

DION WHITEHEAD, PHD

AI ENGINEER · SCIENTIFIC SOFTWARE ARCHITECT · ML
INFRASTRUCTURE · OPEN SCIENCE



I'm an AI engineer and scientific technologist who builds and leads the infrastructure that turns AI into working science. I have a PhD in Bioinformatics and Evolutionary Biology and 15+ years spanning machine learning platforms, LLM/agent systems, and open-source scientific software going from training pipelines and GPU clusters at Sony AI, to reinforcement learning APIs, to browser-based tools where AI generates interactive scientific visualizations and workflows on demand. I care deeply about AI-accelerated science: making compute, models, and data shareable, reproducible, and usable by researchers. Long-standing interests in complex systems, molecular simulation, evolution, biological networks, deep time, and UX.

- <https://dionjw.dev>
- LinkedIn: [linkedin.com/in/dionwhitehead/](https://www.linkedin.com/in/dionwhitehead/)
- [@dionjwa](#) (GitHub)
- [@metapages](#) (GitHub)
- New Zealand & US citizen
- Lives in Berkeley

Open Source Application Portfolio Highlights

- [framejs.app](#) — AI MCP server and skill combined with platform infrastructure to generate and share interactive browser-based visualizations, gated by OAuth account and user profiles
- [Metapages](#) — shareable scientific workflows in the browser
- [container.mtfm.io](#) — public compute queues for scientific workflows (runs compute jobs in <https://metapage.io>)
- [websocket-router](#) — websocket router for scientific workflows any other application: connect app or device real time with disposable unguessable channels
- github.com/metapages

- github.com/dionjwa

Technology

I build across the technological stack: AI agents, AI skills, AI MCP servers, full-stack web, mobile apps, cloud computing infrastructure, machine learning platforms, and scientific research tools.

Scientific Workflows: High-level focus on building collaborative scientific workflows that are shareable and reusable.

AI-Powered Web: Built browser-based tools that use LLMs to generate interactive visualizations and data dashboards on the fly. Integrated AI generation into collaborative scientific workflows.

AI & Agents: Agentic coding workflows (Claude Code, Cursor, Copilot), multi-agent orchestration, RAG pipelines, prompt engineering, fine-tuning, evaluation frameworks. Daily driver of LLM-assisted development across the full stack.

MCP & Infrastructure: Developed MCP servers for container-based compute, enabling AI agents to provision and orchestrate remote execution environments. Experience bridging AI capabilities with cloud infrastructure.

ML Platform: 5+ years building ML infrastructure at Sony AI — training pipelines, experiment tracking, model serving, GPU cluster management. Full stack from Kubernetes to React dashboards.

Core Stack: TypeScript, Python, Go, Docker, Deno, Node.js. AWS, GCP, Terraform, Kubernetes. React, Preact. Postgres, Redis. Git, CI/CD.

Focus Areas & Keywords

AI & ML: Large language models (LLMs), AI agents, agentic workflows, multi-agent orchestration, Model Context Protocol (MCP), retrieval-augmented generation (RAG), prompt engineering, fine-tuning, model evaluation, reinforcement learning, machine learning infrastructure (MLOps), training pipelines, experiment tracking, model serving, GPU cluster management, AI-generated visualization.

Science & Research: AI for science, AI-accelerated research, computational biology, bioinformatics, molecular simulation & dynamics, scientific workflows, scientific visualization, reproducible research, open science, scientific data infrastructure, drug discovery tooling, biological networks, evolution, complex systems.

Engineering & Platform: Distributed systems, cloud computing, high-performance / batch compute, containerized compute grids, full-stack development, data pipelines, developer tooling, open-source software, technical leadership & architecture.

Experience

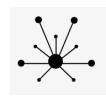


CTO (Fractional)

2025–2026

[Global Ultrasound Institute](#)

Leading the technical direction of the Global Ultrasound Institute, a startup advancing the next generation of point-of-care ultrasound, including AI-assisted tooling for medical imaging, diagnostics, and clinical education.



Scientific Visualization Engineer (UX)

2026

[Open Molecular Software Foundation](#)

Building advanced and shareable scientific visualizations for molecular dynamics open-source software.



Open Science Fellow

2024–2025

[Astera Institute](#)

Founded [metapage.io](#): shareable, reproducible scientific workflows in the browser. Abstracts compute into a universal, open-source grid where any machine or cluster can be plugged in to power your workflows, with AI/LLM generation integrated directly into the pipeline. Combines compute, AI, and visualization all in the browser.



Senior Software Engineer — ML Infrastructure / ML Platform

2019–2024

[Sony AI](#)

Built and architected machine learning infrastructure and researcher-focused tools powering AI research at scale: training pipelines, experiment tracking, distributed model training, GPU cluster orchestration (Kubernetes), and model serving. Full stack from backend ML machinery to React front-end visualization and dashboards for researchers. Co-author on [Sony AI's GT Sophy](#), the deep reinforcement learning agent that outraced champion Gran Turismo drivers (published in **Nature**, 2022).



Senior Software Engineer — Reinforcement Learning Platform

2017–2019

[Cogitai, Inc](#)

Co-designed the technology stack for Cogitai's reinforcement learning and continual-learning APIs, delivering reinforcement-learning models as a cloud service. Built from the ground up, balancing current team abilities with optimal new technology.



Senior Software Engineer

2014–2017

Autodesk Life Sciences

Worked on molecular simulation applications and Genetic Constructor. Integrated scientific tools into more accessible versions.



Lead Mobile Engineer

2012–2014

Idle Gaming

Led a small team that published the iOS and Android versions of Fresh Deck Poker. Developed for both the front-end client and back-end systems, ensuring maximum performance and development velocity.



Software Developer

2008–2010

Three Rings Design, Philadelphia

Prototyped and developed flash social games.



Postdoc

2005–2008

Institute of Evolution & Biodiversity, Münster University

Collaborated with biologists to analyze a large and unique data set, providing unique software tools due to deep understanding of the data, statistical methods, and the possibilities of multiple interconnected software packages.

Education



PhD in Bioinformatics

Bielefeld University, Germany

CeBiTec Graduate School. Grade: 1.0. Thesis



Post-graduate Diploma in Genetics

Massey University, New Zealand



B.Sc. Immunology & Virology

Massey University, New Zealand

Publications

- Outracing champion Gran Turismo drivers with deep reinforcement learning — Wurman, P.R., Barrett, S., Kawamoto, K. et al. **Nature** 602, 223–228 (2022)

- The look-ahead effect of phenotypic mutations — Whitehead, D., Wilke, C., Vernazobres, D. & Bornberg-Bauer, E. **Biology Direct** (2008)
- The AtGenExpress global stress expression data set — Kilian, J., Whitehead, D. et al. **Plant Journal** (2007)
- Reconstructing gene function and gene regulatory networks in prokaryotes — Whitehead, D. **PhD Thesis** (2005)

Full list of publications