

ELMAR RASHO

Toronto, ON · rashoelmar@gmail.com · www.rasho.dev · linkedin.com/in/elmarrasho · github.com/TechElmar

SUMMARY

CS graduate from Western University (2026) with production experience building backend services, full stack web apps, mobile apps, and AI-integrated pipelines — all shipped and running independently. Built PokeDropz, a real-time e-commerce monitoring platform on AWS serving 9,000+ active Discord users with 12+ months of continuous production operation. Shipped Morph, a native iOS app with a multi-step OpenAI pipeline pending App Store approval. Published GPT Mini, a Chrome extension running LLMs locally via WebGPU and WebLLM. Proficient in Python, Java, JavaScript, TypeScript, Swift, C++, and C# — with hands-on experience in REST API design, async event-driven systems, cloud infrastructure, OOP & design patterns, unit testing, CI/CD, and AI-assisted development using Claude Code and Cursor.

EDUCATION

Western University — BSc Honours, Computer Science · Graduated June 2026 · London, Ontario

Coursework: Java OOP & Design Patterns, C/C++, C#, Data Structures & Algorithms, Databases I (SQL), Distributed Systems, Software Engineering, Operating Systems, AI I, Computer Organization, Statistics for Science

TECHNICAL SKILLS

Languages: Python, Java, JavaScript, TypeScript, Swift, C++, C#, Go

Mobile: Swift, SwiftUI, iOS SDK, XCode, XCTest, MVVM Architecture, Kotlin/Android concepts, React Native concepts, App Store submission process

Frontend & Web: React, HTML/CSS, TypeScript, JavaScript (ES6+), Chrome Extension API (MV3), Vue 3 concepts, Next.js concepts, Tailwind CSS, WebSockets, Socket.IO, WebGPU, WebLLM

Backend & APIs: Node.js, Spring Boot, REST API Design & Integration, FastAPI/Flask, Async & Event-Driven Architecture, Microservices, SQL (PostgreSQL, MySQL/MariaDB), Redis concepts, NoSQL concepts, Kafka concepts

AI & Agentic Systems: OpenAI API (Chat Completions, Vision), LLM Integration, Local LLM Inference (WebGPU/WebLLM), Prompt Engineering, Prompt Chaining, Structured Model Outputs, Agentic Workflow Design, MCP (Model Context Protocol), LangChain concepts, RAG concepts, Claude Code, Cursor, GitHub Copilot, ChatGPT

Cloud & DevOps: AWS (EC2, S3), GCP, Azure concepts, Docker, Kubernetes concepts, Git & GitHub, CI/CD Pipelines, Linux/Unix, Production Monitoring & Observability, Logging & Alerting, Agile / Scrum

Practices: OOP & Design Patterns, SOLID Principles, Clean Architecture, TDD concepts, Unit & Integration Testing, Code Reviews, Debugging & Root Cause Analysis, Incident Response, Technical Documentation

PROJECTS

PokeDropz — Inventory Monitoring & Alert Platform · Python · Java · AWS EC2 & S3 · REST APIs · Async Pipeline · Event-Driven · Discord API

- Python backend on AWS EC2 monitoring hundreds of e-commerce SKUs in real time via rotating proxy pool — parses stock status changes and pushes instant alerts to 9,000+ users via Discord webhooks; S3 stores persistent state across restarts.
- Async pipeline with per-request retry logic and exponential backoff; Discord bot layer handles command parsing, role-based notification preferences, and channel routing for 9,000+ users.
- 12+ months in production on Linux — triaged failures, refactored caching to eliminate false-positive alerts, maintained documentation, and shipped continuous improvements with zero manual intervention.
- Used Claude Code and Cursor throughout development; applied full engineering judgment over all AI-assisted output before shipping.

Morph: AI Physique Coach (iOS) · Swift · SwiftUI · XCode · OpenAI Vision API · Chat Completions · Prompt Chaining · Structured JSON · MVVM

- Native iOS app in Swift and SwiftUI that passes a physique photo through OpenAI's Vision API for image analysis, then runs a multi-step prompt chain across Chat Completions to generate a rating, critique, personalized meal plan, and workout program.
- JSON schema constraints enforced at every pipeline stage ensure clean data contracts between model calls and eliminate parsing failures; MVVM architecture decouples UI from the AI pipeline — each layer independently testable and swappable.
- Completed full App Store submission process including build, TestFlight, and App Review submission; pending Apple approval.

GPT Mini — Chrome Extension · JavaScript · TypeScript · Chrome MV3 · WebGPU · WebLLM · Local LLM Inference

- Published Chrome extension that runs LLM inference entirely in-browser via WebGPU and WebLLM — no API calls, no data leaving the device. Selected text on any webpage is processed locally by the on-device model and the response is rendered inline with zero page disruption.
- MV3 architecture: background service worker manages WebLLM model loading and inference orchestration; TypeScript content script handles DOM manipulation, response rendering, and contextual UI anchoring with clean separation of concerns throughout.