

Summary

Senior Software Engineer and Tech Lead with **6+ years** at Oracle, Google, and Amazon. Senior Bar Raiser with **30+ hires**. Led **5 engineers** to build a cloud-native medical record output platform from **0 to 1**, capable of handling **10M** daily requests and bursts of **1M requests** through an asynchronous, distributed architecture with **7 microservices**. Domain expertise spans Healthcare, Search, and Finance, with specialization in **AI/ML**, including **generative AI, RAG, and Agentic AI**.

Seeking technical and leadership challenges at a fast-paced startup operating at scale.

Work Experience

Oracle

Apr 2024 - Present

Senior Member of Technical Staff (SDE III) | Promoted to Tech Lead (Oct 2025)

Bengaluru (Remote)

- Built the India Clinical Reporting team from scratch and now lead **8 engineers** across **3 products**: MRO Gen2, XR Gen1, and Dex Gen1.
- Delivered a cloud-native report generation platform producing HTML and PDF reports for **30+ clinical concepts** across **16+** healthcare organizations in production. Onboarded **25+ cross-functional teams** to MRO, reducing onboarding time from **2 weeks** to under **1 week** through **AI SKILLS**, documentation, and partial automation.
- Own MRO Gen2's technical architecture, strategy, delivery, and production readiness. Defined and delivered a **1-year** roadmap. Planned and unblocked work for **5-8 engineers**. Delivered **4** successful releases (**2** quarterly, followed by **2** bimonthly) and increased delivery throughput by **100%** over **10 months**, measured by completed Jira story points, through execution improvements and **AI** adoption.
- Built MRO v1 as a cloud-native POC on OCI to modernize legacy Gen1 Clinical Reporting, drove adoption across multiple concept teams, and re-architected it into MRO v2 with **7 Java/Micronaut microservices** and a React UI to handle **10x** the Gen1 load.
- Architected a **3-lane** asynchronous platform capable of handling **10M requests per day** and bursts of **1M requests**. Routed **~1%** of workloads exceeding **10 minutes** or **1 GB** of output to a dedicated lane, maintaining **p80 < 3 s**, **p90 ≤ 1 min**, and **p99 ≤ 10 min** SLAs with **3 ms/page** report generation. Migrated traffic from MRO v1 to v2 without downtime.
- Coached the team on on-call incident response and served as the escalation owner for critical production issues. Restored **~1,000-page** reports to the end-to-end latency SLA, reduced heap usage by **40%**, and cut report generation time by **25%**.
- Led **3 engineers** to deliver the Report Ingress Service for Gen1 Clinical Reporting, involving **6 APIs**, in **1 quarter + 1 month** against a **2-quarter** plan. Enabled the phased migration from WebSphere to EJS, reducing processing hangs, lowering infrastructure and third-party costs, and speeding up delivery of new features.
- Turned highly ambiguous technical and product scopes into architecture and execution plans for the SI DRZ region build and EHRC modernization program.
- Proposed **AI-powered** template builder and prompt-based report generation initiatives, built POCs, and aligned **4 product managers** to advance the ideas.
- Changed a partner team's operating model to increase delivery velocity by **30%** and coached engineers across teams on **AI** adoption.

Google

Sep 2021 - Feb 2024

Software Engineer II - Google Search

Bengaluru

- Developed Exam OSRPs and the interactive Exam Quiz for **100M+ educational users** on Google Search, serving **1B+ queries per day**. Launched the localized Periodic Table in Hindi, Tamil, and Telugu and designed a scalable Language API to extend Indic-language access across lower-tier cities.
- Improved OSRP triggering for **40%** of education queries through the diOSRP migration, shifting Search query understanding from an entity-driven to an intent-driven architecture and impacting **0.05%** of Search traffic.
- Improved early Bard **LLM** responses through training and evaluation. Developed Search Features in Google's monorepo (**2B+ lines**, **20M+ files**) using C++, Java, Python, Go, gRPC, Protocol Buffers, Blaze, Ant, Boq, Guice, and Spanner. Analyzed educational OSRP triggering by running evaluations on **50K+ queries**.

Amazon

Jul 2020 - Aug 2021

Software Development Engineer

Hyderabad

- Engineered duplicate-detection and anomaly-detection systems for **100K+ invoices daily**. Developed APIs, microservices, dashboards, infrastructure, and CI/CD using Java, Kotlin, TypeScript, Python, Elasticsearch, DynamoDB, Redshift, Coral, Smithy, Ruby on Rails, and AWS CDK. Handled **24-hour** on-call support.
- Implemented **3 anomaly-detection rules** using statistical analysis and **Machine Learning** (Isolation Forest), detecting over **\$100M** in invoice anomaly value within the **1st month**.
- Shortened anomaly reporting time from **24 hours** to **~5 minutes** by migrating from a daily cron job to an event-driven workflow for invoice creation and updates in the cross-team Gandalf-Veritas integration project.

HackerEarth

Jan 2020 - Jun 2020

Problem Curator Intern

Bengaluru

- Authored and validated algorithmic problems, reference solutions, test cases, and code stubs in C++, Java, Python, and JavaScript for hiring assessments and internal coding contests at Google, Facebook, Nokia, PayPal, Salesforce, and Infosys. Also contributed to HackWithInfy, which drew **167,000+ participants**. Raised Nokia's post-contest Net Promoter Score from **7** to **8** by improving problem quality, resolving issues faster, and supporting participants.
- Delivered a **2-day, 10+-hour** Advanced Data Structures and Algorithms workshop at MNIT Jaipur to **120+ attendees** on behalf of HackerEarth.

