

Adrian Wedd

AI Safety Engineer & Independent Researcher

Cygnet, Tasmania, Australia

[GitHub](#) [LinkedIn](#) [Email](#) [PDF](#) [Print](#) [Watch Me Work](#)

ABOUT

AI safety researcher and systems developer building the systems I study. Currently freelancing and shipping independent projects — from a red-teaming framework for embodied AI with 104 tracked research issues, to a voice-first agentic phone system, to a four-site healthcare digital presence under AHPRA advertising regulations, to a multilingual restaurant site in Bali. Previously led systems integration and cybersecurity for Tasmania's public housing sector. Active across 15+ repositories with 1,200+ tracked issues spanning agent orchestration, AI governance, CLI tooling, and client delivery. Background in environmental advocacy, government service, and infrastructure management. Self-taught across the stack. Most interested in where AI systems break.

CORE COMPETENCIES

AI Safety & Evaluation

Red-teaming, adversarial testing, failure-first methodology

Frontier AI Models

Claude API, multi-agent systems, evaluation frameworks

Risk Assessment

Pre-mortem analysis, FMEA, failure mode identification

Policy Translation

Technical findings to actionable governance for decision-makers

EXPERIENCE

Freelance Developer & AI Consultant

Independent 2025 - Present

Delivering client web projects and independent AI research while exploring full-time opportunities. Current client work includes a multi-site healthcare digital presence under AHPRA advertising regulations, with Cliniko API integration and phased project delivery. Concurrently building open-source AI safety tooling, multi-agent frameworks, and CLI automation tools.

- Delivered four-site healthcare digital presence with AHPRA regulatory compliance and Cliniko API integration
- Built and maintained 10+ active open-source repositories spanning agent orchestration, AI governance, and CLI tooling
- Developed failure-first methodology for AI safety evaluation across multi-agent systems

Python, JavaScript, HTML/CSS, Cloudflare Pages, Claude API, GitHub Actions, Cliniko API

Systems Analyst / Acting Senior Change Analyst

Homes Tasmania (formerly Department of Communities Tasmania) 2018 - Feb 2026

Leading systems integration and digital transformation initiatives for Tasmania's public housing sector. Specializing in API development, cybersecurity enhancement, and automation solutions that improve service delivery for vulnerable communities.

- Enhanced Housing Management System integration with external services using RESTful APIs and SFTP
- Led cybersecurity initiatives improving system security and reducing vulnerabilities
- Developed automation scripts in Python, PowerShell and JavaScript streamlining operations
- Pioneered use of generative AI for data analysis and decision-making processes

- Improved operational efficiency through strategic systems integration and process optimization

Python, PowerShell, JavaScript, RESTful APIs, SFTP, SQL, Azure

ITS Client Services Officer

University of Tasmania 2015 - 2018

Provided comprehensive IT support and technical solutions for university staff and students. Specialized in Microsoft technologies, troubleshooting complex technical problems, and developing procedural documentation.

- Delivered first-line IT support resolving complex technical problems across diverse user base
- Developed comprehensive procedural documentation improving support team efficiency
- Supported Microsoft technology stack ensuring optimal performance for academic environment
- Maintained high user satisfaction through prompt problem resolution and clear communication

Microsoft 365, Windows Server, Active Directory, ITIL, ServiceNow

Director

Digital Agency PTY LTD 2015 - 2018

Founded and operated digital marketing agency specializing in helping nonprofits and small businesses maximize their digital impact. Expert-level certified across Google and Bing advertising platforms.

- Enabled nonprofits and small businesses to measure digital effectiveness with Google Analytics
- Attracted new audiences through strategic Google AdWords and Google Ad Grants implementation
- Achieved expert-level certification across Google and Bing advertising platforms

- Delivered measurable digital marketing results for diverse client portfolio

Google Analytics, Google AdWords, Bing Ads, Digital Marketing, Campaign Management

Second Level IT Support Engineer

The Wilderness Society Inc. 2012 - 2015

Managed comprehensive IT infrastructure for national environmental organization. Specialized in lifecycle management of diverse server environments, network infrastructure, and communication systems.

- Managed lifecycle of 60+ Windows/Debian/Ubuntu servers ensuring optimal performance
- Maintained heterogeneous network infrastructure supporting distributed organization
- Administered Google Apps infrastructure and VOIP PBX systems
- Executed complex upgrades, implementations, migrations and decommissioning projects

Windows Server, Debian/Ubuntu, Network Infrastructure, Google Apps, VOIP PBX, Server Management

Communications and Logistics Coordinator

Greenpeace Australia Pacific 2010 - 2012

Coordinated high-profile environmental campaigns combining technology, logistics, and strategic planning. Specialized in risk assessment, security planning, and ICT support for direct action initiatives.

- Planned and coordinated logistical implementation of direct actions and protests
- Developed comprehensive risk assessments and security plans for campaign activities
- Provided ICT support for actions ensuring reliable communication systems
- Delivered activist training programs building organizational capacity

ICT Systems, Communications, Risk Assessment, Security Planning, Logistics Management

PROJECTS

NotebookLM Automation

[GitHub](#)

CLI automation toolkit for Google NotebookLM with research quality controls, deterministic exports, template engine, JSON schema validation, and comprehensive integration testing.

