Richael

Auckland, New Zealand | huangd6666@gmail.com
dh8116.github.io | github.com/dh8116 | x.com/RicaV42

PROFILE

I'm 16, in Auckland, building Soulor (soulor.app): an AI app that lets a manager rehearse a hard conversation before they have it, and gives the business a one-page report that behaviour changed.

STAGE

- Live in six languages, billing switched on (Pro $7.99/mo, team plan $15/seat): 15 users, 0 paying. One outside user came back across 18 days; retention is the problem I'm working on now.
- Sales so far: cold email to training and development teams (2 company replies), a Product Hunt launch in October, and short-form video with tracked links.
- 2 users have run a real situation through it and written feedback.
- Target: 3 paying pilots by 20 November 2026.

THE BUYER

- Managers and team leads at mid-size companies, bought by training and development leads: a two-week pilot, with one metric agreed before it starts, reported per cohort at the end.
- Paid from the team's coaching and training budget. I'm still sizing that budget, so there is no market figure here yet.

HOW I WORK

- Diagnosed my own funnel: sign-ups were stalling at the confirmation email (8 of the first 14 accounts confirmed), so I put Google sign-in first, with no email to wait for, and shipped it the same day.
- A tester said the voice call "won't respond". I traced it to a call that never listened and only played canned lines, switched it off and rebuilt it as a real listening call the same evening.
- I measure before I claim: making each simulated person's relationship a stake cut stock replies ("no worries", "I get it") from 5 to 1 on messages held out from tuning.

TECHNICAL PROOF

- Nine Triton GPU kernels in nine weeks, each benchmarked against PyTorch; fused cross-entropy runs 1.51x faster with 1.67x less peak memory at vocab 131,072.
- Fine-tuned Qwen3-14B with LoRA for $1.30 of compute and built cold-start routing that took replies from 5.7s to 2.3s.
- Run the whole stack alone: my own multi-provider gateway (Groq, Gemini) that benches a provider at its per-model quota and alerts on outages, chat encrypted at rest, 200+ pull requests.
- Public benchmark (soulor.app/benchmark, model-judged): 8.0 against Character.AI's 7.9 and ChatGPT's 3.9 on sounding like a real person.
- Leading a five-person research paper on verifier error in reinforcement learning for language models (RLVR); draft due by the end of 2026, targeting COLM/TMLR.

EDUCATION & AWARDS

Year 11, Auckland, New Zealand (IB Math AA HL, NCEA Level 2). Expected graduation 2028

- AIME 10/15 | AMC 12 Honor Roll (Top 5%) | HiMCM and IMMC regional Honorable Mentions