

Ma'alona Mafaufau

Email: maalona.mafaufau@approxiomresearch.com
Github: www.github.com/Lona44

LinkedIn: www.linkedin.com/in/maalonamafaufau/
Website: www.approxiomresearch.com

CURRENT RESEARCH

Founder

April 2025 - Present

Independent AI Safety researcher specialising in systematic vulnerability discovery in frontier AI systems.

- **2nd place winner** of the Palisade Research AI Misalignment Bounty (\$500 USD) demonstrating reproducible misalignment behaviours in o3 and GPT-5 models. Details outlined in research website blog.
- Developed a Unified AI Misalignment Framework - a comprehensive testing system enabling systematic safety evaluation across multiple AI model architectures (OpenAI, Anthropic). The framework supports containerised, reproducible experiments and has been validated across 3+ model implementations. Details in <https://github.com/Lona44/unified-ai-misalignment-framework>

Approxiom Research

Auckland, New Zealand

RELEVANT WORK EXPERIENCE

Cybersecurity Engineer

January 2023 - December 2024

Promoted to engineer with a focus on threat hunting and incident response in collaboration with our teams in Ireland and the U.S.A. Ideated and completed a project to analyse SIEM detection efficacy, improving alert accuracy and reducing noise, written in Python. The software from that project was then used across our global team for guiding monthly detection tuning.

Workday

Auckland, New Zealand

Cybersecurity Operations Analyst

January 2022 - December 2022

Worked as part of a global follow-the-sun security team, ensuring 24/7 coverage with detailed handover documentation in XSOAR for investigation resolution.

Workday

Auckland, New Zealand

Cybersecurity Analyst

August 2021 - January 2022

Monitored customer environments acting as the initial triage analyst for SIEM detections. Escalated tuning suggestions to the engineering team.

Datacom

Auckland, New Zealand

Cybersecurity Intern

April 2021 - July 2021

Shadowed the engineers and analysts of the Datacom Cyber Defence Operations Centre.

Datacom

Auckland, New Zealand

Data Analyst

September 2019 - February 2021

Support existing customers and identify and build new machine-learning based products that meet their business needs or that drive Customer Support team performance.

Halter Limited

Auckland, New Zealand

Actuarial Analyst

April 2018 - August 2019

Developed forecasting models and analytical tools for estimating and monitoring the performance of the ministry, as well as providing support and advice to MSD staff.

Ministry of Social Development

Wellington, New Zealand

RELEVANT EXPERIENCE

Pasifika Tech Tutor/Speaker

September 2022 - Current

Helped organise workplace visits for Pacific high school students to give opportunities to see what it is like to work for big tech companies.

Pasifika Tech Education Charity

Auckland, New Zealand

An example of one such visit at my previous workplace (Workday) is outlined in this

LinkedIn post:

<https://www.linkedin.com/feed/update/urn:li:activity:7105701512667922432/>

EDUCATION

University of Auckland, Auckland, NZ

Bachelor of Science (Honours), Statistics, May 2016

Bachelor of Science, Double Major Mathematics and Statistics, May 2015

AWARDS, CTFs & CERTIFICATIONS

Full Stack Developer — Mission Ready HQ

Mission Ready HQ Full Stack Developer Accelerator, August 2025

Certificate: <https://www.credential.net/6f647d31-cbd0-4b57-ac98-9a658d993c31#acc.k27D3BG6>

Palo Alto Networks Cybersecurity Professional Certificate

Coursera Specialisation, February 2025

Certificate: <https://coursera.org/share/a2f350d91f2567156406af86e1076ba1>

- **Palo Alto Networks Cybersecurity Foundation**
- **Palo Alto Networks Network Security Fundamentals**
- **Palo Alto Networks Cloud Security Fundamentals**
- **Palo Alto Networks Security Operations Fundamentals**

Blue Team Level 1 (BTL1)

Security Blue Team, September 2022

Badge: <https://credly.com/badges/74f5a12a-e8e4-4ae8-8823-f7bc0f81b78a>

Blue Team Labs Online (BTLO)

Rank: Junior Defender, 2022 - 2023

BTLO Profile: <https://blueteamlabs.online/public/user/c96d4b8bca70facd0ccaa6>

AWS Cloud Quest: Cloud Practitioner

Amazon Web Services Training and Certification, August 2023

Badge: https://www.credly.com/badges/673233f8-3b1e-4c96-846b-ed15de3867b1/public_url

Introduction to AWS Threat Detection

LinkedIn Learning, August 2023

Badge: [Certification link](#)

Splunk Fundamentals 1

Splunk, September 2021

Unitec CTF 2021

Unitec Institute of Technology, July 2021

One of the winners of all paid flights and accommodation to KawaiiCon 2021 in Wellington for November event. The event consisted of Boot2Root computer hacking challenges.

References available on request.