

CoreFlow one-pager

# Gareth Roberts | Founding Engineer

## I make risky AI changes boring to ship.

AI can write implementation. Production still needs proof. My edge is building the verification, rollout, and rollback systems that let model promotions, route changes, backfills, and heavy-compute deploys ship without production roulette.

<b>25M</b>	<b>4.4x</b>	<b>&lt;24h</b>
<b>Scale</b>	<b>Risk</b>	<b>Compute</b>
Users served via multilingual multi-agent coaching and behavioural evaluation.	Throughput improvement in regulated insurance with validation, monitoring, and audit discipline.	Heavy-compute geospatial ML pipelines compressed from 6+ weeks to under one day.

<p><b>Why I fit this role</b></p> <p><b>Production AI at scale.</b> Built multi-agent systems and evaluation frameworks where model behaviour had to be measured, not guessed.</p> <p><b>High-consequence rollout.</b> Shipped AI inside regulated insurance where bad behaviour creates incidents, not just retros.</p> <p><b>Heavy compute comfort.</b> Led terabyte-scale ML systems and tight iteration loops under real delivery pressure.</p> <p><b>Founding temperament.</b> Recruited teams, set technical direction, and owned end-to-end delivery in 0-to-1 environments.</p>	<p><b>What I'd own first</b></p> <ol style="list-style-type: none"> <li><b>Evidence-gated model promotions.</b> Tie behavioural checks, telemetry, and rollback rules to route changes and model promotion.</li> <li><b>Make scary changes boring.</b> Codify patterns for failover, backfills, migrations, and canaries so safety isn't reinvented each time.</li> <li><b>Close the repo-to-reality gap.</b> Pull metrics, logs, incident history, and user signals into the verification loop.</li> <li><b>Leave behind reusable machinery.</b> A small tool or workflow that the team can keep using after the trial.</li> </ol>
---	---

Sydney | on-site | paid trial ready | gareth.roberts@ieee.org | 2.works