Arjun Omray

+91 6396936059 · arjunomray@gmail.com · arjunomray.com

SUMMARY

Backend Engineer with 1.7+ years of experience driving core API design, database modeling, and server-side architecture in fast-paced startup environments. Led end-to-end development of critical backend services using Python, TypeScript, PostgreSQL, and Redis, taking features from concept to production deployment. Strong advocate for automated CI/CD pipelines, containerized workflows, and resilient distributed systems.

EXPERIENCE

Agniverse AI New Delhi

Software Development Engineer Jul 2025 – Aug 2026

- Sole backend engineer: built and owned all backend services across 4 consumer products with 50,000+ MAUs.
- Built a semantic search tool using vector embeddings and pgvector to power the video search library.
- Built CI/CD pipelines (GitHub Actions, Docker, AWS Fargate), automating up to 50 deployments/day at peak.
- Built an iOS device manager for fleet Appium/WDA automation with automatic failure recovery.
- Built Claude MCP servers integrating Meta Ads and payment webhooks for automated revenue analysis.

Backend Product Intern Nov 2024 – Jun 2025

- Built auth, business logic, and multimodal LLM APIs (OpenAI, Gemini) in Node.js/Express on MongoDB.
- Built Next.js landing pages with Meta Pixel and CAPI integration for Meta Ads conversion tracking.

ACHIEVEMENTS

- **Django PR #18413**: Added the `ColorInput` widget to Django core forms for native HTML5 color inputs.
- **Django PR #18352**: Resolved admin CSS regression, fixing paginator border alignment on changelist tables.
- **Django PR #18306**: Resolved admin UI regression, correcting container padding on the admin login form.

PROJECTS

What The Docs | *Python, FastAPI, PostgreSQL, Redis, Docker*

- Built an async RAG backend fusing pgvector HNSW and PostgreSQL GIN search via RRF in a single CTE.
- Built an arq queue with SHA-256 ingestion and Redis caching; benchmarked ∼3,800 QPS at 2.8ms p50.
- Built a token-aware chunker with tiktoken and PyMuPDF, preserving page citations for verifiable answers.

Rate Limiters | *TypeScript, Bun, Express*

- Implemented fixed-window, sliding-log, and token-bucket algorithms in TypeScript behind a unified interface.
- Built an Express server emitting `X-RateLimit-*` and `Retry-After` headers; benchmarked ∼800k ops/sec.
- Benchmarked 1M requests across 1,000 simulated users in Bun, profiling throughput and memory consumption.

TECHNICAL SKILLS

Languages: Python, TypeScript, JavaScript, SQL

Frameworks: FastAPI, Django, Express, Node.js, Next.js

Databases: PostgreSQL, MongoDB, Redis

Infrastructure: Docker, AWS EC2, AWS Lambda, AWS Fargate

EDUCATION

ABES Engineering College

Bachelor of Technology in Computer Science