

Richard Woollcott

AI Systems Engineer

Bridgwater, Somerset, UK · rich@appmilla.com · linkedin.com/in/richard-woollcott

github.com/guardkit · huggingface.co/RichWoollcott · youtube.com/@RichWoollcott · guardkit.ai · lookitsrich.com

PROFILE

I'm a software engineer and architect building agentic AI systems, with a particular focus on agent reliability, evaluation and local model serving. I'm looking for hands-on AI systems work where I can take responsibility for how applications and agents are built, tested and operated.

I built a software factory with multiple cooperating agents that plan and implement software, while separate agents verify the results by building and running the features in a Docker sandbox. The factory runs locally on my Nvidia DGX Spark using local models and fine-tuned LoRA adapters, enabling six role-specific variants of the Gemma 4 base model to run alongside the Qwen 'workhorse'.

I also built Study Tutor, an AI study platform that uses a fine-tuned model to guide students through questions. Students can hold spoken conversations through a mobile app with push-to-talk or a Reachy Mini robot, with the conversation synchronised across both interfaces.

Alongside this, I architected and built an AI fintech proof of concept for FinProxy and led the Kiro pilot and rollout for AI-assisted development across DHL eCommerce and Evri. This builds on more than twenty years of delivering software, including the SSE customer app used by up to 500,000 people a month.

KEY SKILLS

- Agent harness development (github.com/guardkit)
- Fine-tuning SLM's on Nvidia DGX Spark using Unsloth, measured against frozen eval exams
- Multi-agent systems built with the Claude Agents SDK, LangChain DeepAgents SDK and LangGraph, co-operating over NATS JetStream
- AI assisted development adoption and governance using Amazon Kiro and Claude Code
- Server-side RESTful API development using C# and Kotlin with ASP.NET on Azure, Ktor, & Spring Boot
- .Net MAUI & Xamarin mobile app development for Android & iOS
- React Native and React SPA development
- Kotlin Multiplatform (KMP) development for Android (Kotlin) and iOS (Swift)
- Native Android development using Kotlin
- Software Engineering & Design – Kotlin, C#, C++ & exposure to Swift
- Mobile, Backend, Desktop & Embedded app architecture, design & development
- Team Leadership · Patterns & Best Practices · Stakeholder Engagement · Requirements Capture
- Xamarin Certified Mobile Developer

WORK EXPERIENCE

Technical Co-founder — Appmilla – AI R&D (guardkit.ai) | September 2025 – Present

Creating an AI augmented development eco-system guardkit.ai using Agent Harnesses and fine-tuned models (Google Gemma 4 MoE) using datasets created on an Nvidia DGX Spark via an agentic harness. These are built with Claude Agents SDK and LangChain DeepAgents SDK, with a Postgres-based fleet memory service replacing an earlier Graphiti temporal knowledge graph. These run on a NATS Jetstream message bus with agents co-operating to comprise a software dark factory – an orchestrated suite of agents including fine-tuned architect, coach and QA verifier models, with a product owner model pending.

- **AutoBuild** — Autonomous code generation pipeline using an adversarial Player-Coach pattern: a Player agent implements features while a separate Coach agent validates against quality gates, iterating up to five turns per task. BDD test validation proved brittle, so the Coach now validates by running the built feature, with an agent-driven fix loop. A recent run built 43 tasks across 6 features with only 3 high-impact human decisions.

- **Forge** — Pipeline orchestrator agent that subscribes to NATS JetStream events, manages Git branch operations, and runs the guardkit pipeline end to end: feature-spec → feature-plan → AutoBuild → task review.
- **Agentic Dataset Factory** — Domain-agnostic pipeline that generates and quality-checks training data for fine-tuning, built on a Player-Coach adversarial loop (LangChain DeepAgents SDK): one agent writes each training example, a second scores it against the domain's rubric and sends it back for revision. Adding a new domain is a config change, not code. Six ways of sourcing examples — from ingested PDFs via RAG (Docling + ChromaDB), generated from scratch where no source material exists, from a bank of pre-written task briefs, mined from real production agent runs, created by planting a known bug in working code so the correct answer is known in advance, and as near-identical pairs differing in a single detail to teach fine distinctions. Every example is validated by something outside the model — a compiler, the live service's own parsers, quote-checking and train/test leakage gates — before it enters the corpus. Feeds Unsloth QLoRA fine-tunes (Gemma 4 MoE, Qwen3) on a DGX Spark, graded by fleet-evals against frozen exams. Six domains shipped.
- **Study Tutor** — GCSE AI tutor using a fine-tuned Gemma 4 26B MoE model with Socratic questioning aligned to AQA exam specifications, served via llama-swap on local hardware. Students create and resume study sessions in the Flutter mobile app with tap-to-talk, or talk to a Reachy Mini robot, with the streaming sessions working together so a conversation with the robot is shown live in the app. Per-subject RAG collections are architected for seven further subjects.
- **Jarvis** — Personal AI assistant and intent router serving as the primary interface to the Ship's Computer fleet, built on the DeepAgents SDK with three launch skills surfaced via a Slack adapter, with trace-richness instrumentation for LangSmith-based debugging.

