

LUCAS GOLD

AI Engineer

Toronto, ON | lucasegold01@gmail.com | lucas.gold | github.com/lucas-gold

SUMMARY

Software developer who builds production LLM systems and the products around them. Five years shipping to real users across mobile, web and backend; now building extraction and retrieval systems in daily use, plus a fully on-device AI assistant for macOS.

TECHNICAL SKILLS

AI & LLM: LLM APIs (OpenAI, Anthropic), RAG, hybrid retrieval, vector search, embeddings, LLM tool calling, structured outputs, Whisper, MLX

Backend: Node.js, FastAPI, REST API design, WebSockets, GCP Cloud Functions

Languages: TypeScript, JavaScript, Python, SQL, Swift

Frontend: React, Next.js, React Native, Expo, Zustand, HTML5/CSS3

EXPERIENCE

Software Developer, AI & Automation

March 2025 – Present

Above Ground Art Supplies — Toronto, ON

- Built an LLM + pdfplumber invoice extraction pipeline handling 130+ vendor layouts with no per-vendor templates — 94% of unseen invoices extracted fully correct (97% on known layouts), 3.5% of line items escalated for review.
- Cut invoice processing from over an hour each to near-instant, reclaiming ~40 hours a month for the purchasing and finance team that uses the tool daily.
- Built a hybrid-retrieval knowledge base over the live inventory catalogue — vector similarity plus keyword search, 92% recall@5 — used daily by 30+ staff.
- Replaced the company Mailchimp subscription with an in-house SendGrid campaign system: queued delivery, targeting from internal customer data, and a monitoring dashboard built on the metrics the business acts on.

Software Developer

August 2021 – February 2025

NextGen Kitchens — Vancouver, BC (Remote)

- Built the MVP of a white-label restaurant commerce platform and grew it to 20+ brands across iOS, Android, web, kiosk and POS from one React Native and React codebase on Node.js and GCP.
- Co-planned and drove a second-generation backend rebuild with one other engineer — a consolidated order pipeline with explicit fee and disbursement handling, migrated by parallel run with zero downtime.
- Shipped multi-provider payments across Stripe and Square — card-present on Stripe Terminal and Square hardware, Apple Pay, Google Pay — with automated multi-restaurant payout splitting through Stripe Connect.
- Gathered requirements directly from restaurant owners and shipped growth features — loyalty, referrals, promo codes, deep linking — as one of 2-3 engineers, mentoring 4 interns through code review.

SELECTED PROJECTS

Fennel — On-device AI voice companion for macOS

2026

Swift, SwiftUI, Python, MLX, SQLite, Whisper | fennel.garden | github.com/lucas-gold/fennel-ai

- Shipped a fully on-device AI assistant — SwiftUI app plus a Python/MLX backend in one self-contained macOS app, so no model, audio, or user data leaves the machine.
- Built the real-time voice pipeline: Silero VAD endpointing, Whisper recognition, streaming inference with prefix-cache reuse, Kokoro synthesis, barge-in interruption and echo cancellation.
- Gated model support behind an automated evaluation harness — a candidate must load, hold a conversation and render tool definitions correctly before it ships.

Harvest Rush — Real-time multiplayer arena game

2025-2026

React Native, Expo, TypeScript, Node.js, WebSockets, Fly.io | harvestrush.fennel.garden | github.com/lucas-gold/harvest-rush

- Designed a server-authoritative multiplayer architecture in Node.js: the server owns all game state at a fixed tick and clients send input intent only, making client-side cheating impossible.

EDUCATION

Bachelor of Science, Computer Science

June 2020

Toronto Metropolitan University (formerly Ryerson) — Toronto, ON | GPA 3.6/4.0 | Distinction, Dean's List 2019 & 2020