

Aditya Dey

Kolkata, India | +91 79085 31553 | adey9217x@gmail.com | linkedin.com/in/aditya-dey | github.com/AditHash | adithash.in

SUMMARY

Backend and AI/ML Engineer with 1.5 years of production experience building RAG systems, agentic AI pipelines, and event-driven microservices on AWS. Delivered systems processing **2M+ documents** and **400K+ insurance claims**. Leads GenAI development within the team, mentors junior engineers on LLM integrations, and holds 3 AWS certifications including Machine Learning Engineer Associate.

EXPERIENCE

Backend Developer - AI

Dec 2023 – Present

Workmates Core2Cloud Solutions Limited, Kolkata, India

- High-Volume Insurance Document Intelligence:** Architected an async FastAPI microservice ingesting and digitizing **2,025,300+ documents** and processing **422,800+ claims**. Integrated AWS Textract with open-source OCR tools (Surya, Paddle) for automated extraction into MongoDB. Delivered a RAG-based QA chatbot using Qdrant and AWS Bedrock for natural language querying over scanned records, deployed on AWS ECS/ECR, eliminating manual data-entry bottlenecks at scale.
- Scalable Policy Recommendation Engine:** Engineered a sub-100ms retrieval system across **11,000+ policy records** using PostgreSQL and pgvector. Implemented hybrid search combining dense vector embeddings with cross-encoder re-ranking to maximize recommendation relevance. Collaborated with stakeholders to secure the system with RBAC and API rate limiting inside a Dockerized ECS deployment.
- NBFC Loan Advisory Chatbot (WhatsApp):** Designed and shipped a customer-facing conversational AI over WhatsApp, serving real-time loan, EMI, and account balance queries. Built a RAG pipeline on AWS Bedrock for policy grounding, with structured database fallback and multi-turn session management, improving customer resolution time and reducing support load.
- AI-Driven Infrastructure Monitoring Agent:** Developed an autonomous incident-response system using LangGraph, AWS Lambda, and Boto3 that monitors CloudWatch alerts and triggers self-healing workflows, significantly cutting L1 investigation time. Instrumented full trace observability via LangSmith for proactive bottleneck detection.

PROJECTS

- Real-Time Fraud Detection Pipeline** – github.com/AditHash
Developed an event-driven pipeline using Apache Kafka for high-throughput transaction streaming, with a two-stage rule-based and scikit-learn ML inference architecture across 3 Kafka topics. Shipped a live SSE dashboard with full Docker Compose orchestration.
- Voice-RAG Hybrid Assistant** – github.com/AditHash
Built a low-latency voice assistant on AWS Bedrock Nova with chat-scoped RAG (PDF uploads and web search), real-time WebSocket streaming, barge-in interruption handling, and session-isolated ChromaDB vector stores.

SKILLS

Languages	Python, C++, SQL
AI and ML	RAG Systems, Agentic Workflows, Vector Embeddings, LangChain, LangGraph, LangSmith, Hugging Face, Scikit-learn
Databases	PostgreSQL, pgvector, MongoDB, Qdrant, ChromaDB, OpenSearch
Cloud	AWS Bedrock, AWS SageMaker, AWS Textract, AWS ECS, AWS Lambda, AWS S3, AWS CloudWatch, Docker
Backend	FastAPI, Flask, Pydantic, Apache Kafka, WebSockets, REST APIs, Microservices

EDUCATION

Techno India University, Kolkata

B.Tech in Computer Science and Engineering, 2024

GPA: 8.46 / 10

CERTIFICATIONS

AWS Certified Machine Learning Engineer – Associate	Dec 2024
AWS Certified AI Practitioner	Sep 2024
MongoDB SI Associate	Sep 2025
AWS Certified Cloud Practitioner	Feb 2024

AWARDS AND RECOGNITION

Rising Star – AI, Workmates Annual Business Summit	2026
AI Innovation Spark, Workmates Annual Business Summit	2026