Software Engineer (Contract) — DHL eCommerce UK (now Evri Premium, part of Evri Group) | August 2025 – August 2026

DHL eCommerce UK, now Evri Premium within Evri Group following the October 2025 merger, provide logistics solutions across the United Kingdom. This project is to create a greenfield .Net MAUI mobile application for iOS and Android, running on BYOD devices and Zebra TC58 scanners, used by delivery drivers to manage the loading and delivery process. The app is now in production with drivers across the UK.

- Responsible for app architecture and design, feature development, and the research and application of industry best practices including Modern Clean Architecture
- Responsible for DevOps setup including Entra authentication, Firebase app distribution and Crashlytics
- Created the Amazon Kiro Pilot and subsequent rollout for AI augmented development, used by teams across both DHL eCommerce and Evri Group. Implemented the Kiro steering files, agents and hooks, and the associated governance.
- Produced developer documentation for the Kiro rollout via a Confluence wiki, and delivered briefing demonstrations and videos to support adoption of AI-augmented development.

AI Systems Engineer (Contract via Appmilla) — FinProxy | March 2026 – August 2026

FinProxy are a UK FinTech creating a protection and collaboration platform for Lasting Power of Attorney for financial affairs. I architected and built the proof of concept platform, which connects a donor's bank accounts via Open Banking and checks every transaction against the instructions and preferences in their LPA, with oversight for their attorneys.

- Designed and implemented the FastAPI backend, using guardkit and autobuild to design, spec and build the features. These included the Moneyhub Open Banking integration, LPA document extraction, AI rule-checking and attorney oversight, secured with Keycloak.
- The AI checks run on local hardware with no cloud AI on the critical path. Demonstration runs caught 4 of 4 planted rule-breakers with no false alarms, and gave identical results across repeated runs.
- Delivered two frontends from a single origin, the original HTMX engineering interface and a designed React SPA, along with technology and investor decks and an AWS deployment proposal.

AI Systems Engineer (Contract) — Elm Legal | August 2025

Architected and developed a proof of concept AI-powered legal research system specializing in UK wills, probate, and estate planning. This greenfield project demonstrates advanced AI orchestration for professional legal services, providing real-time legislation access and OSCOLA-compliant legal analysis.

- Designed and implemented a multi-agent architecture surfaced by a Fast API, using LangChain and LangGraph for complex legal workflow orchestration, achieving 95% citation accuracy.

- Integrated real-time UK legislation search via legislation.gov.uk API with multiple search strategies including historical acts (regnal year citations), modern legislation and CLML XML parsing.
- Developed comprehensive legal reasoning engine implementing IRAC methodology (Issue, Rule, Application, Conclusion) with confidence scoring and self-verification loops for professional-grade legal analysis.

Software Engineer, then Lead Software Engineer (Contract) — De La Rue, then Crane Authentication (Crane NXT) following the acquisition | July 2024 – August 2025

DeLaRue provide Authentication Solutions to business and governmental customers. Crane NXT completed its purchase of DeLaRue Authentication in May 2025, and my contract continued through the acquisition across two projects.

Certify Inspector app (July 2024 – April 2025). The Certify Inspector app was a greenfield .Net MAUI mobile app for Android and iOS, used by government tax inspectors in the Arab states of the Persian Gulf. The app is part of the Certify platform and works with the latest tax stamps incorporating Quantum Ink.

- Architected, designed, and implemented the Consumer SDK used by third-party developers.
- Responsible for feature development of the app, and the integration of the Quantum Ink scanner, which involved liaison with the supplier, and the design and architecture of a common solution for the Certify Inspector app and Consumer SDK.

Production Line Activation (May 2025 – August 2025). The PLA project consists of an enterprise-grade, multi-market, multi-tenancy federated system with market-based isolation. This was deployed on Azure virtual machines providing the backend for a customer-facing web app core to the manufacturing and distribution process for Certify tax stamps.

- Designed, and implemented the core framework underpinning all three API's providing a FastEndpoints implementation with functional error handling via the Either Monad and observability via OpenTelemetry integrated with Serilog Structured Logging and ProblemDetails.
- Responsible for the Docker Swarm deployment and Azure DevOps pipelines, and integration with core components including Keycloak for Authentication and YARP Gateways.

Technical Co-founder — Appmilla – .NET MAUI & Open Banking R&D | April 2024 – June 2024

Appmilla is the app development agency I co-founded in 2020, and the vehicle through which I provide contracted services. Between contracts I investigated technologies related to client projects and worked on proof-of-concept and demo apps/experiments including:

- Modern .NET MAUI app series of technical articles and demo apps showcasing Reactive, Functional and declarative UI approaches
- Open Banking API clients and example apps for the Moneyhub and Yapily platforms in .NET

Software Engineer (Contract) — VSL | September 2023 – March 2024

VSL is a start-up company based around an app providing a Web3 sports social platform, engaged at the pre-funding stage for Research & development of the mobile app and server-side applications. Kotlin Multiplatform technology shares much of the implementation between platforms while retaining a native UI using SwiftUI & Jetpack Compose.

