

RAJ CHHAPARIYA

AI / Data Engineer · Data & LLM Systems

Bengaluru, India · rajchhapariya8@gmail.com · [Portfolio](#) · [GitHub](#) · [X](#) · [LinkedIn](#)

PROFESSIONAL SUMMARY

AI / Data Engineer with hands-on experience building and evaluating reliable data pipelines, in-process analytical engines, and retrieval-augmented LLM systems. Proficient in Python, SQL, and TypeScript, with practical experience designing deterministic query guardrails and assessing AI outputs for factual grounding and citation faithfulness. Supported by full-stack engineering capabilities across PostgreSQL, DuckDB, and Next.js.

TECHNICAL SKILLS

Data & AI Core:	Python, SQL, DuckDB, Pandas, NumPy, Scikit-learn, XGBoost
LLM Systems:	Retrieval-Augmented Generation (RAG), OpenAI API, BM25, Pydantic, Output Evaluation
Databases & Storage:	PostgreSQL, Supabase, MongoDB, Redis, REST APIs, Firebase
Supporting Web:	TypeScript, JavaScript, Next.js, React, Node.js, Streamlit
Tools & Analytics:	Git, GitHub, Plotly, Power BI, Matplotlib, Seaborn

AI & DATA SCIENCE PROJECTS

Agentic RAG Knowledge Assistant [Python, OpenAI API, BM25, NumPy, Pydantic] [GitHub](#) Aug 2026 – Present

- Built a retrieval-augmented generation pipeline that plans queries, retrieves information, checks evidence quality, and generates answers only when supported by evidence. Weak evidence triggers a retry or refusal, and every answer includes source citations.
- Combined semantic search with BM25 keyword search using **Reciprocal Rank Fusion (RRF)** to improve retrieval accuracy. Added handling for multi-part questions and contradiction detection.
- Built an evaluation framework measuring retrieval accuracy, citation quality, hallucination rate, and cost to systematically assess output quality.
- Tested extensively with a mocked API and budget limits: **87 automated test cases** pass without live API calls.

AI Data Analyst Agent [Python, LLMs, DuckDB, Streamlit] [GitHub](#) Aug 2026 – Present

- Built an agent that answers plain-English questions about datasets using four fixed tools (SQL queries, charts, statistics, and clarification), avoiding open-ended code execution for safer, more predictable behavior.
- Restricted the system to read-only SQL and built a deterministic verification check that validates the model's numerical claims against the underlying data to catch unsupported or incorrect assertions.
- Evaluated the system on **20 questions and edge cases**; it selected and executed the correct tool in every case, with limitations documented.
- Built a Streamlit interface supporting CSV upload, automatic data profiling, interactive charts, and report export.

PROFESSIONAL EXPERIENCE

Founder & Full-Stack Engineer — Truly (trulyapp.online) [Live](#) Apr 2026 – Present

- Architected an intentional social video platform for young urban Indians (ages 18–30) with structured 60-min circles, double-blind matching (2-hr SLA), Panda CSS, and Neon PostgreSQL with Drizzle ORM.
- Implemented pessimistic row-locking (`SELECT ... FOR UPDATE`) to eliminate overbooking across 8-seat circles, UUIDv7 keys, and Transactional Outbox event processing.

Founder & Developer — Resume Roaster (atsroast.com) [Live](#) Jul 2026 – Present

- Built and launched an AI-powered resume tool using Next.js, TypeScript, Supabase (PostgreSQL with Row Level Security), and OpenAI API, featuring a spatial ATS parser, zero-duplicate action verb rewriter, and ATS-safe PDF export.

Restaurant Owner & Operator — Varanasi, India Jan 2025 – Jul 2025

- Owned and operated an independent restaurant for six months, managing staff operations, budgeting, menu pricing, supplier negotiations, and vendor relationships.

OTHER PROJECTS

- Satta Darshan:** Civic legislative intelligence directory with Next.js, MongoDB, and Puppeteer scrapers. [GitHub](#)
 - SubsSync:** B2B SaaS subscription billing with Razorpay UPI AutoPay and webhooks.
 - Namma Metro Guide:** Android transit navigation app with offline routing & fare calculation (in testing).
- Aarohana Trails:** Freelance client trekking platform with custom CMS and Prisma ORM. [Live](#)
 - GSTify:** Invoicing software with Drizzle ORM, Neon PostgreSQL, and pdf-lib.

EDUCATION

Bachelor of Technology (B.Tech) in Information Technology
Dr. A.P.J. Abdul Kalam Technical University
2020 – 2024

CERTIFICATIONS

- Google Data Analytics Professional Certificate** (May 2023)
- Cloud Computing Certification**, Honeywell (Grade A, Oct 2023)