

Ankit Raj

Solutions Architect — Cloud, AI & Enterprise Transformation

Stockholm, Sweden | rajankit749@gmail.com | +46 72 774 7756 | ankitraj.cloud | linkedin.com/in/raj-ankit | github.com/restlessankyyy

PROFESSIONAL SUMMARY

Solutions Architect with 10+ years of experience designing and implementing highly scalable, flexible, and resilient cloud architectures for large-scale enterprise customers across automotive, fashion, and retail industries. Proven track record as a trusted technical advisor to global organisations — partnering with C-level stakeholders and engineering teams to craft cloud migration strategies, modernise legacy workloads, and accelerate cloud adoption. Deep technical expertise across AWS, Azure, and Google Cloud with hands-on delivery in serverless, containers, AI/ML, data analytics, infrastructure as code, networking, and security. Demonstrated ability to connect technology with measurable business value — driving 35% cost optimisation, 70% latency reduction, and platform reliability at 99.9% SLA for 200+ product teams. AWS Certified (DevOps Engineer Professional, Developer Associate) with a passion for educating, training, and building cloud solutions for diverse enterprise customers.

CORE COMPETENCIES

Solutions Architecture	Cloud Architecture Design, Well-Architected Reviews, Migration Planning, Technical Advisory
AWS Services	EC2, Lambda, Fargate, ECS, Bedrock, SageMaker, S3, RDS, DynamoDB, CloudFormation, CDK, API Gateway, Step Functions, EventBridge, CloudWatch, IAM, VPC, Route 53, CloudFront
AI & Machine Learning	GenAI/LLM Strategy, AWS Bedrock, SageMaker, Vertex AI, OpenAI, Prompt Engineering, RAG, MLOps
Multi-Cloud	AWS, Microsoft Azure, Google Cloud Platform — Architecture, Migration & Governance
Infrastructure	Terraform, CloudFormation, Bicep, Kubernetes, Docker, Helm, ECS, GKE, AKS
Data & Analytics	Data Pipelines, BigQuery, Kafka, EventGrid, GraphQL, REST, gRPC, Real-time Streaming
Security & Networking	Zero Trust Architecture, IAM, RBAC, VPC Design, OAuth 2.0, Entra ID, Sentinel, Compliance
Languages	TypeScript, Python, JavaScript, Go, SQL
Leadership	Technical Advisory, Executive Presentations, Workshops, Mentoring, Strategic Planning, Stakeholder Management

PROFESSIONAL EXPERIENCE

Lead Cloud Architect — AI & Multi-Cloud

2024 — Present

Global Fashion Retail, Stockholm

- Serve as trusted technical advisor to executive stakeholders, translating business requirements into scalable cloud architectures across AWS, Azure, and Google Cloud supporting 200+ product teams globally
- Designed multi-agent AI workflows with MCP tools and RAG pipelines for automated incident response, security governance, and compliance monitoring across GitHub repositories at enterprise scale
- Architect and lead implementation of enterprise AI platforms on AWS Bedrock, Azure AI Foundry, and GCP Vertex AI — designing end-to-end solutions with multi-model orchestration, auto-scaling inference, and A/B model evaluation
- Drive cloud migration and modernisation strategy for legacy workloads, enabling digital transformation across fashion retail and supply chain operations
- Led multi-million SEK FinOps programme — designed cost optimisation architectures reducing cloud spend by 35% through right-sizing, reserved capacity, and automated governance
- Established cloud governance framework at enterprise scale — policies, security baselines, RBAC, and compliance controls for 1600+ subscriptions across multi-cloud estate
- Develop technical content including architecture blueprints, workshops, and best-practice documentation — enabling broader field teams and upskilling 8+ engineers in cloud architecture and system design
- Lead enterprise architecture governance using SAP LeanIX, Zachman, and TOGAF frameworks — aligning technology decisions with long-term business strategy
- Led RFP evaluation and cloud strategy for AWS vs Azure — authored technical comparison frameworks, total cost of ownership analyses, and vendor scoring matrices to drive executive decision-making on multi-cloud adoption

Lead Software Engineer — Platform Engineering

2023 — 2024

Global Fashion Retail, Stockholm

- Engineered multi-cloud platform backbone on Azure & Google Cloud — achieving 99.9% uptime SLA for 200+ product teams across global operations
- Designed and built Internal Developer Platform with self-service provisioning — EventGrid + Functions architecture reducing team onboarding from days to minutes
- Evaluated, benchmarked, and deployed AI/ML tools (OpenAI, Claude, Gemini) — shipped developer productivity solutions improving team velocity by 40%
- Conducted deep-dive architecture reviews and design exercises, creating world-class solutions and reusable patterns for cross-team adoption

Senior Software Engineer — Distributed Systems

2020 — 2023

Global Fashion Retail, Stockholm

- Designed Zero Trust microservices architecture — 15+ services with sub-100ms p99 latency, demonstrating deep infrastructure and networking expertise
- Built high-throughput GraphQL & REST APIs handling 50K+ RPM across distributed systems on Azure + GKE with automated scaling and rate limiting
- Led CI/CD standardisation using GitHub Actions across 200+ repositories — 95% pipeline success rate, establishing DevOps best practices at enterprise scale
- Resolved complex technical problems through root cause analysis across distributed systems, improving platform reliability and developer experience

