

John Wilkie

London, UK · john1091w@gmail.com · jbwilkie.dev · github.com/JBWilkie

SUMMARY

Software engineer specialising in AI engineering - Developing performant and efficient LLM applications through observability, evaluation pipelines, and context engineering, all built on strong backend and infrastructure engineering foundations. As CTO & co-founder of Nur Opus I delivered and maintained bespoke LLM solutions and evaluation pipelines for multiple clients. I'm currently expanding into LLM post-training including SFT & reinforcement learning methods as an active area of learning and research.

SKILLS

- **LLM engineering:** Evaluation pipeline design (deterministic rules, calibrated LLM judges), tracing & observability (application instrumentation, token/cost/latency), context engineering, guardrail & hallucination reduction, agent regression testing
- **Backend & infra:** Python, Elixir, RESTful API design, PostgreSQL, Docker, CI/CD, AWS, observability (Sentry, PostHog), software testing strategies
- **Engineering:** SDK design & developer experience, distributed systems, Object-Oriented design
- **LLM post-training** (area of active learning): PyTorch, SFT, DPO, LoRA/QLoRA, reinforcement learning, synthetic data and dataset design

EXPERIENCE

CTO & Co-founder - Nur Opus, Dubai

Feb 2025 – April 2026

- Co-founded an LLM consultancy and led engineering - Architecture, infrastructure, tooling, and delivery of bespoke LLM applications, evaluation pipelines, and context-engineering work across multiple verticals
- Recognised recurring problems across numerous client engagements - Evaluation, observability, easy application instrumentation - and started a pivot from bespoke delivery to a scalable product: Orient, a self-hostable LLM engineering platform
- Built Orient's Python SDK and evaluation system: Production calls stream into the platform as traces with token, cost, and latency, continuously evaluated by reusable deterministic and LLM-judge evaluators
- Designed the improvement loop to keep humans in control: Curate failing traces into regression datasets, automatically compare prompt candidates against them, and record upgrade decisions backed by evaluation evidence

Software Engineer (Python) - V7 Labs, London

Oct 2023 – Feb 2025

- Sole author and maintainer of [darwin-py](#), V7's public Python SDK - led feature development and authored extensive client-facing documentation, tutorials, and best-practice guides
- Refactored the SDK's core I/O layer for concurrency, delivering an up-to-8x throughput improvement
- Owned V7's backend processing of 2D & 3D medical imaging formats (DICOM, NIfTI) - triage, bug-fixing, and new platform capabilities for medical files
- Built a new, extensible end-to-end testing engine for the SDK's three core flows (data upload, annotation upload, data download)
- Developed SDK-based API bindings for the features clients most commonly struggled with, producing a measurable drop in related support cases
- Built, deployed, and maintained public integrations with external partners, expanding the addressable market for both businesses

Technical Support → Escalation Engineer - V7 Labs & Verkada

Sep 2021 – Oct 2023

- First support engineer at V7: Built the team's processes, troubleshooting tooling, and client-facing technical content; regular hands-on work with Python, REST APIs, SAML, and AWS/GCP/Azure. Contributions to [darwin-py](#) from this role led to the engineering position

- At Verkada, resolved 2,000+ technical issues and built internal troubleshooting tools, working with engineering to triage software, hardware, and platform problems

EDUCATION

MSc Artificial Intelligence - Distinction - Loughborough University

- Probabilistic & statistical foundations of ML; supervised/unsupervised learning, computer vision, NLP; Python throughout
- Master's project: assessing the predictive quality of datasets for improving autonomous-vehicle safety in pedestrian-rich environments

BEng Mechanical Engineering with Diploma in Industrial Studies - 2:1 - Loughborough University

- Dissertation: Software for autonomous verification of the calibration quality of a 3D surface scanner.

SELECTED PROJECTS & OPEN SOURCE

- [darwin-py](#): Concurrency refactor (up-to-8x speedup) · End-to-end test coverage of the three core flows · Audio-waveform generation for efficient audio labelling · Multi-file grid upload for computer-vision labelling · SDK-based API bindings that measurably reduced support load
- **Orient**: Self-hostable LLM engineering platform - instrument an application with the Python SDK, evaluate it continuously against live production traces, and promote changes on recorded evaluation evidence. Open-source release in preparation
- [Infinite Bazaar](#): Experimental team project in agentic machine economies - AI agents establish verified on-chain identities, publish paid services, and trade with one another autonomously. 3rd prize at Coinbase's "Agents in Action" hackathon (2025); presented at the AWS Builder Loft, San Francisco
- **Writing**: Building Orient in public, LLM engineering in practice, and post-training learned from scratch - jbwilkie.dev/writing