Josh Rochon

Sr. Software Engineer

206.240.5182 | Seattle, WA
joshdrochon@gmail.com
linkedin.com/in/josh-rochon

SUMMARY

Senior software engineer with 7+ years building and running production web applications, APIs and cloud infrastructure for Starbucks, Microsoft, BECU and other clients. Modernized legacy .NET systems, served on call for critical releases and mentored junior engineers. Recently completed the Gauntlet AI fellowship, where my capstone was voted #1 by the cohort and staff. There I built a self-service developer platform on AWS with Terraform and least-privilege IAM, LLM agents gated by evals in CI/CD, and a fine-tuned 1.7B model that outscored Claude Sonnet on its target behavior. I use Claude Code, Cursor and Copilot daily.

SKILLS

AI & LLM: Anthropic Claude API, AWS Bedrock, LangGraph, LangSmith, RAG (BM25 + vector search, Cohere Rerank), LLM evaluation (golden datasets, pass/fail rubrics, LLM-as-judge), red-teaming (OWASP LLM Top 10), QLoRA fine-tuning, PyTorch, Hugging Face, Google Colab

Languages: TypeScript, JavaScript, Python, C#, SQL, GraphQL, Bash, PowerShell

Backend & Data: Node.js, Express, ASP.NET Core / .NET, Django, REST APIs, webhook delivery (retries, dead-letter queue, replay), PostgreSQL, MySQL, Microsoft SQL Server, FHIR

Cloud & DevOps: AWS (Lambda, S3, EC2, IAM, Elastic Beanstalk, Aurora, CloudFront, VPC, SSM, Secrets Manager), Terraform, Docker, Kubernetes (EKS), CI/CD (GitHub Actions, GitLab CI, Azure DevOps, Jenkins)

Frontend: React (Hooks, Context, Redux), Next.js, Apollo Client, Vue, Angular, HTML5, CSS3, Sass, Less, styled-components, Webpack

Testing & Quality: Jest, React Testing Library, Playwright, Moq, Jasmine, Selenium, load testing, WCAG 2.1 (axe, Lighthouse)

Tools: Claude Code, Cursor, GitHub Copilot, Git, Jira, Figma

PROFESSIONAL EXPERIENCE

Gauntlet AI | AI Engineering Fellow | Austin, TX | Jul 2026 – Sep 2026

Selective, full-time fellowship with a deployed project each week. Capstone voted #1 by the cohort and staff.

- Audited and improved a fork of the U.S. Treasury's open-source Ship app (TypeScript, React, Express, PostgreSQL). Cut p95 API latency from 1,276 ms to 103 ms at 50 concurrent connections, shrank the production bundle 82% and reduced queries on a key user flow from 50 to 37.
- Turned Ship into a public developer platform with an OAuth 2.0 server, HMAC-signed webhooks with retries, a dead-letter queue and replay, a rate-limited API with generated OpenAPI docs and a TypeScript SDK. Deployed it on AWS, with a CI check proving a new developer could go from a clean container to a verified webhook in 56 seconds, passing 20 of 20 runs.
- Ran Ship's AWS infrastructure as Terraform code (Elastic Beanstalk, Aurora PostgreSQL, S3, CloudFront, VPC, SSM) with encrypted remote state and a monthly cost alarm. Cut the server's IAM role from full admin to least privilege, tested against the live account, and rebuilt all 82 resources from code in a teardown drill that found three defects.
- Built a LangGraph agent for Ship that checked project data every 3 minutes for stalled work, review bottlenecks and sprint-miss risk, and asked a human before acting. Only ~1% of scans needed a model call in testing, keeping cost flat at scale. Deployed with Terraform and traced in LangSmith.
- Audited OpenEMR, an open-source health records system, for security and HIPAA gaps, then built a clinical co-pilot agent inside it with a verification gate that blocked any claim the agent couldn't trace to the patient's record.
- Built an ingestion pipeline for scanned lab PDFs and intake forms, run by a supervisor and two worker agents, with a strict schema and source citation on every extracted field. Gated merges on a 69-case eval set scored against 96 pass/fail rubrics, all passing, with a 0% error rate under load at 50 concurrent users.
- Built a four-agent red-team platform and ran 20+ attacks in seven OWASP and OWASP LLM Top 10 categories on the co-pilot, which held against every behavioral attack, and the platform filed five vulnerability reports on dependencies, a defense-in-depth gap and a model-config regression.