Cloud Consultant — AML/KYC Modernisation

2020

Cognizant Technology Solutions - Ikano Bank, Stockholm

- Engaged as cloud consultant to architect and deliver proof-of-concept migration of Ikano Bank's AML & KYC APIs from Axway on-prem API gateway to AWS serverless
- Designed event-driven architecture using Lambda, API Gateway, Cognito, IAM, and DynamoDB — processing 500K+ daily financial transactions with Terraform IaC
- Implemented AWS Organisations for multi-account governance, CloudWatch for end-to-end observability, and enabled API governance and billing controls across environments
- Migrated identity and access management from Curity to AWS Cognito — consolidated OAuth 2.0/OIDC flows, reducing operational overhead and licensing costs
- Achieved 70% latency reduction and improved system resilience through serverless patterns, auto-scaling, and managed services replacing legacy infrastructure
- Presented technical proof-of-concept to executive stakeholders — cost-benefit analysis and live demos drove consensus for full AWS adoption across the bank
- Built CI/CD pipeline with GitHub Actions and Terraform replacing legacy manual deployments — deployment frequency increased 10x with automated quality gates
- Authored migration playbook and architecture decision records (ADRs) enabling the bank's internal teams to continue the modernisation programme independently

[Volvo Cars, Gothenburg](#)

- Served as technical consultant to Volvo Cars engineering teams — designing and implementing enterprise observability platform (ELK, Grafana, Prometheus) on AKS monitoring 100+ automotive microservices
- Managed CI/CD pipelines for OpenShift & Kubernetes clusters; reduced deployment cycle time from hours to 15 minutes across automotive software delivery
- Authored runbooks, architecture documentation, and incident response procedures — achieving 99.5% SLA compliance across production environments
- Partnered with cross-functional teams to identify and resolve infrastructure bottlenecks, translating automotive-specific requirements into optimised cloud solutions

Cloud DevOps Engineer

2019 - 2020

[Cognizant Technology Solutions](#)

- Designed and implemented AWS cloud architectures for 5+ enterprise clients - CI/CD pipelines with CodePipeline, Docker, ECS/Fargate across diverse industry verticals
- Developed React-based Cloud Management Dashboard providing real-time infrastructure visibility and cost analytics to enterprise customers

Cloud Developer

2016 - 2019

[Cognizant Technology Solutions](#)

- Won 1st place at Enterprise AWS Game Day 2018 (200+ participants) - demonstrating deep AWS technical expertise under production-pressure scenarios
- Automated resource tagging and policy governance across AWS and Azure environments using AWS Config Rules, Azure Policy, and custom Lambda functions - ensuring compliance and cost attribution across enterprise accounts
- Provided technical guidance and knowledge transfer to client teams, building cloud fluency across enterprise organisations

KEY SOLUTIONS & ARCHITECTURE PROJECTS

Enterprise AI Platform — Multi-Model Orchestration (Global Fashion Retail)

Designed distributed AI orchestration platform serving 200+ teams. Multi-model inference pipelines with auto-scaling, A/B evaluation, and real-time serving. AWS Bedrock, Vertex AI, OpenAI, Python, TypeScript.

AML/KYC Serverless Migration — AWS Architecture (Cognizant / Ikano Bank)

Architected full cloud migration from Axway on-prem API gateway to AWS serverless. Lambda, Cognito, API Gateway, DynamoDB, CloudWatch, Terraform. Curity to Cognito identity migration. 70% latency reduction, 10x deployment frequency.

Enterprise FinOps Platform — Cloud Cost Optimisation (Global Fashion Retail)

Multi-agentic AI system reducing multi-million SEK cloud spend. Automated anomaly detection, right-sizing recommendations, and governance dashboards across AWS, Azure, and GCP.

OpenClaw Tesla — AI Vehicle Agent (Open Source)

AI agent plugin for Tesla Fleet API with 21 voice-controlled actions. Demonstrates IoT, automotive API integration, OAuth 2.0, and intelligent agent orchestration. TypeScript, REST APIs.

Rubin — AI Coding Assistant (Open Source)

Privacy-first local AI coding assistant with 14-tool agent mode, MCP server integration, streaming chat. Demonstrates AI/ML product thinking and developer tooling expertise. TypeScript, VS Code API.

CERTIFICATIONS

- **AWS Certified DevOps Engineer — Professional**
- **AWS Certified Developer — Associate**
- **Microsoft Certified Professional (MCP) — Exam 533: Azure Infrastructure**
- **Microsoft Certified: Azure Fundamentals (AZ-900)**
- **OpenHack: AI-Powered Knowledge Mining · Containers · Security, Compliance & Identity · DevOps**
- Google India Udacity Scholar

EDUCATION

Bachelor of Technology — Computer Science & Engineering

2012 — 2016

[Galgotias University, Greater Noida, India](#)

Specialisation in Cloud Computing & Virtualisation. Projects: IoT Heart Rate Monitor, Big Data Analytics.

PUBLICATIONS & THOUGHT LEADERSHIP

[Why FinOps Is Failing the Enterprise — And What Should Replace It](#) — Enterprise FinOps implementation and cloud cost governance (Medium)

[Google Just Compressed AI to 3 Bits With Zero Accuracy Loss](#) — Deep dive into TurboQuant KV cache compression (Medium)

[I Gave an AI Agent Root Access to a Sandboxed Linux Box](#) — NVIDIA NemoClaw AI agent sandbox experiment (Medium)

[What is Kubernetes?](#) — Container orchestration deep-dive (FAUN / Medium)

[OpenAI & DALL-E Experimentation](#) — AI image generation and practical applications (Medium)

[Building Apps on Google Cloud Run](#) — Serverless deployment patterns (Medium)

[The Future of Quantum Computing](#) — Emerging technology landscape (Medium)

[Azure DevOps is the new VSTS](#) — Enterprise DevOps transformation (Medium)