

EMPLOYMENT

Senior Software Engineer –Tech lead

Amazon

Sept 2014–Present

Consumer Robotics (Prompt Engineering, LLMs, Distributed Systems, Data Platform, IoT, AI)

[Technical Lead]

1. Technical lead, mentor and design bar-raiser in an org of 28 Software Engineers split between three locations (Seattle, Bay Area and Bengaluru) focusing on Agentic AI and Data platform for robots.
2. Most recently, led a team of 11 engineers and 3 scientists to deliver an AI platform that improves both conversational and action-planning abilities for amazon consumer robot devices.
- 3.

[Stateful Multi-agent AI platform for Consumer Robots]

4. Overall categorized into four different platform verticals. (a) Inference Compute (b) Agent & Inference Management (c) Agent inter-operability and (d) Observability and evaluation.
5. My role is to design E2E architecture that supports multiple agents distributed between device and cloud. The framework handles message passing between agents, interruptions, dependency management, parallelized retrieval workflows and re-use of embedding strategies which reduced overall UPL (User Perceived Latency) by ~40% as the system optimized prompt assembly by parallelly aggregating prompt data and caching static sections.
6. I single-handedly designed version control system to track both inter and intra inference and agent dependencies distributed between device and cloud increasing dev velocity by ~25% as team did not have to spend time manually rebasing, resolving conflicts, or root-causing regressions.

[LLMs, GenAI, RAG, Versioning, API design, Embedding Models, Vector Database, Prompt Engineering]

[AI-enabled Embodied Conversations]

8. Drove the development of embodied conversations for multi-modal robot devices. This involved (a) development of a markup language to generate serialized inputs that can be used to ultimately render different animations, sounds, gestures on the device. (b) ability to generate runtime synchronized choreographies based on context (mood, sentiment, task, environment, etc.). and (c) prompt engineering to enable LLMs infer the above markup text.

[LLMs, ECS, Kubernetes, NLP, Prompt Engineering]

[Data Platform & Communication Framework]

10. Identified and built consensus around the need & scope of a dedicated multi-tenant data platform for robot devices. Designed E2E platform including on-device middleware and cloud components handling >10K TPS. Data aggregated on device (stored in-memory in a graph-based data structure, periodically synced to cloud, stored in DynamoDB) is vended to clients through push (SNS, SQS), poll (query APIs) and streaming (Kinesis).
11. Proposed & delivered a dynamic rules engine that streamlined creation, verification and dynamic deployment of rules driving robot's character & communication with the user. This rule engine is built on top of the above data platform and allows developers to write custom queries to filter events, and enable common behaviors (status indicators, speech, animations, body motion) in response to these events.

[AWS AppConfig, GraphQL, AWS AppSync, DynamoDB, SQS, SNS, AWS Kinesis, Lambda, KMS]

[Routines]

13. Lead architect for Smart Automation Platform (aka Routines) for Astro. Worked with 10+ engineers across teams to design a distributed system that aggregates data from the robot and smart home sensors in cloud into meaningful signals, execute tasks in a pre-configured sequence, and allow customers to mix & match triggers, actions and conditions to create custom experiences.
14. Drove migration of high traffic services (handling all speech to Amazon Robot Devices) from EC2 to ECS (containerized infrastructure) resulting in ~15% cost reduction, reduced startup and deployment times, auto-scaling amongst other benefits.
15. *[ECS, EC2, Docker, Distributed Systems]*

Amazon B2B Growth (Backend, Distributed Systems, Big Data, Analytics)

1. Launched ABS (Amazon Business Seller) registration pipeline in EU region (includes France, Italy and Spain).
2. Developed E2E Data Pipeline to process raw order history and find Top Products with an option to filter by desired segment (Buyer Industry, Size and Region) refreshed daily.
3. Lead architect for Custom Seller Dashboard i.e., custom actionable analytic dashboards for different tiers of Business sellers. *(Spark, Dryad, Job Scheduler, AWS Data Pipelines, DDB, Lambda, SNS, Hive, Hadoop)*

Amazon Pay (Backend, High-QPS services)

1. Lead developer for the registration pipeline which was used to on-board new merchants onto the Amazon Pay platform in India, Japan and Europe. This streamlined the on-boarding process and significantly improved (~30%) the Rtl (Registration to Launch) ratio.
2. Steered the transition of all Merchant-related operations, services and components from Seattle based Payments team to my team. Ramped up myself and then 20+ engineers onto different merchant swim lane products and services.
- 3.

Software Engineer

Kuliza (now Finvolv)

Feb 2012–Aug 2014

1. Designed and developed the workflow using open-source tools (Pentaho Kettle) to automate the entire process of Data Migration for 5000 (Ford and Lincoln) websites from a SQL (Microsoft SQL Server) based in Java to a NO-SQL (MongoDB) environment based in Python.
2. Developed iOS applications for The Hindu (one of the top 10 newspaper apps on the Apple App Store in India) and Deccan Herald (another popular daily in India).
3. Developed a RESTful API suite to assist in location-based marketing, customer service and point of sale applications using BLE (i.e., low-energy Bluetooth) technology.

EDUCATION

Graduate

Jaypee Institute of Information Technology

- B.Tech. in Computer Science & Engineering, CGPA (10 point): 8.5; Coursework: Object Oriented Programming Data Structures; Computer Organization & Architecture; Algorithms; Artificial Intelligence; Compiler Design; Computer Networks; Databases.

SKILLS

- LLMs; OpenAI; Anthropic Claude; Prompt Tuning; RAG; API design; Vector Database; Embedding Models; VDB
- Java; C++; Python; Ruby
- MSSQL; MongoDB, DynamoDB; S3; Redshift; HDFS
- Distributed Systems; IoT; Android; AWS; Big Data; Analytics; Data Platforms; LLMs
- Ruby on Rails; Django; Spring; Eclipse IDE; XCode; Android Studio; IntelliJ-; JIRA; Agile; Slack