

# Dipesh Singh

AI & Backend Systems Engineer

✉ [sindhideshow@gmail.com](mailto:sindhideshow@gmail.com) [dipesh.dev](https://dipesh.dev)

[in dipesh-singh253](https://www.linkedin.com/in/dipesh-singh253) [github dipeshsingh253](https://github.com/dipeshsingh253) [twitter dipeshdotdev](https://twitter.com/dipeshdotdev)

## Experience

---

### TradeBrains

Sep 2025 – Present

SDE-2

On-Site

- Led the backend of a live education platform as the sole backend engineer, owning features, reliability, and production issues end-to-end.
- Built a production AI(Langgraph) query system serving **165** daily queries and used by **2,500** unique users in the last 3 months.
- Reduced per-query AI cost by **70%** through prompt refinement, context summarization (every 10 messages), and routing simpler requests to smaller models, while improving response quality.
- Managed on-prem infrastructure and AWS services (EC2-based deployments), including routing, NAT, and service networking.
- Diagnosed and fixed a Docker networking issue affecting an internal office-wide tool; identified root cause within **1–2** hours and restored service without cloud migration.
- Designed and built a time-based algorithmic trading platform supporting paper and live trading using FastAPI, PostgreSQL, and Dramatiq.

### Zivy

Jan 2025 – Aug 2025

Backend Engineer

Hybrid

- Built AI-powered features including summarization, draft suggestions, and personalized writing-style adjustments.
- Managed AWS infrastructure (ECS, EC2, Route53, ALB) using Terraform and Terragrunt to maintain consistent and reproducible environments.
- Self-hosted Loki for centralized logging (**60GB+ logs/day**), reducing observability costs by over **50%**.
- Developed Slack integrations and automation workflows used in core product flows.

### Blursday

Oct 2024 – Dec 2024

Backend Engineer (Contract)

Remote

- Delivered AI-powered Figma and Canva integrations under tight deadlines, including a dynamic brand design PDF generator built in one week.
- Built and shipped a Canva app in **2** days.
- Migrated AI workloads to Azure-hosted OpenAI to utilize available credits efficiently and preserve long-term GCP limits.

### Logwintech

Apr 2023 – Sep 2024

Software Engineer

Hybrid

- Built and maintained 2–3 Spring Boot microservices using Java, Kafka, and REST APIs.
- Contributed to production issue resolution and service stability in distributed systems.

### Triplekiwi

Jul 2022 – Mar 2023

Full Stack Developer (Apprenticeship → Full-Time)

Remote

- Developed React frontend and Spring Boot backend services across internal and client-facing applications.
- Resolved production issues and supported reliability of microservices-based systems.

## Projects

---

### Zhiro | zhiro.in

- Built and deployed a full-stack read-it-later platform including backend, webapp, React Native mobile app, and browser extension.
- Implemented authentication (email/password, Google OAuth) and background jobs using Dramatiq + Redis for RSS sync and article fetching.
- Designed efficient database schema (256 articles) and caching for articles, summaries, and TTS features.
- Deployed on VPS using Docker, Nginx/Caddy, and GitHub Actions for CI/CD.

## Open Source Contributions

---

### **OWASP ZAP (zap-extensions)** | [zapproxy.org](https://zapproxy.org)

- Reduced unnecessary initializations in zap-extensions, improving performance and lowering resource usage.

### **Keycloak** | [keycloak.org](https://keycloak.org)

- Improved Admin UI by displaying the realm display name on the welcome page for better clarity.
- Resolved Windows dev-server compatibility issue, restoring local development support.

### **Quarkus QE** | [quarkus.io](https://quarkus.io)

- Fixed Vert.x and HttpClient resource leaks in test infrastructure, improving stability and reliability.

### **Springwolf** | [springwolf.dev](https://springwolf.dev)

- Refactored JSON example generation logic to improve code clarity and maintainability.

## Skills

---

**Backend:** Java, Python, FastAPI, Django, Spring Boot, REST APIs, Kafka, PostgreSQL, SQL

**AI Systems:** LLM integration, prompt optimization, RAG-based systems, Langchain, Langgraph

**Infrastructure:** Linux, Docker, Docker Compose, AWS (EC2, ECS, Route53, ALB, Bedrock), Terraform, Terragrunt, Nginx, Caddy

**Observability & CI/CD:** Loki, GitHub Actions

**Frontend (working knowledge):** React Native, Svelte, NextJs, React

## Education

---

**Bachelor of Computer Applications (BCA)**

*Indira Gandhi National Open University*

2023 – 2026

*India*