

Suhaas Borra

AI Engineer

AI engineer building agent runtimes, hybrid RAG and fine-tuning systems. I build Python microservices for agent runtimes, hybrid RAG, multi-provider model routing, data ingestion and memory.

mehersuhaas123@gmail.com | Hyderabad, India | github.com/Suhaas-code

PROJECTS

genx github.com/Suhaas-code
Multitenant enterprise AI platform: 12 microservices, agent runtime, hybrid RAG and MCP connectors.

hdai github.com/Suhaas-code/PalmRecognitionPublic
Hand detection as input: turn any camera into a touchless gesture controller.

documint ai github.com/Suhaas-code
Legal document intelligence: six-stage OCR, parsing and knowledge-graph pipeline.

meta synthetic data kit github.com/meta-llama/synthetic-data-kit
Open source contributions to Meta's synthetic dataset generator for LLM fine-tuning.

EXPERIENCE

AI Engineer·Nainovate Technologies 2025 – present

AI Developer Intern·Brillius Technologies 2024 – 2025

ML Intern, Radar & Seismology·India Meteorological Department 2023 – 2024

TOOLKIT

ai / ml	LLMs, PEFT, LoRA, RLHF, HPO, RAG, hybrid search, reranking, agentic AI, MCP
backend	Python, FastAPI, gRPC, REST, SSE, Node.js
data / infra	MongoDB, Qdrant, Pinecone, LiteLLM, Docker, Kubernetes, GCP
other	React, computer vision, OCR, document AI, Git

~/.PLAN

~/.plan — what i'm doing

[now] genx agent runtime: tool loop, retries, streaming
[now] hybrid retrieval: dense + lexical, reranking eval
[next] write up the genx architecture as a public case study
[next] small experiments in speech + vision, on-device first
[done] portfolio v2: canvas ascii portrait, terminal, guestbook

last edited 2026-11

mehersuhaas123@gmail.com Hyderabad, India