

Jayrup Nakawala

London, UK | jayrupnakawala@gmail.com | github.com/CaptainJack2491 | linkedin.com/in/jayrupnakawala | jayrup.me

RESEARCH EXPERIENCE

Dissertation: Moral Cover and Watchful Eyes

github.com/CaptainJack2491/Dissertation

Nov 2025 — May 2026

- Architected a sandboxed evaluation framework with a Virtual File System (VFS) to safely execute and monitor autonomous LLM tool-use across multiple providers (Anthropic, OpenAI, Google).
- Built a provider abstraction layer for glass-box reasoning extraction and a post-hoc interrogation system to adversarially probe agent intent.
- Designed a 3-prong automated judging pipeline (regex, LLM blackbox, LLM glassbox) and ran **960 experimental trials** across exploratory and confirmatory studies.
- Quantified a **“Stealth Gap”**: deceptive intent held steady (62–67%) under high oversight while output-level detectability dropped (47.8%→35.6%), showing output-only monitoring underestimates risk.

PROJECTS

Hybrid-Cloud Research Infrastructure & Zero-Trust Mesh

Jan 2024 — Present

- Developed a hybrid-cloud networking mesh using Tailscale/WireGuard to unify a Dell Precision local node (Ubuntu) with 3 distributed VPS instances into a single secure plane.
- Engineered a custom ingress solution using VPS nodes as forward proxies to bypass CGNAT, ensuring secure, high-availability access to internal research tools without port forwarding.
- Orchestrated 15+ containerized services via Docker Compose, implementing automated backups, resource monitoring, and a centralized NVMe-backed storage tier.
- Optimized a headless development workflow (Arch Linux, Neovim, Tmux) with dotfile versioning for consistent, rapid environment replication across heterogeneous hardware.

Jayrup.me: Fast Personal Portfolio

jayrup.me

Nov 2025 — Present

- Engineered a dual-interface web server using Nginx to serve context-aware content (HTML vs. Plaintext) based on User-Agent detection (e.g., curl, wget).
- Architected a Quarto-based publishing pipeline that generates synchronized web and CLI-friendly documentation from a single Markdown source.
- Automated Deployment: Optimized a secure rsync-over-SSH deployment workflow from a local Arch Linux environment to a distributed VPS cluster.
- Focus: Having a light-weight portfolio that is accessible via the terminal

EDUCATION

University of East London

London, UK

Bachelor's of Science (hons) Data Science and Artificial Intelligence

Sept 2023 — May 2026

- Relevant Coursework: Math for computing, Data Structure and Algorithm, Networking, Database systems, Big Data analysis, Project Management, Artificial Intelligence

SKILLS

Programming: Python, Bash, SQL

ML/AI: PyTorch, Transformers, LangChain, scikit-learn

Infrastructure: Docker, Nginx, Caddy, Tailscale, WireGuard, Quarto

Systems: Linux (Arch, Ubuntu), Git, rsync, Tmux, Neovim