

Aryam Shinde

+1 (480) 791-6580 | aryamshinde2@gmail.com | aryamshinde.com

EDUCATION

Arizona State University

B.S. Computer Science | Cert: NVIDIA DLI, Generative AI & LLMs

Tempe, AZ

EXPERIENCE

Senior Software Engineer

August 2025 – Present

Holiday Channel LLC

- **Cut seller listing time by 70%** (10 min to 3 min per item) by shipping a vision powered auto listing pipeline where GPT 4 Vision extracts attributes from S3 stored images (category, condition, defects) and an LLM generates SEO titles and descriptions before the seller types a word
- **Scaled the pipeline for bursty seller traffic** with SQS async workers, multi image batching, and Redis cached embeddings for near duplicates, holding p95 latency under 8s while cutting Vision API spend by 30%
- **Eliminated 20+ hours per week of manual admin work** by automating compliance reminders (Resend + Node-cron, 7/3/1-day triggers) and syncing onboarding data into Google Workspace via OAuth 2.0 (Calendar, Sheets, Drive)
- **Shipped a multi tenant platform for admin, mentor, and student roles** on Express.js and Next.js with NextAuth.js JWT sessions and route-level RBAC enforcing tier specific access on every endpoint

Software Engineer

December 2022 – July 2025

DNA Traffic

- **Shipped an AI summarization Chrome extension to 1,200+ active users** (React, Manifest V3), with 3.5s median response on Hugging Face BART large CNN and exponential backoff retries holding availability above 99% during upstream API incidents
- **Cut API spend and latency by 35%** by caching the last 50 summaries per user via Chrome Storage and short circuiting duplicate requests at the content script layer; improved output quality with a Readability.js pipeline that strips ads and boilerplate before tokenization
- **Moved 8,000+ daily Salesforce records into Redshift** with a Python and AWS Lambda ETL using pandas for normalization, replacing a brittle nightly CSV export and giving leadership same day pipeline visibility
- **Replaced a weekly manual sales report** with Grafana dashboards on Redshift (DirectQuery, 15 min refresh) covering KPIs, regional performance, and funnel conversion, now the standing artifact in the 6-person leadership weekly

Software Engineer Intern

June 2022 – August 2022

Winwire

- **Reduced critical CVE exposure from 9 days to 3 days** by automating dependency scanning (Safety, pip audit) in GitHub Actions and blocking vulnerable PRs from merging
- **Quantified human layer risk across an 80 person org** with a phishing simulation tool (Flask, Celery, PostgreSQL) that surfaced a 23% click-through rate and drove the company's mandatory security training

PROJECTS

Glassbox | Python, FastAPI, LLM function-calling, agent loop, React

github.com/aryamshinde

- **Built a transparent AI agent that streams its reasoning live**, an agent loop (FastAPI, Groq function-calling) that decides between three tools (Wikipedia, calculator, URL fetcher), reads results, and iterates up to an 8-step ceiling before answering
- **Streamed every step to a React timeline over Server-Sent Events**, emitting typed events (Thought, ToolCall, ToolResult, FinalAnswer) rendered as the model's chain-of-thought, expandable tool-call cards, and a highlighted final answer
- **Designed a one-file Tool protocol and a safe calculator AST walker** (no eval), plus per-IP rate limiting with a bring-your-own-key fallback, and deployed it as a split Render/Vercel app

TECHNICAL SKILLS

- **Languages:** Python, JavaScript, TypeScript, Java, SQL
- **Frameworks:** React, Next.js, Node.js, Express.js, Flask, Angular
- **AI / ML:** GPT-4 Vision, Claude, Hugging Face Inference API, BART, prompt engineering, NVIDIA DLI GenAI
- **Data & Infra:** PostgreSQL, MySQL, MongoDB, Redis, Redshift, S3, SQS, Airflow, Spark
- **Cloud & DevOps:** AWS (Lambda, S3, Redshift, SQS), GCP, Docker, Kubernetes, GitHub Actions, Cloudflare
- **Tools:** Git, Postman, Grafana, PowerBI, Supabase, ServiceNow