SAMUEL BAGUMA

Mountain House, CA 95391 • 202-631-7391 • kabsam99@gmail.com • linkedin.com/in/samuel-baguma-37a45110/

ML/AI Security Engineer | Adversarial ML Defense | Secure LLM Deployment

Accomplished security engineer specializing in Trustworthy AI and MLSecOps. Hands-on experience in securing AI/ML systems through threat modeling, adversarial testing, data protection, and deploying robust defenses against evolving attack vectors. Skilled in designing AI security controls, model monitoring pipelines, and multi-agent defenses. Passionate about building secure, fair, explainable, and privacy-preserving AI systems.

SKILLS

ML/AI TOOLS & FRAMEWORKS: Scikit-learn, TensorFlow, PyTorch (basic), Hugging Face, LangChain, FAISS, Gradio, OpenAI API

MLSECOPS TOOLS: Adversarial Robustness Toolbox, TextAttack, ModelScan, NeMo-Guardials, Agentic Security

CONCEPTS: NLP, Transformers, RAG, Prompt Engineering, LLM Agents, Anomaly Detection, Feature Engineering, Vector Search

LANGUAGES/STANDARDS: Python, SQL, NIST AI RMF, ISO/IEC 42001, and MITRE ATLAS, OWASP AI

Machine Learning Projects

HealthBot – Medical Symptom Checker | *Personal Project, 2024*

[Live demo: huggingface.co/spaces/sbaguma/HealthBot]

- Built and deployed a LLM-powered **RAG chatbot using LangChain, Transformers, and FAISS** for medical knowledge retrieval.
- Implemented prompt injection safeguards and model output sanitization to defend against manipulation attacks.
- Designed an explainable reasoning pipeline using a ReAct-style agent framework for secure multi-step decision-making.
- Simulated **prompt injection attack** for Denial of Service, phishing and jailbreak

ML-Powered Log Anomaly Detection | ***Personal Project**, 2024*

- Developed a **multi-agent anomaly detection model** to identify security threats in SIEM logs using Scikit-learn and TensorFlow.
- Achieved **87% precision** in filtering real threats from noise, reducing analyst fatigue by **~40%**.
- Simulated **adversarial evasion attacks** and enhanced robustness via defensive distillation techniques.

Professional Experience

Information Security Engineer | Patelco Credit Union, Dublin, CA • 2024 – Present

- Conduct adversarial model scanning and vulnerability assessments on ML pipelines, strengthening model integrity.
- Designed defenses against model extraction, inversion, and data poisoning using differential privacy techniques.
- Led AI Red Teaming exercises, simulating prompt injection, model hijacking, and adversarial attacks.
- Built continuous model monitoring dashboards to detect performance drift and anomalous outputs in production.

Manager, Security Operations |23andMe, Sunnyvale, CA • 2022 to 2024

- Established the company's **first AI security operations program**, integrating MLSecOps into enterprise security strategy.
- Conducted **AI-specific threat modeling** with STRIDE and MITRE ATLAS to secure LLM and genomic models.
- Leveraged **Adversarial Robustness Toolbox (ART)** to detect and mitigate **evasion, poisoning, and inversion attacks**.
- Automated **model vulnerability scanning** and secured LLMs against **prompt injection** and **DoS attacks**.

Lead Security Engineer |Optimize Health, Remote, CA • 2021 to 2022

- Led the secure deployment of cloud-native health platforms, integrating ML-enabled threat detection tools such as AWS GuardDuty and Inspector.
- Collaborated on designing a HITRUST-certified architecture featuring ML-powered audit logging pipelines and automated compliance rule engines.
- Improved incident response readiness by developing and executing a comprehensive incident response plan.

Senior Security Engineer |23andMe, Sunnyvale, CA • 2020 to 2021

- Strengthened corporate IT security by implementing and managing key infrastructure security solutions.
- Optimized threat detection capabilities by configuring SIEMs and enhancing log correlation and alerting processes.
- Deployed advanced threat intelligence products and developed detailed threat reports to inform risk management strategies.

Senior Security Analyst| Walmart eCommerce (via Fisec), Sunnyvale, CA • 2019 to 2020

Addressed vulnerabilities by utilizing Static (SAST) and Dynamic (DAST) techniques to review application code for potential security risks. Augmented application protection by deploying Run-Time Application Self-Protection (RASP) and anti-bot techniques to reinforce security.

Security Analyst |Google (on-site via Accenture), Mountain View, CA • 2018 to 2019

Catapulted application security by conducting code reviews via static and dynamic analysis techniques to identify potential vulnerabilities. Enhanced malware defense through application reverse engineering in support of detailed malware analysis efforts.

(Previous roles at Ministry of Finance, and Pride Microfinance available upon request)

Educational Background

Master of Engineering in Cybersecurity(Machine Learning Minor) | University of Maryland, College Park, MD

Master of Science in Computing | Makerere University, Kampala, Uganda

Bachelors of Science in Statistics | Makerere University, Kampala, Uganda

Professional Development

Machine Learning & AI: Machine Learning Engineer (2025), Meta Back-End Developer, Cryptography (UMD, 2024)

Cloud & Security: AWS Certified Security – Specialty (2023), CCSFP (2022), CISSP (2017), Splunk Certified User (2018)

Architecture & Frameworks: Architecting with GCP (2018), TOGAF9, COBIT5, ITIL Expert (2014)

Professional Affiliation

Member, Project Management Institute, 2011 to Present

Member, Certified Information Systems Security Professional (CISSP), 2017 to Present