- Architecture and implementation of the core cross platform iOS and Android app using Kotlin Multiplatform including User registration and OAuth login using Keycloak, provisioning a Metakeep Web3 wallet and Amity Social views for Feed/Clubs/Live chat.
- Architecture and implementation of the backend API's for AWS using Kotlin with Ktor and Spring Boot Cloud Stream with RabbitMQ.

Software Engineer (Contract) — ClearAccept | June 2022 – September 2023

ClearAccept is the Payments Platform for the ClearCourse group companies; group company TISSL required an embedded payments solution for their EPOS. The TISSL Pay solution is comprised of an Ingenico Axium Dx8000 terminal running Android with the Pay at Table app, enabling table service staff to retrieve orders and take payments with options to split the bill.

- Contracted to develop the native Android Pay-at-Table application in Kotlin and integrate this with the back-end and the Ingenico platform. This was the first such development on the Ingenico platform.
- The app connected to the back-end using gRPC with a bi-directional stream and the technology stack included Coroutines with Structured Concurrency, Flows, XML Views & Hilt.

Software Engineer (Contract) — SSE Energy Services Ltd, then OVO Energy Ltd following the acquisition |

July 2018 – October 2023

SSE Energy Services were one of the UK's 'Big Six' energy providers. I created a new mobile application 'MySSE' for customers, providing features related to billing, payments and meter readings and usage. I architected, designed and developed the app in a very small team. The app was extremely well liked by customers with a rating of 4.8 in the Apple AppStore and 4.6 in the Google Play Store, and had over a million downloads with up to 500K Monthly Active Users. Ovo Energy acquired SSE Energy Services in December 2019; active development continued until June 2022, and I supported the customer migration until we retired the app in October 2023.

- Contracted to design, implement and support the new Xamarin mobile application on Android and iOS. This contract ran from inception through production deployment in December 2019 and finally the app retirement in October 2023.
- Architected the application to use a mix of native and shared UI, using Xamarin Forms content via the Native Forms feature, and implemented behavioural feature testing built on ReactiveUI, testing from the view model through application services to the API queries using Http interception.
- Built a React Native integration hosting the OVO app inside the SSE Xamarin shell, so customers would move to OVO as they migrated, after a week of Callstack React Native training.

EARLIER CAREER

- 09/2017 – 07/2018 Invatech Health — Contract C# Software Engineer
- 08/2016 – 09/2017 Pelican Exchange — Contract Xamarin Developer on iOS & Android
- 08/2013 – 08/2016 United Kingdom Hydrographic Office — Contract C# Software Engineer

TALKS & PUBLIC WORK

- DDD South West 2026 – first conference talk, “2026: The year of the Software Factory”; the Reachy Mini robot demo is published as a YouTube video, along with the slides and repos
- YouTube channel documenting the move from software engineering to AI systems engineering – youtube.com/@RichWoollcott
- Open source: GuardKit (guardkit.ai) and the guardkit GitHub organisation, with fine-tuned models on Hugging Face

EDUCATION & CERTIFICATIONS

- Xamarin University Certification Track – Certified Mobile Developer
- BTEC Higher National Diploma in Computer Studies

TECHNICAL SKILLS

Programming Languages: C#, Python, Kotlin, C++, C++11, C, C++/CLI.

Operating Systems: iOS, Android, Azure, MacOS, Windows, Linux

Frameworks: ASP.NET, FastAPI, .NET MAUI, React Native, Jetpack Compose, SwiftUI, Jetpack libraries, .NET Framework, Xamarin Native and Xamarin Forms on iOS & Android, FastEndpoints, ReactiveUI, DynamicData, Akavache, WPF

AI / Machine Learning Engineering: LangChain, LangGraph, DeepAgents SDK, Claude Agents SDK, MCP, NATS JetStream, ChromaDB, Graphiti with FalkorDB, Unsloth QLoRA fine-tuning, llama-swap, vLLM, Open WebUI, LangSmith, Docling.

Tools: Android Studio, IntelliJ IDEA, Xcode, Rider, Visual Studio, Visual Studio Code, Azure DevOps, GitHub, Git, JIRA, Linear. Agile & SCRUM.

Testing: Automated UI and Unit Testing using TDD/BDD. pytest, Kotlin.Test, Kotest, XUnit, NUnit, Moq, Specflow, Gherkin.

Asynchronous Development: Kotlin coroutines, Python asyncio, Task Parallel Library (TPL)

Database Development: PostgreSQL, Realm/MongoDB, ChromaDB, FalkorDB, Exposed, Microsoft Entity Framework (EF), SQLite, SQL Server

DevOps: CI/CD with Azure DevOps pipelines for mobile and server-side apps, Docker & Docker Swarm, Firebase App Distribution & Crashlytics, AppCenter