Shivam Dixit

Senior Software Engineer & Forward Deployed Engineer | Applied AI, Agentic Systems & Distributed Backends
(+91) 8318559932 | shivamdixit7077@gmail.com | linkedin.com/in/shivam-dixit-fde | github.com/shivam-25

SUMMARY

Six years building production backend systems in regulated finance — American Express, Goldman Sachs, BlackRock and two funded fintechs. Designs event-driven, scalable distributed systems in Java, Rust and Python on Kafka, Spring Boot, PostgreSQL and AWS across payments, settlement and low-latency trading. Embeds with business teams to take agentic AI to production: maps stakeholders and scopes requirements using SPIN and RAPID, then ships multi-agent LangGraph systems with RAG, gated by evaluations that block regressions.

TECHNICAL SKILLS

Languages	Python, Java, Rust, Go, TypeScript, SQL
Applied AI	LangGraph multi-agent orchestration, tool-calling, RAG, hybrid retrieval, embeddings, MCP servers
LLM Evaluation	Recall@K, MRR, nDCG, groundedness & hallucination checks, LangSmith, Ragas, promptfoo
Backend & Data	Spring Boot, FastAPI, Axum, Kafka, NATS JetStream, ClickHouse, PostgreSQL, pgvector, Redis, Cassandra
Cloud & Platform	AWS, Bedrock, GCP, Docker, Kubernetes, ArgoCD, CI/CD, observability, incident response

PROFESSIONAL EXPERIENCE

American Express · *Software Development Engineer III* *India · Jan 2025 – Present*

- Embed with operations analysts to ship **BTA Resolve**, a six-agent LangGraph system auto-reconciling settled Business Travel Account card transactions against GDS/TMC itineraries, supplier invoices, T&E policy and HRIS data — **five source systems with no common join key**. **Cut reconciliation from roughly five days of analyst effort to 5–10 minutes in production.**
- Own the evaluation harness every prompt and graph change must pass: **Recall@K and MRR** over a labelled match set, **groundedness and hallucination checks** asserting each policy call cites its retrieved clause, and a golden regression suite in CI.
- Apply **RAG** with hybrid retrieval over supplier invoices and the T&E policy corpus, extracting structured fields from free-form documents to ground every policy-breach decision.
- Architected a concurrent DB2/PostgreSQL framework with independent transaction boundaries and config-driven rollback, de-risking a live production migration that held **99.9% data consistency**.

Filament Finance · *Founding Engineer* *Remote · May 2024 – Jan 2025*

- Launched a decentralised perpetual futures exchange as founding engineer, owning the backend end-to-end: a central limit order book and comb-pool liquidity model on Spring Boot, Kafka, Redis and MongoDB, key-partitioned for per-market ordering.
- Drove **order-matching latency down to ~50 ms**, with **50–100 ms end-to-end processing**; delivered REST/WebSocket order flow and smart-contract integration for on-chain collateral, margin and PnL settlement.

Goldman Sachs · *Software Engineer, Associate (promoted from Analyst)* *India · Apr 2022 – Apr 2024*

- Built and owned trade-flow and post-trade compliance surveillance controls across the front-to-back trading architecture at **5M trades/day**, investigating and clearing the exceptions raised.

ZestMoney · *Software Development Engineer I* *Gurugram, India · Aug 2021 – Apr 2022*

- Built event-driven microservices (Java, Spring Boot, AWS SQS/SNS, MySQL) across the consumer **lending** lifecycle — origination, settlement, payments and closure — at **50,000 loans/day**.

BlackRock · *Analyst* *Gurugram, India · Jul 2020 – Aug 2021*

- Built event-sourcing microservices (Java, Spring Boot, Kafka, Cassandra) for ETF order fulfilment on behalf of authorised participants at **2M orders/day**, with retry and recovery logic that guaranteed every order completed.

INDEPENDENT CONSULTING — PART-TIME, CONCURRENT

Click · Backend Data Engineer, Rust Trading & Real-Time Analytics (Aug 2025 – Oct 2026) · Rust backend behind a live trading venue: **10M+ events/day** ingested with exactly-once delivery into ClickHouse, powering real-time analytics over **490M+ rows**.

SurfCash · Backend Engineer, Stablecoin Neobank (Jan 2026 – Sep 2026) · TypeScript backend for a consumer **stablecoin** wallet: Circle wallets, MoonPay and Helio ramps, deBridge, bank rails, video-KYC/**AML** onboarding and BNPL origination.

PROJECTS — FDE ACADEMY

Ramco Payce — Support Triage Agent (Nov 2026) · Scoped a payroll provider's support queue with nine stakeholders — **15,000 tickets a month costing USD 90,450 to handle**. Shipped a PII-masking agent that classifies, de-duplicates and drafts cited answers, routing the rest to a person; hybrid retrieval on Bedrock, **measured run cost USD 87 a month**.

SetuHaul — Dock Exception Agent, Logistics (Nov 2026) · Late trucks trigger 20–35 driver messages an hour, and two drivers can be promised the same dock. Built a conversational exception desk over a deterministic capacity core — the model never decides feasibility or books — so **one booking wins from 20 simultaneous claims, 10 runs of 10**, the rest offered the next slot at once. Validation exposed a live rule gap that was rejecting every slot at one warehouse.

EDUCATION

FDE Academy · PGP in Forward Deployed Engineering & Applied AI Solutions, Cohort 3 *Jul 2026 – Feb 2027*

Vellore Institute of Technology · B.Tech, Computer Science & Engineering, CGPA 9.06/10 *Jul 2016 – Jun 2020*