JOSHUA URIANZA

SENIOR AI SECURITY ENGINEER | ACTIVE TS/SCI

Salt Lake City, UT • 385-985-8433 • jkurianza@gmail.com • linkedin.com/in/joshua-urianza • github.com/RootReap3r • rootreap3r.com

PROFESSIONAL SUMMARY

Senior AI security engineer hands-on with generative AI since 2020 and with AI jailbreaks and prompt injection since 2022. Specializes in MCP server security architecture, agentic system threat modeling, and prompt-injection defense: building code-enforced tool-permission boundaries and adversarial evaluation harnesses resistant to injection-based scope escalation. 10+ years in cybersecurity, including 7+ years at the TS/SCI level across nuclear C2, SAP/TEMPEST, and federal cloud environments, with a track record of finding zero-days missed by enterprise tooling through manual exploit analysis. M.S. AI candidate at Georgia Tech researching adversarial robustness in agentic systems.

CORE SKILLS

AI Security: MCP security architecture, prompt injection (direct, indirect, XPIA), jailbreaking, tool poisoning, agentic threat modeling, LLM red teaming, RAG security, AI incident containment

Frameworks: OWASP Top 10 (LLM, MCP, API, Web), MITRE ATLAS and ATT&CK, NIST AI RMF, NIST 800-53 Rev 5, RMF/ATO

Cloud & Platform: AWS (IAM, Bedrock Agents), Azure (Sentinel, Entra, Defender for Cloud, Purview), GCP, Kubernetes (RBAC, admission control), Terraform, multi-tenant isolation

Detection & Telemetry: Splunk, Microsoft Sentinel, OpenTelemetry (GenAI semantic conventions), CloudWatch, Sysmon, Zeek, STIX 2.1

Languages: Python, Go, Bash, PowerShell

EXPERIENCE

Senior AI Security Engineer, Wilson Sonsini Goodrich & Rosati (WSGR) *Sep 2026 to Present*

AI Red Teamer / AI Security Engineer (Contract), Mercor | LinkedIn *Jul 2026 to Sep 2026*

- Landed 100+ paid, reviewer-validated prompt-injection and jailbreak findings (direct and cross-domain) across client LLM products, scored against a structured red-team taxonomy.
- Surfaced MCP tool poisoning, schema injection, and cross-server privilege escalation by exploiting over-privileged servers across client AI tool infrastructure.
- Built a multi-tenant AI chat widget, embeddable on any site in under 30 minutes, with CORS allowlisting, prompt-injection screening, XSS-safe rendering, CSRF protection, and timing-safe auth.

Senior Cybersecurity Engineer, Accenture Federal Services *Oct 2022 to Aug 2026*

- **Zero-Day Discovery & Detection:** Drove accreditation for a $147B TS/SAP cloud program by manually reverse-engineering zero-days missed by enterprise tooling, building an AI correlation engine (NVD/CVE/CNNVD) to automate triage, and engineering Splunk detections that cut C2 dwell time from hours to under 2 minutes, securing full ATO across 10 NIST Rev 5 control packages.
- **Threat & Risk Assessment:** Owned adversary-informed assessment and IR exercise programs across all four classification tiers; ran 2,730 assessments correlating 524 vulnerability scans and 476 audits, and built 16 quarterly adversarial scenarios from threat model through after-action report.
- **Compliance & Governance:** Authored accreditation packages (~2,500 controls, SCTMs, SSPs) as Lead ISSO for Sentinel and Nuclear Command & Control programs.

Security Program Manager, DS&S *Oct 2020 to Oct 2022*

- **Personnel & Insider Threat:** Directed personnel security and insider threat for 2,350+ AFNWC personnel with TS/SAP access, including counterintelligence vetting, clearance adjudication, and need-to-know validation for Restricted Data; managed direct reports, hiring, and performance reviews.
- **Physical & Nuclear C2 Security:** Owned security for $41M+ in SAP/TEMPEST facilities and networks carrying NC2 design data and DOE Restricted Data, running adversarial assessments across every control layer with zero attributable incidents.

Cyber Automation Lead, Sebastian Tech Solutions (DISA) *Mar 2020 to Oct 2020*

- Built SOAR and PowerShell/Python automation for account remediation, access provisioning, and IR triage, cutting escalation decisions on compromise cases from minutes to seconds across DISA, White House, Pentagon, and 7th Fleet environments.

Cyber Investigations Lead, Netflix *Apr 2018 to Mar 2019*

- Led response to a credential-stuffing campaign compromising 900+ accounts, correlating device, IP, geolocation, and VPN signals to restore 100% of accounts to their owners.

Incident Response Lead, The Results Company *Aug 2015 to Apr 2018*

- Advanced from help desk analyst to IR Lead, resolving 10,264 tickets and coordinating a multi-person IR team while sustaining a 96% stakeholder satisfaction score.

MILITARY SERVICE

Weapons Director, Air National Guard *Jan 2017 to Jan 2023*

- Directed up to 30 aircraft ($2B+ in F-22/F-35/F-15/F-18 assets) across 100+ sorties in live coalition exercises including Red Flag-Alaska and Northern Lightning, on full-time orders Mar 2019 to Mar 2020.

PROJECTS

- **Wraith** | *AI Red-Team Engagement Harness.* Python harness unifying AI red-team tradecraft across 11 attack modules (prompt injection and jailbreaks, MCP/RAG/A2A, adversarial-ML evasion, data poisoning, membership inference, model supply chain, agent SSRF), with a code-enforced scope gate, verbatim request/response evidence capture, scoring and report scaffolding, and an MCP server so agents can drive it; 260 tests.
- **Seraph** | *AI Defense Control Verification Library.* Blue-team counterpart to Wraith: 15 detection and guard modules (indirect injection, taint-based exfiltration, MCP privilege boundaries, agent egress SSRF, supply-chain startup hooks, agent tampering with monitoring) mapped to a shared attack taxonomy, each shipping measured true- and false-positive rates; includes a deployable egress interceptor validated end to end against Wraith's SSRF payloads; 160 tests.
- **RMF-AI-COPILOT** | *NIST 800-53 Rev 4 to Rev 5 Migration Pipeline.* Generates SCTM, eMASS export, and POA&M items from SSP content, grounding every AI judgment against the full 1,189-entry Rev 5 catalog; provider-agnostic (Anthropic, OpenAI, Bedrock GovCloud) and deployable air-gapped/NIPR with zero code changes.
- **AI IR Tabletop** | *Adversarial LLM Incident-Response Engine.* LLM plays a state-consistent threat actor across 23 scenarios and 9 MITRE-mapped APT profiles, with attacker and grader split across separate API passes and auto-generated after-action reports, cutting exercise prep from weeks to hours.

CERTIFICATIONS & AWARDS

AI Security: HTB COAE (#213 worldwide) | CompTIA SecAI+ (beta cohort) | OSAI+ (in progress) | Azure AI Fundamentals

Offensive: PWPE (<100 holders worldwide) | PNPT | eCPPTv3 | CompTIA PenTest+

Core: CompTIA SecurityX (CASP+) | CompTIA Security+

Awards: Security Team of the Year (AFNWC) | Minuteman III Personnel Award | ICBM Personnel Award

EDUCATION

M.S. Artificial Intelligence, Georgia Institute of Technology, Atlanta, GA

In progress, part-time. Research: adversarial robustness and red teaming of LLM agentic systems.

B.A.S. Cyber Operations, Defense & Forensics, University of Arizona, Tucson, AZ

NSA Center of Academic Excellence in Cyber Operations; malware reverse engineering and digital forensics.