ZAID ALI SYED

AI Engineer / Software Engineer Intern | LLM Systems, RAG and Backend | Available full time from mid-December 2026

sidzaid72@gmail.com | +91-7972097242 | linkedin.com/in/zaidalisyed | github.com/zaidwhy
solstine.dev | ORCID orcid.org/0009-0003-4313-1510 | Aurangabad, India

Final-year B.Tech IT student who ships backend and LLM systems and measures whether they work. Founding engineer intern on a live ed-tech platform taking payments (27 merged pull requests, 24 audit findings fixed, 1 critical, a production incident root-caused over SSH); 2 research reports with DOIs; 3 PRs merged into open-source Python libraries (13.4k stars combined).

EDUCATION

B.Tech (Honors with Research), Information Technology - MGM University, Aurangabad Jul 2023 - Jun 2027 | CGPA 8.19/10

EXPERIENCE

Founding Software Engineer Intern - Zabira Academy (Remote) Aug 2026 - Present (ends Dec 2026)

- Shipped 27 merged pull requests to a live ed-tech platform taking payments (~1,400 files, 336 API endpoints): a full-codebase audit found 24 security and quality issues (1 critical), all fixed under review.
- Root-caused a production incident (broken password reset) over SSH to a missed migration, a disabled exec(), a database lockout and a MySQL 8 / PDO auth-plugin mismatch; fixed 3 more customer-reported bugs the same day.
- Built the project's first CI-gated end-to-end suite on a portable PHP + MariaDB + Playwright stack (113 assertions).

AI Intern - Pratyaya Limited Sep 2026 - Present (ends Dec 2026)

- Built a customer-support analytics dashboard from an empty repository (Python, FastAPI, pandas): daily to month-to-date performance, per-agent response, resolution and CSAT, bot versus agent resolution; 285 automated tests.
- Reconciled every figure to the platform's own reports: 10 of 10 aggregate totals match exactly, 53 of 55 per-agent response-time averages agree within one second; a figure that differs is clickable and explains why.
- Classified bot-to-agent handoff reasons from transcripts (read-only, test-enforced API client): unexplained "other" 10% to 1%.

ENGINEERING PROJECTS

CivilizationOS - multi-agent society simulation (live) *Python, FastAPI, React, Three.js, NetworkX, Ollama/Gemini/OpenRouter*

- 10 autonomous agents and 5 AI councils on a 3-tier model router (local Ollama, free tier, paid API) that runs locally at zero inference cost; 76 API tests, React + Three.js front end.
- Designed TCMF, causal-graph memory retrieval, benchmarked against 6 baselines over 300 scenarios: the first fusion scored recall@5 0.02 where its own signal reaches 1.00; normalized additive fusion recovered it (215 benchmark tests in CI). civilization-os-murex.vercel.app

Personal LLM - local-first memory and RAG kernel (live demo) *Python, FastAPI, SQLite, ChromaDB, sentence-transformers*

- Memory and RAG engine (SQLite + ChromaDB + hybrid model router) imported by 3 apps; citation-grounded answers that refuse when context is missing. 381 offline tests across the three, keyless CI. zaidwhys-personal-llm-demo.hf.space

RESEARCH AND COMPETITIONS - ARCHIVED REPORTS WITH DOIs (CC BY 4.0); ONE SHARED TASK

COLD READ - the anonymity half-life belongs to the reader *Python, Ollama, local LLMs*

- The reader, not the text, sets the anonymity half-life: 50, 800 or 1600 words across three same-size models; control held. doi.org/10.5281/zenodo.22309660

AUGUR - where time-locked language models actually stand *Python, local GGUF models*

- A model trained only on pre-1938 text says the year is 1899; a date anchor moves one model forty years, the other refuses. doi.org/10.5281/zenodo.22309658

SemEval-2027 Task 1 (RETECO) - registered solo entrant, retrieval sub-tracks *Python, BM25, bge-small fine-tuning, cross-encoder reranking, RRF*

- Fine-tuned bge-small + BM25 hybrid, mxbai cross-encoder rerank, RRF: dev nDCG@10 (official domain-macro) 0.115 to 0.297; official test Jan 2027. datascienceuibk.github.io/RETECO/

OPEN SOURCE - 3 PRS MERGED INTO LIBRARIES WITH 13.4K STARS COMBINED

- **Merged** jd/tenacity #668 - fixed static typing of @retry-decorated instance methods under mypy strict.
- **Merged** dry-python/returns #2480 - relaxed future() and future_safe() parameter types from Coroutine to Awaitable.
- **Merged** memgraph/gqlalchemy #390 - added an IS NOT NULL operator to the Cypher query builder, with tests and docs.

LEADERSHIP & ACHIEVEMENTS

- IEEE Student Branch, Vice Chair (2026), Treasurer (2025). GDG Tech4Sprint, Special Sponsor Award (Winner).

SKILLS

Languages & Web: Python, TypeScript, SQL, PHP; React, Next.js; deployed on Vercel, Render
AI / ML: LLMs, Retrieval-Augmented Generation (RAG), vector search (ChromaDB), knowledge graphs (NetworkX), multi-agent systems, tool calling, structured outputs, model routing, LLM evaluation, Ollama
Backend & Systems: FastAPI, REST, WebSockets, Pydantic, pandas, pytest, SQL, Docker, Git, GitHub Actions CI/CD, Playwright, Linux / SSH