

Joshua Fitz-Morris

Lead engineer of an AI document platform used daily by a Fortune 5 technology company.

Vancouver, BC · Canadian citizen (TN-eligible), open to relocating to San Francisco or remote
joshuafitzmorris09@gmail.com · 778-233-7067 · linkedin.com/in/jfitzmor · joshuafitzmorris.github.io

Experience

Lead Software Engineer, [Ollon](#)

2024 – Present

AI document intelligence and agent platform

Since Nov 2025

- Took the platform from inception to production in 28 weeks as technical lead and owner. Its seven GCP services have turned 70,000 vendor estimates and bids into human-reviewed records.
- Architected multi-tenant LLM extraction with Gemini tool calling and pgvector RAG over 26,000 documents.
- Shipped a deny-by-default MCP server for agents: 40 tools over 21 BI views, OAuth 2.1 with PKCE, per-user tenant scoping, and a CI test that fails any client-data tool without a scope check.
- Built the eval harness (16 hand-verified documents) that measures prompt and schema changes; raised quantity recall from 37% to 99% and reached 96.3% precision at 91% coverage on taxonomy mapping.
- Led four to five engineers at a time as final code reviewer: merged 117 PRs from eight teammates, wrote 92 of 134 docs and every CI/CD guard, and shipped 40 releases in the eight weeks after v1.0.
- Scoped and wrote the signed SOW for its next phase: anonymized retrieval over about 93,000 PDFs.

Context Studio, in-house document-intelligence product

Since Jul 2026

- Architect and owner: wrote the PRD and 9 of 15 ADRs, designed its model layer for Gemini, Claude and Bedrock, built its AWS platform, and caught inflated eval scores (F1 0.971, corrected to 0.937).

Stellantis, AI contract intelligence

- Cut legal and procurement search time by about 90% with MCP retrieval over 1,000+ vendor contracts.

Green & Spiegel LLP, immigration case management

- Led a five-person team on a \$1M+ engagement that launched a platform now used in several countries. Designed its schema-driven form engine, which saves hundreds of hours of manual legal review a month.

Full Stack Engineer, [Ollon](#)

Nov 2021 – 2024

- Rolled out Okta and Azure AD SSO for 1,000+ **Credit Suisse** users, shipped appointment booking for a 1,500-user clinic platform with SQL up to 9.5x faster, and cut mobile CI from 2 days to 2 hours.

AI Evaluator (contract, part-time), [G2i](#)

May 2026 – Present

- For a **frontier AI lab**, score coding-model output against calibrated rubrics and write expert training data.

Skills

AI: agents, tool calling, RAG, embeddings, MCP, LLM evaluation, Claude, Gemini, TypeSafe Jev, Codex

Engineering: Python, TypeScript, SQL, FastAPI, React, Next.js, PostgreSQL, pgvector, Docker, Terraform

AWS: ECS Fargate, Aurora, RDS, S3, SQS, Lambda, CloudFront, Route 53, IAM, Secrets Manager, Bedrock

GCP: Cloud Run, Cloud Tasks, Pub/Sub, Vertex AI, AlloyDB, BigQuery, Cloud Storage, Secret Manager

Education

Simon Fraser University, BSc, Computing Science

2021

AWS Certified Solutions Architect – Associate

2025