

Experience

Amazon

Software Dev Engineer I

June 2026 – Present

Bellevue, WA

BlackRock

Software Engineer

Aug 2025 – May 2026

Atlanta, GA

- Architected a **local-first** config management MFE plugin using **React**, **TypeScript**, **module federation (single-spa)**, **IndexedDB** for offline persistence, **SQLite WASM** for ad-hoc querying, and **AJV/JSON Schema** for runtime validation.
- Reproduced core **Git** primitives from scratch: three-way conflict detection, staged change tracking, and PR diff rendering; supported ingestion from **Azure DevOps** REST API, local files, and ZIPs.
- Extended a **Java Spring Boot** private markets data exchange server: added bond type features, real estate attributes, and propagated client names through the exposure manager to the downstream calculation server.
- Implemented a resource-oriented **gRPC** endpoint (**Protobuf + Spring Boot**) to retrieve portfolios by template, collaborating cross-team on API and **JPA** entity design.

Amazon

Software Development Engineer Intern

May 2025 – July 2025

Seattle, WA

- Architected and deployed an **end-to-end ETL pipeline**, from **Andes** to **S3** to **AWS Glue** with **AWS CDK** and **Cradle**, processing over **300 billion** digital subscription records for downstream analytics.
- Engineered a **retrieval-augmented LLM agent** using Amazon Bedrock, combining structured data queries with modular prompts to expand metric coverage from 4 to 12 key business insights.
- Designed a SQL-driven evaluation framework for LLM outputs and integrated it with a Lambda-based regression harness, detecting schema drift and improving answer accuracy from **82%** to **95%**.
- Refactored and extended **CDK infrastructure**: removed redundant boilerplate, added two new stacks to auto-provision S3 buckets and Cradle job roles, accelerating onboarding of future agents.
- Contributed to **multi-agent routing** design and shadowed on-call, resolving customer-facing incidents related to subscription activations.

Projects

Go Semantic Signal Pipeline

- Architected a concurrent RSS ingestion service in **Go** (worker pools, rate limiting, backoff) with configurable concurrency; reduced end-to-end latency $\sim 5\times$ via parallel embedding and **vector search**.
- Implemented two-stage deduplication (exact ID + embedding-based cosine similarity using **OpenRouter + Pinecone**) with configurable thresholds and instrumentation of embedding/query/upsert latency.
- Exposed **Prometheus** metrics (throughput, duration, dup counts, errors) and structured **zerolog** run stats to support benchmarking and observability.
- Designed cost-aware **LLM routing** (task-specific primary/fallback models with timeouts) and optional LLM-based **knowledge graph extraction** persisted to **PostgreSQL**.
- Built cron-driven ingestion and daily summarization daemon with retention pruning, reset/benchmark flags, and multi-stage **Docker** deployment.

Skills

- **Languages:** Java, TypeScript, Python, Go, C++, SQL
- **Frameworks:** React, Spring Boot, Node.js
- **Cloud & Databases:** AWS (Bedrock, CDK, S3, Glue, Lambda, DynamoDB), PostgreSQL, Redis
- **Tools:** Docker, Git, gRPC, Prometheus

Education

Georgia Institute of Technology

B.S. Computer Science, GPA: 3.91/4.00

Aug 2022 – May 2025

Atlanta, GA