

MARC-DONALD TONGA NOUDJA

AI Red Team Engineer · LLM Security Researcher · AI Safety & Evaluation

Yaoundé, Cameroon (Open to Remote, Worldwide) | +237 659 382 005 | tonganoudjamarcdonald@gmail.com
linkedin.com/in/tonga-marc-donald-3b437738b | github.com/T-MARC-DONALD | limitless-cyberhaven.vercel.app

PROFILE SUMMARY

AI Red Team Engineer and LLM Security Researcher finishing a B.Sc. in Cybersecurity (ICT University, expected 2026) with seven years of freelance security engineering experience. Author of *PromptShield*, an automated LLM red-teaming suite that surfaced a novel “encoding paradox” across 11 providers; also built *redforge* and *AI-IDS*. Fluent in English and French.

CORE COMPETENCIES

- **LLM Security:** Prompt Injection, Jailbreaks, Obfuscation, Prompt Extraction, Multi-Turn Escalation
- **AI Evaluation & Safety:** Red Team Eval, Rubric Scoring, Preference Ranking, Safety Labeling, LLM-as-a-Judge
- **Adversarial ML:** OWASP LLM Top 10, MITRE ATLAS, AI Guardrails, Data Poisoning Defense, Fuzzing
- **AI Governance & Risk:** AI Governance, Responsible AI, AI Risk Management, AI Incident Response
- **Security & Networking:** Penetration Testing, Intrusion Detection (IDS), Threat Modeling (STRIDE), Vulnerability Scanning
- **Programming & ML:** Python, JavaScript, TypeScript, SQL, PyTorch, scikit-learn, Hugging Face Transformers
- **Engineering & Cloud:** Docker, Git, Linux, Flask, Django, React, Node.js, REST APIs, AWS, Azure, GCP, Splunk

PROFESSIONAL EXPERIENCE

AI Red Team Engineer & LLM Evaluation Researcher, PromptShield (Independent Research)

Remote | 2025 to 2026

- Uncovered a novel “encoding paradox” vulnerability in frontier LLMs, as measured by a 66.7% encoding-attack success rate on Nemotron Ultra 550B versus 40.0% on Llama 3.3 70B, by engineering 55 adversarial prompts across eight injection categories and evaluating 67 model configurations across 11 providers.

AI Red Team Tooling Engineer, redforge (Independent Open-Source Project)

Remote | 2026

- Built an automated LLM red-teaming toolkit that runs 10 attack families against any model endpoint, as measured by ~1,800 lines of code across 31 modules with four pytest suites, by creating a provider-agnostic adapter layer over async httpx with auto-discovery attack registration.

Machine Learning Security Engineer, AI-IDS (Independent Research)

Remote | 2024 to 2025

- Achieved near-perfect F1 classification of 14 network attack classes, as measured by 2.8M+ labeled flows processed with 180% expanded attack-class coverage, by training a two-stage Random Forest pipeline on CIC-IDS2017 and building a real-time flow assembly module in Python and Scapy.

Cybersecurity Tool Developer, Automatic-Porter, Vulnera & NMS (Open Source)

Remote | 2024 to 2025

- Delivered three open-source security tools for port visibility, vulnerability scanning, and network anomaly monitoring, as measured by 15 maintained public repositories plus GitHub Pull Shark and YOLO achievements, by designing each tool around real-time detection and alerting workflows.

Freelance AI Consultant & Full-Stack Developer

Self-Employed | Remote | 2018 to Present

- Shipped 10+ production web applications with zero reported security incidents across seven years of bilingual client engagements by integrating input validation, CSRF protection, secure authentication, and API hardening into every build.

EDUCATION

B.Sc. in Cybersecurity, ICT University, Yaoundé, Cameroon | Expected 2026

Relevant coursework: Machine Learning, Network Security, Cryptography & Cryptanalysis, Information Systems Security, Linux Administration, Digital Forensics.

Web Development Certification, ISTIA Cameroon | 2015 to 2018 | Honor Roll, top of class three years running

CERTIFICATIONS

AI Security: Security in the Age of Artificial Intelligence, Coursera | Aug 2026

Microsoft Enterprise AI Governance, Ethics & Security, Microsoft | Sep 2026

AI Security Fundamentals: LLM Threats & OWASP 2026, Packt | 2026