Bash Python CI/CD JSON Schema

Orbitr

[GitHub](#)

Multi-track polyphonic AI sequencer with concentric ring interface. Inspired by Playtronica's Orbita, enhanced with Meta's MusicGen for real-time sample generation. 36 issues tracked across phased delivery.

HTML/CSS/JS MusicGen Web Audio API

Evolve Evolution

Four-site digital presence for a healthcare practitioner: AHPRA-compliant chiropractic clinic site, coaching/personal brand platform, author hub, and business evolution site. Shared CSS design token system, Cliniko API integration for appointment management, GA4-to-GTM analytics migration, Cloudflare Pages deployment with serverless functions, and quarterly regulatory compliance workflow. 170+ issues tracked across phased delivery.

HTML/CSS/JS Cloudflare Pages Cliniko API GA4/GTM Structured Data

Tanda Pizza

Static website for a Bali-based restaurant, localized to 5 languages with WhatsApp ordering integration and gluten-free menu filtering.

HTML/CSS/JS i18n WhatsApp API

Failure-First Embodied AI

[GitHub](#)

Red-teaming and benchmarking framework for embodied and agentic AI systems. Focused on recursive failure analysis, recovery evaluation, and human-in-the-loop safety. Includes automated attack adapters (TAP), cross-modal vulnerability research on VLA models, arXiv paper synthesis pipeline, and a customer discovery track for commercial AI safety services.

Python AI Safety Red-teaming NotebookLM arXiv

TEL3SIS

[GitHub](#)

Real-time telephony platform with LLM-powered conversations, tool use (calendar, SMS, email), tri-layer memory (Redis/SQLite/vectors), safety oracle for pre-execution checks, and human handoff. Built on vocode-python with Whisper, GPT-4, and ElevenLabs.

Python Vocode Whisper Redis Docker

Cygnet

[GitHub](#)

Multi-agent platform coordinating land acquisition, build planning, and 3D print operations for affordable housing. 28+ specialized agents, FastAPI backend, React frontend. 795 issues tracked, 432 commits across Python, TypeScript, Kotlin, and Swift.

Python FastAPI React TypeScript Docker

VERITAS

[GitHub](#)

Legal intelligence platform addressing the efficiency-trust deficit in the legal market. Knowledge graph with Cypher queries, document analysis pipeline, and transparency-first AI reasoning. 230 issues, 784 commits across Python, TypeScript, and Jupyter notebooks.

Python TypeScript Neo4j/Cypher Jupyter Docker

ADHDo

[GitHub](#)

AI assistant for ADHD executive function support with crisis detection, circuit breaker psychology, and local-first processing. NDIS-aware for Australian users. 100 issues, 73 commits.

Python AI Safety Accessibility

NeuroConnect

[GitHub](#)

24/7 voice helpline with wired STT-to-LLM-to-TTS pipeline over Twilio, NDIS evidence logging, and executive function scaffolding. Multi-provider LLM fallback chain including Claude CLI.

Python Twilio Whisper Docker

RLM-MCP

[GitHub](#)

MCP server enabling Claude Code to process 1M+ character documents through session-based chunking, BM25 search, and artifact provenance tracking.

Python MCP BM25 Claude Code

Agentic Research Engine

[GitHub](#)

Next-generation multi-agent research system with LangGraph orchestration, self-correction loops, and autonomous learning. 1,001 commits across Python with CI/CD cost optimization and enterprise security integration.

Python LangGraph Multi-Agent Systems Docker

SKILLS

Programming Languages

Python	Primary
JavaScript	Primary
TypeScript	Secondary
PowerShell	Secondary

AI & Automation

LLM Integration	Primary
AI Safety & Evaluation	Primary
Process Automation	Primary
MCP Development	Secondary
Voice/Telephony AI	Secondary

Infrastructure

Systems Integration	Primary
Cybersecurity	Primary
Linux/Windows Server	Primary

DevOps

Docker	Secondary
--------	-----------

GitHub Actions Primary

Cloudflare Pages Secondary

Frontend

React Secondary

Data

SQL / PostgreSQL Primary

ACHIEVEMENTS

Systems Integration Excellence

Enhanced Housing Management System integration with external services using RESTful APIs and SFTP, improving data exchange and operational efficiency for Tasmania's public housing sector.

2018-2024

Cybersecurity Leadership

Led cybersecurity initiatives at Homes Tasmania, improving system security posture and reducing vulnerabilities across critical housing management systems.

2020-2024

AI Safety Research & Red-Teaming

Developed failure-first red-teaming framework for embodied AI systems with 104 tracked research issues, cross-modal vulnerability analysis on VLA models, and automated attack adapters. Pursuing commercial AI safety services.

2025-2026

Voice-First AI Agent Platform

Built TEL3SIS, a real-time telephony platform with LLM-powered conversations, tri-layer memory architecture, safety oracle, and tool orchestration across 226 tracked issues.

2024-2026

Environmental Campaign Technology

Managed IT infrastructure for The Wilderness Society and coordinated Greenpeace campaigns, combining technology expertise with environmental advocacy.

2010-2015

Professional Certifications

Achieved certifications across Google Analytics, AdWords, and Bing Ads platforms.

2015

Automation & Process Improvement

Developed automation scripts in Python, PowerShell, and JavaScript that streamlined operations across multiple organizations.

2018-2024

Adrian Wedd

AI Safety Engineer & Independent Researcher

Last updated: Feb 26, 2026, 07:12 AM

© 2026 Adrian Wedd