Skills

Languages	Python, Java, C++, TypeScript, Go
Backend	Distributed Systems, System Design, Microservices, APIs, Event-Driven Architecture, Asynchronous Processing, Kafka, gRPC, WebSockets
Data & Search	MySQL, PostgreSQL, Spanner, Redis, DynamoDB, Elasticsearch, Knowledge Graphs, Amazon S3, Azure Blob Storage, Semantic and Vector Search, Amazon Redshift
AI & ML	Generative AI, LLMs, RAG, AI Agents, LangChain, LangGraph, MCP, Scikit-Learn, Isolation Forest, CNN, XGBoost, TensorFlow
Cloud & Infra	AWS, OCI, Azure, Docker, Kubernetes, Terraform, AWS CDK, CI/CD, Jenkins, Grafana
Technical Leadership	Architecture, Technical Direction, System Design, Production Readiness, Performance Optimization, Code Reviews, Technical Documentation, Mentoring, Hiring, Cross-team Delivery, Roadmapping, 1:1s

Achievements

Professional & Leadership

- Contributed to **30+ hires**, conducting **150+ total interviews** across bar-raiser (**50%**), coding and screening (**40%**), and design (**10%**) rounds.
- Served as a mentor and judge at OraHacks, an Oracle-wide hackathon, evaluating **12 teams** and their AI projects and guiding prototypes toward production readiness.
- Mentored **3 engineers** to take ownership of individual components: **1 UI** and **2 backend services**.
- 1st student** from the university to receive a direct SDE full-time offer from Amazon during **3rd-year** on-campus internship hiring. Served as class representative in the **1st** and **2nd years**.

Certifications

- Cloud (OCI): Architect Associate (2026), Developer Professional (2025), Generative AI Professional (2025), and AI Foundations Associate (2024).
- CodeChef Certified Data Structures and Algorithms Programme (**CCDSAP, Advanced**).
- Earned **30+ technical certifications** over the past **10+ years** across various computer science domains, including Machine Learning, Deep Learning, and Data Science.

Coding, Exams & Academic Excellence

- Qualified for the **ACM-ICPC** Asia Regionals, finishing **85th** at the Gwalior-Pune Regional, and earned **World Rank 1** in CodeChef's October Challenge 2018 (Div 2).
- Placed **5th** in the national **onsite** coding contest "ALOHOMORA" at NIT Durgapur and **21st** in a national **onsite** contest at LNMIIT Jaipur.
- Ranked **2nd** among **175 teams** in LPU's University Coding Contest, solved **1,000+ competitive-programming problems** during college, and participated in **3 hackathons** focused on machine learning and web development. Scored in the **99.98th percentile** in eLitmus Quantitative Aptitude.
- Ranked **86th** among **60,946 candidates** in the 2014 Rajasthan Pre-Engineering Test (RPET) and qualified JEE Advanced with an All India Rank of **15,550** among **126,000 candidates**, from a total JEE candidate pool of **15 lakh+**.

Education

Lovely Professional University	Jul 2016 - May 2020
B.Tech. in Computer Science Engineering	Jalandhar, Punjab
Specialization: Machine Learning	CGPA: 9.06/10

Personal Projects

tech-lead.in	Live
------------------------------	----------------------

*Next.js, React.js, TypeScript, Microfrontends, Python, FastAPI, Gemini AI, Supabase, PostgreSQL, Azure, Blob Storage, Razorpay, Go, MongoDB, **RAG**, Twilio, hCaptcha, Aiven (Valkey, Kafka, OpenSearch)*

- Launched a production-ready portfolio website using a microfrontend architecture to present independently deployed projects together, with quality gates for LCP $\leq$ **2.5s**, CLS $\leq$ **0.1**, **85+** Lighthouse performance, and **90+** accessibility and SEO.
- Benchmarked the FastAPI backend on a single Uvicorn worker with **10 concurrent clients**, sustaining **450+ requests/s** with **36 ms** p99 latency and **0% errors**.
- Integrated Gemini AI for portfolio Q&A and prototyped **RAG**-based retrieval.
- Integrated Razorpay for payment processing and implemented a role-based login and access-control flow for private application areas.

Articles and notes across technology, leadership, and hiring. AI projects include:

- Intelligent Report Builder**: An Agentic AI workspace for creating report templates, reviewing quality, investigating failures, and predicting delivery time using LLMs, RAG, LangChain, LangGraph, and human-in-the-loop review.
- Code Atlas**: A local-first code knowledge graph that maps repository symbols and relationships to provide focused context for debugging, code review, impact analysis, and coding agents.
- Personal Assistant with RAG**: A context-aware assistant that uses grounded document retrieval, specialized AI agents, separated memory, and governed tool calling across personal, finance, and work modes.

Online Profiles

GitHub	LeetCode	Codeforces	CodeChef	SPOJ	Kaggle
	1811 (Knight, Top ~5%)	1813 (Expert)	2022 (5-Star)		