

ASHWANI JHA

Software Engineer • Distributed Systems, Search Infrastructure & Applied AI
Dubai, UAE • +971 50 265 8196 • +91 96328 33382 • ashwanijha04@gmail.com
LinkedIn • GitHub • Friday (project) • Open to relocation

SUMMARY

Software engineer with 7+ years building large-scale distributed systems and applied-AI products at Amazon and Propertyfinder. Currently leading the search infrastructure powering the GCC's largest property portal — 2M+ listings across 5 countries on DynamoDB + OpenSearch with real-time Kinesis sync. Prior work at Amazon spanned Just Walk Out (real-time CV anomaly detection, RAG ops bots, AI root-cause agents), Amazon Pay, and AWS CloudFormation. Comfortable owning problems end-to-end: architecture, code, on-call, and the cross-team alignment that makes large rollouts actually ship.

CORE SKILLS

Languages	Java, Python, Go, Node.js / TypeScript, JavaScript, SQL
Distributed Systems	Microservices, event-driven architectures, CQRS, partitioning, fault tolerance, high-availability design
Data & Search	DynamoDB, OpenSearch / Elasticsearch, PySpark, AWS Glue, Kinesis, Kafka, SQS, S3
AWS	Lambda, ECS, EKS, API Gateway, CloudFormation, CDK, Bedrock, SageMaker, IAM, OpenTelemetry
Applied AI / ML	RAG systems, vector stores (Qdrant), LLM agent architectures, ARIMA forecasting, market basket analysis, CNNs
Frontend	React, Next.js
Tooling & Ops	Docker, Kubernetes, CI/CD (CodePipeline, Jenkins), Grafana, on-call & incident response

EXPERIENCE

Software Engineer • Propertyfinder

Sep 2025 – Present | Dubai, UAE • GCC's largest property portal — UAE, Bahrain, Qatar, Saudi Arabia, Egypt

- Lead the search team powering property search for millions of seekers across 5 GCC markets — owning the end-to-end stack from ingestion and indexing to relevance, freshness, and on-call.
- Led the platform migration of 2M+ listings onto a new DynamoDB + OpenSearch architecture using a dual-write strategy with AWS Kinesis for real-time sync between source and search index, exposing per-segment drift and lag metrics and running a continuous reconciliation pipeline that diffed source-of-truth against the new index every few hours, surfaced mismatches to Slack, and auto-triggered re-migration before customer impact — cutting media migration latency 6× (1 hr → 10 min) along the way via a producer-consumer Glue/PySpark redesign.
- Built and own the enterprise listings APIs consumed by internal product surfaces and external enterprise partners — designed for high read throughput and predictable latency, with versioning, auth, and rate-limiting baked in to support partner integrations without regressing the consumer experience.
- Led the Digital Successor program — a company-wide AI platform letting any engineer or PM spin up their own Slack-native bot to answer team questions, run code reviews, and auto-generate code and design docs from internal context. Turned scattered tribal knowledge into queryable infrastructure and gave non-technical teams a self-serve path to automation.
- Built the external listings ingestion pipeline (search aggregator + DLD import) that normalizes heterogeneous third-party data — amenities, attributes, pricing, media, metadata — into Propertyfinder's schema, with inline field-contamination checks and an SQS + LRU-cache redesign that unblocked a TPS ceiling and shipped 29k listings/week with zero integration regressions.
- Led the search team's operational excellence program — reduced production incidents from 200+ to near-zero within 2 months by overhauling observability (Grafana dashboards for ingestion throughput, error rates, segment-level latency, and dual-write drift), tightening alerting thresholds, and instituting on-call hygiene practices that caught regressions before they paged.

Software Development Engineer II • Amazon — Just Walk Out

Aug 2022 – Sep 2025 | Bangalore, India • Real-time CV/sensor systems for Amazon Go physical stores

- Led a 10-engineer pod within Just Walk Out spanning backend, ML, and data engineering — owned the team's roadmap, hiring, and growth: ran sprint planning and design reviews, set technical direction across the AI / anomaly-detection stack, mentored 4 engineers through promotion cycles, and conducted 20+ hiring interviews to grow the team.
- Built a real-time anomaly detection system fusing camera and sensor signals to identify and rank the most impactful in-store issues, reducing operational cost ~40% and freeing ~60% of engineering bandwidth previously spent on manual root-causing.
- Shipped an AI auto-root-cause agent that diagnoses and proposes fixes for Amazon-codebase bugs, cutting median root-cause time from ~3 days to ~1 hour for the on-call rotation.
- Designed an internal RAG bot on AWS Bedrock + SageMaker over SOPs and internal docs, reducing onboarding and operational task time ~30% across new hires and support engineers.
- Built an ARIMA sales-forecasting model for Amazon Go retail stores driving capacity, inventory, and bulk-procurement decisions, with forecasts at 95% confidence.
- Ran transaction-data analysis (exploratory + market-basket) that informed product placement and promotions — 18% lift in cross-sell revenue and 13% improvement in inventory turnover.

Software Development Engineer • Amazon Pay

Aug 2021 – Jul 2022 | Bangalore, India

- Reduced transaction failures 40% by profiling latency hotspots across Amazon Pay's microservices and tightening service-to-service interactions.
- Shipped Amazon Pay × Alexa voice payments by designing the backend workflows and secure API contracts on EKS, enabling voice-initiated purchases for end customers.
- Wrote production services in Spring Boot on AWS (Lambda, API Gateway, EKS, OpenSearch, S3, SQS/SNS) with CDK-based IaC and CodePipeline CI/CD.
- Built React / Next.js UIs for internal payment tooling and customer-facing flows.

Cloud Infrastructure Engineer • AWS — CloudFormation

Jul 2019 – Aug 2021 | Bangalore, India • Global SME for CloudFormation, ECS, EKS, CodeDeploy

- Cut EKS release-cycle time 70% by building an end-to-end CI/CD pipeline on CodeDeploy for a representative reference architecture, later adopted as an internal pattern.
- Authored training material adopted by 200+ engineers globally — labs, deep-dive tutorials, and POC walkthroughs on CloudFormation features — closing recurring gaps in support coverage.
- Acted as escalation point for high-severity enterprise CloudFormation cases, partnering with TAMs, Solutions Architects, and service teams to land fixes in upstream code.

SELECTED PROJECTS

Friday — Personal AI assistant (Iron Man-inspired cognitive architecture) (ongoing)

- Modular agent runtime with BDI-style reasoning, four-layer memory (episodic, semantic, procedural, identity), and Qdrant-backed vector storage; biologically-inspired modules (Hippocampus, Amygdala, Basal Ganglia, Prefrontal Cortex) decompose recall, salience, action selection, and planning into composable services.
- Multi-channel ingestion including Outlook via Microsoft Graph + IMAP for long-horizon context capture.

Agent-facing brand layer (early stage, 2026)

- Hosted MCP server / API letting brands expose catalog, pricing, policies, voice, and transaction rules to AI agents — infrastructure for an agent-first commerce surface.

EDUCATION

B.E., Information Science • Ramaiah Institute of Technology, Bangalore | CGPA 9.32

Entrepreneurship & Development core team. Side projects: visitor-detection robotics (Raspberry Pi + PS3 Eye), self-driving car prototype using CNNs on Arduino.

CERTIFICATIONS & RECOGNITION

- AWS Certified Solutions Architect — architecting secure, resilient, cost-effective systems on AWS.

- Multiple Amazon Quarterly Awards for Ownership, Bias for Action, Insist on the Highest Standards, and Hire & Develop the Best.
- AWS CloudFormation Global SME — internal training adopted by 200+ engineers worldwide.