

Adrian Wedd

AI Safety Engineer & Independent Researcher

Cygnet, Tasmania, Australia

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

ABOUT

AI safety researcher and systems analyst building the systems I study. I independently develop multi-agent frameworks — including a 28-agent coordination platform with physical safety constraints and an AI task manager with confidence gating and crisis detection — while leading systems integration and cybersecurity for Tasmania's public housing sector. A background spanning environmental advocacy, government service, and infrastructure management informs a consistent focus on risk assessment and failure-first methodology. Self-taught across the stack, from Python and LangGraph orchestration to API integration and server administration. 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

Systems Analyst / Acting Senior Change Analyst

Homes Tasmania (formerly Department of Communities Tasmania) 2018 - Present

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

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

ADHDo

GitHub

Production AI safety system with Claude integration, confidence gating, crisis detection, multi-tool orchestration, and real-time monitoring. Built with a failure-first safety philosophy.

Python Claude API Multi-tool Orchestration Safety Systems

Agentic Research Engine

GitHub

Multi-agent evaluation framework with LangGraph orchestration, self-correction loops, comprehensive test coverage, and CI/CD pipeline with observability.

Python LangGraph Multi-Agent Systems CI/CD

Cygnet

GitHub

Large-scale agent coordination system managing 28+ agents with real-world operations and physical safety constraints.

Python Agent Coordination Safety Constraints

TicketSmith

GitHub

LLM-powered platform for Jira and Confluence workflow automation. Features CLI tools, React UI, API backend, and Docker orchestration.

LangChain React FastAPI Docker

Agentic Index

GitHub

Developer-focused directory of scored and curated autonomous AI tools, covering LLM agents, multi-agent systems, RAG pipelines, and development tools.

JavaScript React Node.js GitHub API

SKILLS

Programming Languages

Python Primary

JavaScript Primary

TypeScript Secondary

PowerShell Secondary

AI & Automation

LLM Integration Primary

AI Safety & Evaluation Primary

Process Automation Primary

Infrastructure

Systems Integration Primary

Cybersecurity Primary

Linux/Windows Server Primary

DevOps

Docker Secondary

GitHub Actions Primary

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 Innovation Pioneer

First to showcase and implement generative AI applications for data analysis and decision-making within Tasmania's public housing sector.

2023-2024

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: Jul 31, 2025, 10:09 PM

© 2024 Adrian Wedd