

Siddharth Ahire

Backend & Applied AI Engineer

Bengaluru, India | +91 7276180147 | siddharth.direct@gmail.com
siddharthinks.com | LinkedIn | GitHub

Education

IIT Madras | B.Tech., Engineering Physics | 2024 | JEE Advanced AIR 1388

Professional Summary

Backend and Applied AI engineer with 2 years building production data systems, autonomous control planes, and multi-provider LLM infrastructure. Combines orchestration outside Databricks with deterministic validation, replay, human approval, and reliable backend boundaries.

Work Experience

Tredence Analytics | Bengaluru, India | Aug 2024 – Present

Data / ML Engineer

Python, Databricks SDK, REST, MCP, Spark SQL, Delta Lake, Photon, Azure AI Studio

- Built a continuously running autonomous control plane outside Databricks to migrate and create production healthcare ETL systems, connecting through the Databricks SDK, REST, and MCP rather than embedding agent logic inside notebooks.
- Designed planner, persistent-memory, scheduler, evaluator, self-improvement, and risk-tiered human-approval layers to orchestrate Jobs, Pipelines, Unity Catalog, Notebooks, Repos, Serving, Genie, and Apps within an air-gapped, low-anomaly-tolerance environment.
- Applied it after human analysis across 100+ production healthcare workflows, increasing delivery from 5 to 40+ per month, peaking at 55, while reducing the cross-functional team from 25+ people to 10.
- Held legacy-to-Databricks result variance below 1% through deterministic SQL parsing, grain, null, and cross-object gates; refactored a roughly 5,000-line monolith into 19 testable modules.

ugo, Inc. | Tokyo, Japan | May 2023 – Jun 2023

Machine Learning Intern – Computer Vision & Robotics

Python, PyTorch, OpenCV, Object Detection, Character Recognition, Jetson Nano

- Built and optimized object-detection and character-recognition stages for dual-arm humanoid inspection robots; separated perception from the downstream decision engine and delivered 7–8x faster recognition on Jetson Nano.

Selected Engineering Projects

Routr

[View Project](#)

Co-owned product / Multi-provider AI gateway

- Built the TypeScript/Hono application for 45+ providers and 250+ models; owned adapters, routing/fallback, circuit breakers, encrypted credentials, observability, reserve-settle accounting, and idempotent billing.

Intent Router

[View Project](#)

Open-source project / Deterministic LLM decision engine

- Built a TypeScript SDK for deterministic intent and model decisions, with capability, residency, and budget gates, exact fixed-point costs, replayable traces, and one-attempt execution.

Aargus

[View Project](#)

Independent product / Multi-tenant signal agent

- Shipped a Next.js/PostgreSQL product with scheduled collection, embedding and LLM scoring, encrypted BYOK credentials, graceful degradation, and human-approved action drafting.

Technical Skills

Languages: TypeScript, Python, SQL, C++ | **Backend/AI Systems:** Node.js, Next.js, Hono, REST APIs, OpenAI-compatible APIs, provider adapters, embeddings, tool calling, structured output, streaming, routing/fallback, circuit breakers, PyTorch, OpenCV | **Data/State:** PostgreSQL, ClickHouse, Redis, Databricks SDK, Spark, Delta Lake | **Reliability/Platform:** deterministic evaluation, human approval, idempotency, authentication, encrypted credentials, observability, Azure, Docker, GitHub Actions