

Richard (Kai) Hallett

kai@oceanheart.ai · github.com/rickhallett · oceanheart.ai

Full-stack engineer with a focus on AI-augmented development. 5 years shipping production code in TypeScript, Python, and Go across retail analytics, social intelligence, and network security. 15 years as a cognitive behavioural therapist before that.

Currently building verification infrastructure for AI-assisted engineering: an evaluation platform, adversarial multi-model code review, and a field taxonomy of LLM output patterns. The CBT background turned out to be directly useful — pattern recognition for when systems produce confident output that is completely wrong.

Looking for a full-stack or platform engineering role where process quality and technical depth are valued.

WHAT I BUILT

Everything below is public. The git logs are the proof.

The Pit

A multi-agent evaluation platform — structured contests between agent configurations with observable traces, rubric-based scoring, failure tagging, and cost visibility. The product and the development process are the same thesis: governance infrastructure first, features second. Every architectural decision is on file, every commit passes through adversarial review, and the evaluation engine grew directly from running the process on itself. Next.js, TypeScript, Drizzle, Neon Postgres, Vercel AI SDK, Stripe.

Sortie

Adversarial multi-model code review, extracted from The Pit into a standalone tool. Three LLMs review a diff in parallel, a fourth synthesises their findings, and merge is gated on convergent severity. Built after a pilot run showed that 74% of findings were caught by only one model. Surveyed 40+ existing tools and 30+ papers before writing a line of code. Python.

Pidgeon · Pidgeon Swarm

A carrier-agnostic shipping rate integration, built twice from the same spec to compare development processes. The first used outside-in TDD with cross-model adversarial review after every step. The second used a design-first parallel agent swarm. The interesting finding wasn't which produced better code — it was that review timing determines what you catch. [Case study](#).

Tells

A field taxonomy of LLM output patterns — epistemic theatre, paper guardrails, analytical lullabies, sycophantic drift. Catalogued from live observation over several months of AI-augmented development. Includes programmatic detectors that annotate text in real time.

EXPERIENCE

2024–present

Software Engineer

Oceanheart.ai

Building The Pit — a multi-agent evaluation platform. Also built the adversarial review pipeline (Sortie), anti-pattern taxonomy (Tells), and cross-model triangulation system that governs the development process itself. All repos public.

2023–2024

Software Engineer

EDITED · Retail analytics

React, TypeScript, Python. Data visualisation features for AI-driven retail analytics SaaS. Partnered with backend engineers (Python/Django) to optimise API integrations serving enterprise clients.

2021–2023

Software Engineer

Brandwatch · Social intelligence

Enterprise data visualisation platform (Monitor project). Built scalable React components while modernising legacy Backbone codebase. Mentored junior developers.

2020-2021

Full Stack Engineer

Telesoft Technologies · Network security

TypeScript, Angular, Node.js. Secure features for cybersecurity applications. Python utilities for data processing and automation.

2019-2020

Software Developer

School Business Services

Vue.js frontend features for school financial management software.

2004-2019

Cognitive Behavioural Therapist

NHS / Private Practice

15 years clinical practice. PGDip CBT (Royal Holloway). The core skill — noticing when someone, including yourself, is producing plausible-sounding nonsense — turned out to transfer directly to AI-augmented engineering.

EDUCATION

PGDip Cognitive Behavioural Therapy

Royal Holloway, University of London

BSc Psychology

UWE Bristol

TECHNICAL

Languages `TypeScript` `Python` `Go` `Bash`

Frontend `React` `Next.js` `Tailwind` `Hugo`

Backend & infra `Node.js` `Bun` `PostgreSQL` `Docker` `Make` `Zod`

AI-augmented dev `Multi-model adversarial review` `Cross-family verification` `Claude / Codex / Gemini`
`Convergence analysis`