Colby Baker

Software Engineer | Backend APIs and LLM systems | Java, Python, TypeScript
Seattle, WA | 850-800-7888 | Colbyco017@gmail.com
colbybaker.dev | github.com/Colbyco | linkedin.com/in/bakercolby

Experience

CVS Health

Software Engineer, Health Agent AI Team *May 2026 – Present*

Patient-facing AI health assistant in CVS Health's consumer health app.

- Built a session-debugging platform (React, FastAPI, Splunk) that rebuilds any patient conversation turn by turn from **8 log sources**; adopted by **30–40% of the team** without a formal rollout, with cold loads cut from **55 s to 25 s**.
- Raised lab-question recall from **84.3% to 92.1%** on 554 real lab panels with a fuzzy matcher and disambiguation.
- Cut welcome-turn latency **~16.5 s**, measured end to end, by replacing three sequential LLM calls with a deterministic route.
- Designed and validated **7 feature-flagged domain specialists**, gated by a ~32k-line offline eval platform: two lab-answer defects fell to zero and blinded reviewers preferred the new answers **64%** of the time (p = 0.0083).
- Shipped a recent-health-summary feature end to end across a LangGraph agent and a TypeScript card service (**14 PRs**); designed field-level encryption and header redaction for PHI/PII in application logs.

Stack: Python, LangGraph, FastAPI, TypeScript, React, Splunk, OpenTelemetry, Kubernetes

Software Engineer, API Scheduling Team *Jun 2024 – May 2026*

APIs and data flows powering appointment scheduling across CVS Health's in-person and virtual care.

- Owned releases for **19 months** and built automation that opens promotion PRs from the latest CI image tags, shipping an **11-service UAT release wave in one day**; migrated **10 codebases** to GitHub Actions.
- Built structured logging and cross-service request tracing (correlation IDs, parent spans, OpenTelemetry) across **3 core services**; a 16.5k-event audit found and fixed lost correlation on **1.6%** of outbound calls.
- Cut the scheduling API's test suite by **58% (5m33s → 2m22s)** and one test group by **97% (185 s → 6 s)** by removing per-test Spring context rebuilds and test-state leaks.
- Built an LLM triage bot that routes ticket logs through a gatekeeper model to **8 domain-expert nodes** and posts root cause and owning team to Slack; mentored a new engineer on it.
- Shipped earliest-availability and distance ranking for employer-sponsored provider search; carried preferred-pharmacy selections across **6 services** through booking, confirmation and rescheduling (Cassandra, MySQL).
- Remediated dependency vulnerabilities across **17 scheduling repositories**, taking one API from **33 findings to 0**.

Stack: Java, Spring Boot, GraphQL, REST, Cassandra, MySQL, Redis, Kubernetes, Harness, GitHub Actions, Snyk, Splunk

Digital Engineering Intern *May – Aug 2023*

- Built a self-service metrics dashboard (React, TypeScript, MySQL) that replaced manual data requests to developers.

Projects

Aspicio · aspicio.app *Next.js, TypeScript, Three.js, Supabase*

Astrology platform with interactive birth charts, a 3D sky and relocation maps. Built the ephemeris engine on NASA JPL DE441; positions match JPL Horizons within 0.0005” across a 1,250-cell test grid. **1,700+ commits** and 600 test files since July 2026.

Spotify Classic Lyrics · github.com/Colbyco *Electron, Node.js, GitHub Actions*

One-click desktop app that restores Spotify's lyrics panel on Windows and macOS: undo, local diagnostics, auto-update and a GitHub Actions release pipeline (4 releases, Oct 2026).

Technical Skills

Languages and frameworks: Java, Python, TypeScript, JavaScript, SQL, Spring Boot, FastAPI, React, Next.js, Node.js, LangGraph, REST, GraphQL

AI systems: LLM applications, multi-agent routing, offline evaluation (graders, blinded review), guardrails, prompt design

Infrastructure and delivery: Kubernetes, Docker, GitHub Actions, Harness, Argo CD, Snyk, Vercel

Data and observability: PostgreSQL, MySQL, Cassandra, Redis, Splunk, OpenTelemetry

Education

Florida State University · B.S. Computer Science *2024*