

Arjun Shekar Bharadwaj

3403 Router Cmn
Fremont CA 94538
📞 +1 (650) 379 5120
✉ arjun.s.waj@gmail.com
🌐 arjunswaj

Software engineer with 10+ years of experience in distributed systems, AI/ML platforms, and developer tools. Skilled in Java, Python, LLMs, and agentic workflows.

Experience

2024–Present

Senior Software Engineer, Microsoft Azure, Mountain View

Building scalable, fault-tolerant compute platforms for batch and streaming workloads. Working with Kubernetes, Apache Airflow, and open-source tools in the Azure Resource Builder team.

- Azure Compute Platform for Streaming and Batch Jobs
 - Built compute platform using Kubernetes and Apache Airflow. Integrated RisingWave, Spark, and Flink for unified pipeline orchestration
 - Designed data ingestion layer with Azure Event Hubs (Kafka) and Storage Accounts for high-throughput workloads
- Developer Experience & AI Tooling
 - Created container solution for dynamic NuGet package updates with zero-downtime deployments
 - Championed MCP servers for agentic workflows and custom prompts to speed up team development cycles
 - Deployed production services on AKS using Helm charts and multi-stage CI/CD pipelines

2015–2024

Software Development Engineer, Amazon, Sunnyvale/Bengaluru

Built large-scale distributed services in Java (~80%) and Python (~20%). Delivered low-latency APIs, payment systems, and multi-device platforms.

- AI-Powered Targeting & Action APIs – Alexa+
 - Built production AI agent system with MCP-like architecture. Enabled LLM-based device targeting for Music, News and Video domains
 - Designed RAG-style context enrichment with token optimization and Redis caching
 - Created prompt engineering pipelines with In-Context Learning. Achieved **90%+ accuracy** through test harness evaluation
- Cost & Latency Optimization – Music Recommendations
 - Rebuilt rule engine using composition and strategy patterns. Cut DynamoDB calls and **lowered costs by 50%** with better tail latency
 - Shipped Java APIs handling millions of daily requests. Tuned latency budgets with retry-with-jitter to meet SLOs at peak load
- Multi-Device Platforms & Leadership
 - Led Vega OS working group with 20+ members (Principal Engineers, Senior Engineers, Product Managers). Aligned API behaviors across device and cloud stacks
 - Ran Q4 readiness with load tests, chaos tests and capacity planning. Mentored engineers on design patterns and test coverage
- Invoice Generation – Global Seller Platform
 - Built async invoice workflow using S3 and secure SFTP to Deutsche Bank. Added connection pooling for reliable bulk transfers

Education

- 2013–2015 **Master of Technology (M.Tech)**, *IIIT-B*, India, *CGPA: 3.70/4.00*
Major: Information Technology
- 2007–2011 **Bachelor of Engineering (B.E)**, *RVCE*, India, *CGPA: 9.00/10.00*
Major: Computer Science and Engineering

Skills

- Languages Java, JavaScript, Python, C#, Haskell
- Cloud Platforms AWS (SQS, SNS, Lambda, DynamoDB, Athena), Azure (Event Hubs, Storage Accounts, Container Apps)
- Container & Orchestration Kubernetes, Docker, Azure Container Registry
- Frameworks & APIs Spring Framework, Spring Boot, Coral, AUI, Angular, jQuery, Proto + gRPC
- Data & Messaging AWS Data pipeline, Datamart, Kafka, Apache Airflow
- AI & ML LLM Integration, Prompt Engineering, Agentic Workflows, MCP Servers, RAG Systems, Langchain, Few-shot Learning