- Fine-tuned Qwen3-1.7B with QLoRA to beat Claude Sonnet on its target behavior, 86% to 71.5%.
- Built Silent Flow, a prototype that transcribes words mouthed silently at a webcam using a pretrained lip-reading model in PyTorch. A LoRA adapter trained on under 3 minutes of my own video cut word error rate from 23.3% to 18.9% on 69 held-out clips (p = 0.014).

Aim Consulting | Senior Consultant, Solutions | Aug 2021 – Apr 2026

Worked as a senior engineer on client teams at Starbucks, Microsoft, BECU and Purple. Mentored junior consultants and client engineers, and wrote runbooks for running and fixing live systems.

Starbucks | Senior React Engineer | Staffing Planning and Availability

- Replaced store managers' manual Excel staffing process with a React/TypeScript app, piloted in 400 stores and rolled out to 10,000+ with 100% manager adoption and 43,000+ sessions.
- Integrated GraphQL APIs through Apollo Client, used code splitting and lazy loading to shrink the bundle considerably, and added memoization to speed up rendering.
- Wrote React Testing Library tests and built dozens of accessible components from Figma designs.
- Helped deliver four sprints of scope in two before launch, with no added headcount.

Microsoft | .NET Developer | Threat and Vulnerability Reporting

- Migrated 22+ .NET projects off deprecated packages, replacing retired ADAL authentication with MSAL and Azure.Identity to clear vulnerabilities that breached the client's security SLA.
- Fixed data-processing bugs in a SQL Server–backed reporting system and introduced 100+ Moq unit tests to catch regressions.
- Took over 22 legacy projects unmaintained for years, rebuilt their undocumented .NET, Service Fabric and NuGet toolchains, and documented onboarding for other engineers.

BECU | Full-Stack ASP.NET MVC Developer | Online Banking

- Built RESTful C#/.NET APIs and full-stack features for online banking used by tens of thousands of members, working cross-functionally with product owners on requirements.
- Migrated legacy ASP.NET Web Forms pages to MVC so they were easier to secure and maintain.
- Built several reusable components that met strict design and WCAG accessibility standards.
- Served on call for the launch of a major new banking feature six months in development, doing root cause analysis on production failures.

Purple | Frontend Developer | Purple.com and Mobile Point of Sale

- Ran 5+ releases for major sales promotions, from cutting the release branch to deploying the Craft CMS site through its AWS pipeline.
- Co-led the Presidents' Day Sale site, used by thousands of shoppers, and built new pages with Vue.js, Craft CMS and Twig from Figma designs.

RedAwning | Software Developer | Property Manager Portal | Jan 2019 – Mar 2020

- Owned the bookings calendar, cutting click-through lag by about a second and fixing localization mismatches between client and backend and event-propagation bugs.
- Rebuilt the React portal mobile-first and migrated its CSS to styled-components.
- Added Redux to remove prop drilling and react-pdf for client-side PDF generation, keeping components cross-browser and WCAG compliant.

GIVE Volunteers | Junior Software Developer | Volunteer Portal | Jun 2018 – Dec 2018

- Built and deployed a full-stack volunteer portal (React/TypeScript front end, Python back end) with a team of four.

EDUCATION & TECHNICAL TRAINING

AI Engineering Fellowship | Gauntlet AI | Jul 2026 – Sep 2026
Full-stack Python/Django Certificate | Coding Dojo | Nov 2020 – Apr 2021
Full-stack React/C# .NET Web Developer Certificate | Epicodus | 2018
Apollo GraphQL Certification | Apollo | 2025
Bachelor of Arts in English | University of Washington | 2013

PUBLISHED CASE STUDIES · AIM CONSULTING

Modernizing Staff Planning with React for a National Beverage Retailer | 2025
Compliance and Technical Debt Issues Resolved for Client Application | 2024