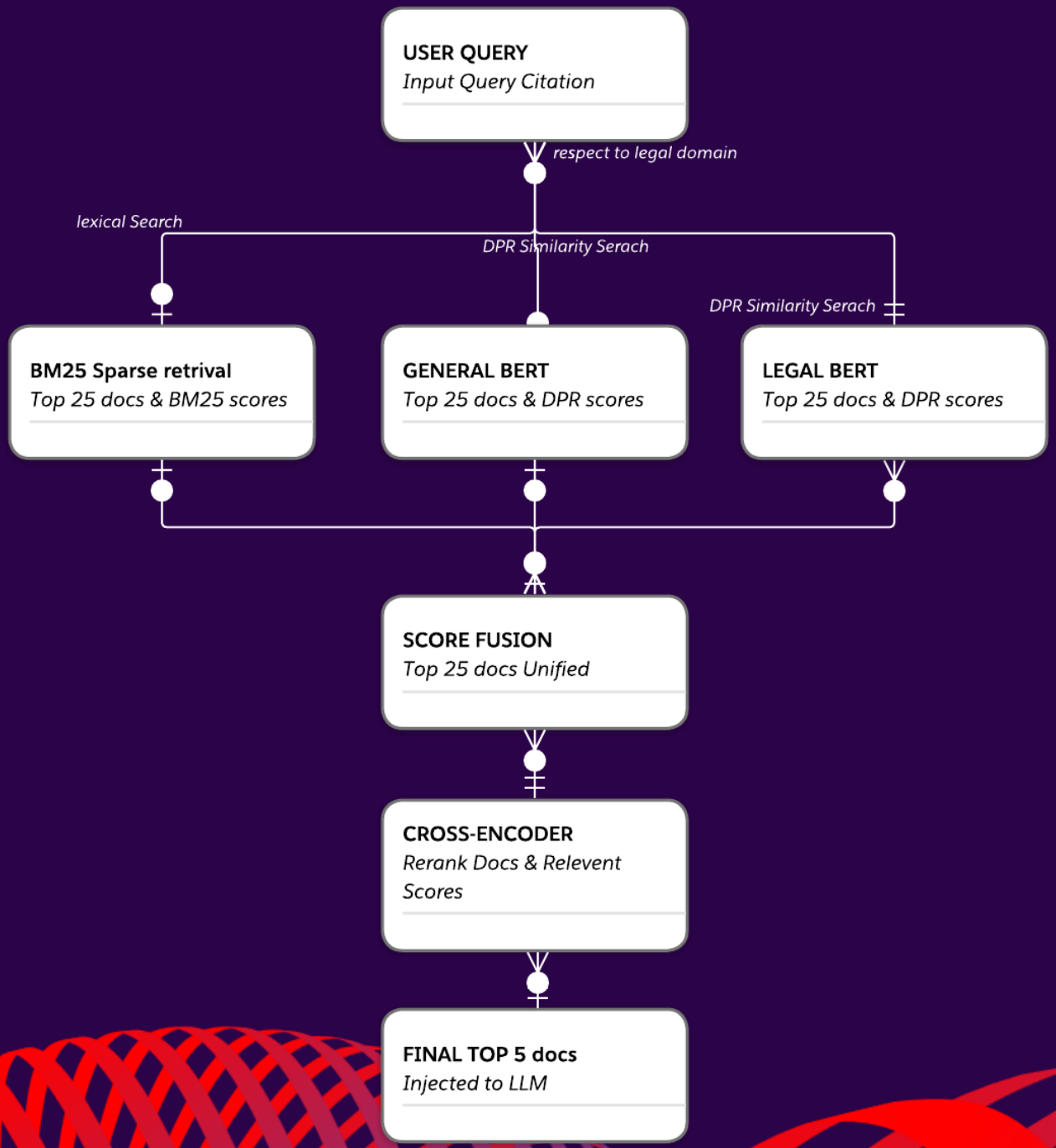
OUR SOLUTION
HYBRID RETRIEVAL & AI-POWERED LEGAL ANALYSIS

OUR SOLUTION EMPLOYS A HYBRID LEGAL CASE RETRIEVAL PIPELINE COMBINED WITH ADVANCED AI-BASED ANALYSIS

USER FLOW

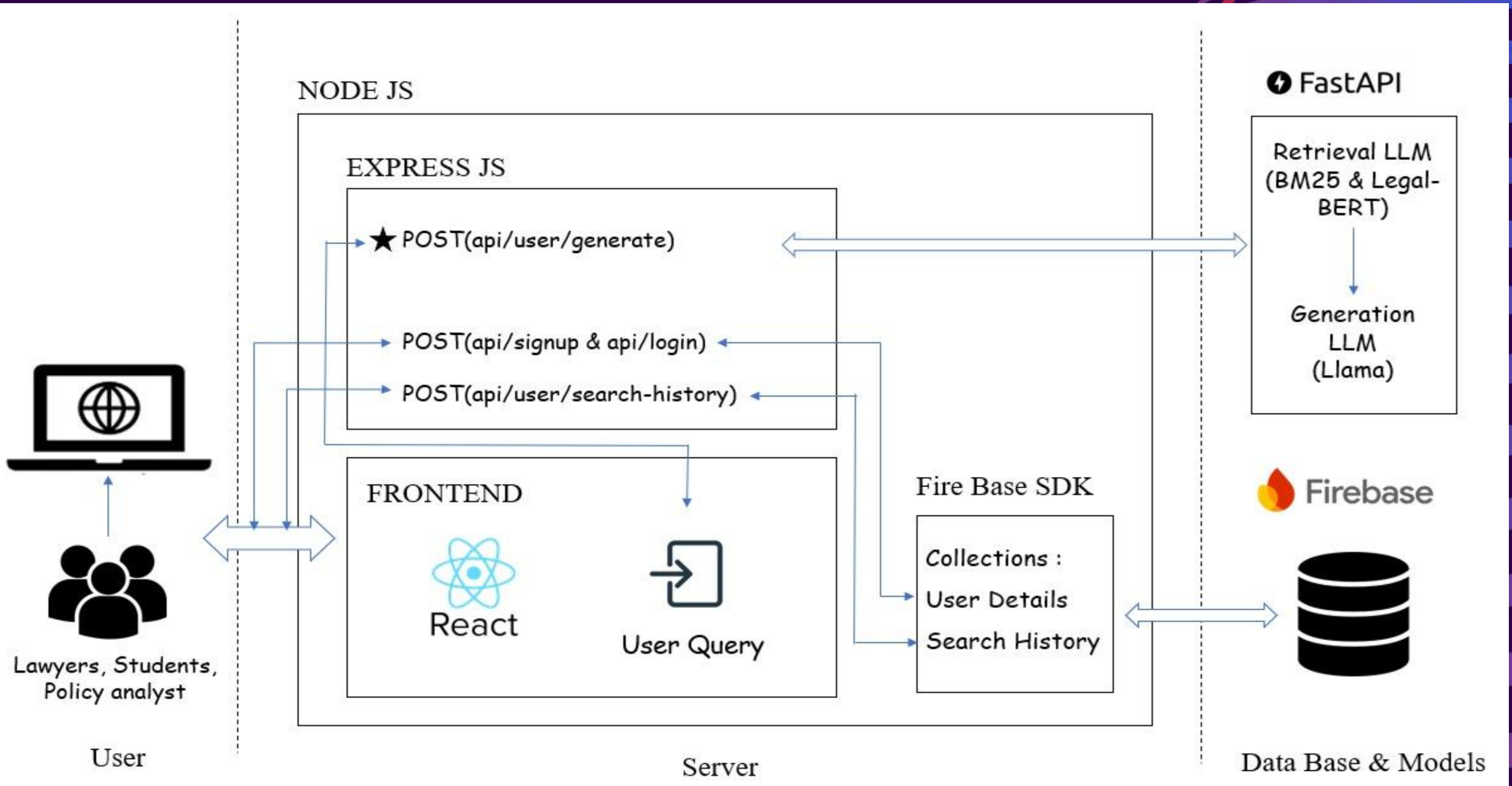


- WE INTEGRATE MULTIPLE COMPLEMENTARY SEARCH METHODS TO CAPTURE BOTH KEYWORD-BASED AND SEMANTIC UNDERSTANDING OF LEGAL QUERIES ENSURING COMPREHENSIVE CASE RETRIEVAL.



- Retrieved cases are then refined using a cross-encoder reranking model.
- The top five relevant cases are fed to a large language model (LLM) that generates concise, citation-grounded legal briefs strictly based on these cases
- Eliminating hallucinations and ensuring trustworthiness.
- The pipeline is designed for scalability and adaptability.

Architecture Diagram



RESEARCH PAPER IDEA & ACTION PLAN

- **IDEA:**
BUILD AN INDIAN CASE-LAW DATASET IN CLERC FORMAT (START WITH SUPREME + HIGH COURTS) AND EVALUATE A HYBRID SEARCH + AI BRIEF GENERATOR ON IT.

- **APPROACH:**
COLLECT JUDGMENTS AND **METADATA**
→ **CLEAN/STANDARDIZE**
→ **CHUNK INTO PASSAGES**
→ **CREATE SIMPLE QUERY–RELEVANCE PAIRS**
→ **RUN HYBRID RETRIEVAL + LLM TO PRODUCE CITED BRIEFS**
→ PACKAGE IN CLERC-LIKE FOLDERS.

- **TOOL:**
ADAPT THE OPEN CLERC PIPELINE AS THE TEMPLATE.

THANK YOU