

MO SADEK

Technical Executive — AI Security, Governance & Strategy

Zero-to-one program building · AI risk & governance · executive and industry translation

San Francisco Bay Area · mo@makesecurity.fun

linkedin.com/in/msadek1 · blog.makesecurity.fun

SUMMARY

A technical leader defined by range rather than a single lane. Deep enough in security and AI systems to do the work personally, from threat modeling to evaluation to prototyping, and organizationally fluent enough to build the structures around it: governance, cross-functional programs, executive alignment. That same investment extends beyond the day job, into open-source tooling, industry standards, and the policy conversations shaping how AI gets deployed. What ties it together is translation, getting researchers, engineers, and executives moving on the same problem, with a standing preference for showing an idea rather than describing it.

SELECTED IMPACT

- **Built a security function from zero at consumer scale.** Early security hire at a platform serving hundreds of millions of users, growing the application-security practice **3 → 12 engineers** and helping steward the wider security organization **7 → 90** through the company's direct public listing.
- **Held an executive staff seat.** Chief of Staff to a Chief Scientist, chairing companywide AI governance and carrying research strategy into C-level rooms across engineering, legal, policy, and product.
- **Sets industry standards, not just internal ones.** Co-lead of a risk area in the flagship peer-reviewed framework for securing autonomous AI agents, and a member of the AI security coalitions writing the field's shared vocabulary.
- **Ships as well as governs.** Author of open-source security tooling for the AI-agent ecosystem, and a hands-on builder of internal AI products adopted by other functions.
- **Operates in public.** Conference stages, a hosted podcast, published research, and a personal writing practice under the "Make Security Fun" banner.

EXPERIENCE

Alice (formerly ActiveFence) — Technical Director

Jul 2025 – Present · San Francisco Bay Area *AI security company. Technical authority bridging research, product, and go-to-market.*

- Co-led the **Rogue Agents (ASI10)** risk area of the **OWASP Top 10 for Agentic Applications (2026)**, a globally peer-reviewed framework for securing autonomous AI agents, referenced by major vendors and by NIST.
- Designed and shipped **Caterpillar**, an open-source scanner that audits **AI agent skills and MCP tool configurations** for malicious behavior before installation. Configurable **LLM-as-a-judge or rule-based** analysis, offline-capable, with CI/CD gating and A-F grading.
- Established a **security-research function** running offensive testing against AI systems and responsible disclosure, converting findings into product requirements and published research.
- Operates an **"inform-through-prototyping"** model with Product and Engineering: turns customer and field signal into working prototypes, then presents them to product leadership as roadmap and net-new feature recommendations.
- Built internal AI tooling adopted across Marketing and Product, including a **RAG-powered browser extension** for industry-specific insight, reusable **agent skills**, and an AI-native content platform with **automated claim fact-checking** and persona-based tonal analysis.
- Established the company's presence across **OWASP, CoSAI, and ISACA** during the GenAI-security market transition, and represents technical direction on conference panels covering guardrails, content filtering, and agentic system design.
- Reframed audience and positioning strategy rather than output volume, producing a **176% increase in content viewership and 58% click-through within 30 days**.

Roblox — Chief of Staff to the Chief Scientist

Jun 2024 – May 2025 *Executive strategy seat spanning AI, Security, Legal, Policy, and Engine organizations.*

- Served as the Chief Scientist's **proxy in executive meetings** and author of executive communications, connecting research direction to engineering execution across the company.
- Established and chaired the companywide **AI Review Board**, running cross-functional reviews of **80+ AI initiatives** in the first year against a risk framework spanning **cost, ethics, performance, infrastructure, and age-appropriate user safety**, with **automated intake and triage** across information security, privacy, patent, and IP.
- Led evaluation of **third-party AI integrations and open-source model implementations** against internal standards, and partnered with legal and compliance on alignment with evolving international AI regulation.
- Built a companywide **patent and IP program**: an automated invention-disclosure system, a **monthly patent showcase**, and an engineering career-advancement framework tying **patent creation to technical-leadership progression at Director+ levels**.
- Advised senior directors on **game-engine feature development** off the back of competitive market analysis, and delivered engine-roadmap feedback in C-level meetings alongside the Chief Scientist.
- Built research-organization infrastructure from zero: a new **research job classification** created with VPs and recruiting to attract academic talent, and an end-to-end **researcher-management system** automating candidate intake, access provisioning, and off-boarding.
- Led program direction for a **Lua-native web layer** replacing Chromium implementations, and established a **device-performance and OOM metrics framework** with data science, including emerging-markets research on low-end devices.

Roblox — Sr. Application Security Engineer (Security Partner)

Sep 2019 – Jun 2024 *Early security hire through the company's hypergrowth years and direct listing; helped build the core security programs from the ground up.*

- Built the **application-security practice from scratch**, covering incident response, live incident leadership, CI/CD integration, code review, and architecture assessment.
- Created a **security partnership program** that made AppSec the front door for engineering, **growing the function 3 → 12**. Led hiring through the first ~40 InfoSec engineers, then owned onboarding and mentorship as the organization scaled **7 → 90**.
- Conceptualized and ran "**Trust by Design**" as the InfoSec representative, delivering **120+ high-level compliance and safety reviews**, and led early regulatory conversations as the company prepared for its direct listing.
- Led **end-to-end security program management for AI features across 4 teams and 15 major features**, covering review, testing, and mitigation.
- Directed a **platform-wide safety-system redesign spanning 30+ components, 7 engineering teams, and 3 organizations**, and ran threat modeling for a **voice-chat safety system**, region-specific applications for China and Vietnam, certificate pinning for education products, and an authentication refactor supporting console quick-login and LMS SSO.
- **Re-architected role-based access control** for the trust and safety platform, rationalizing **400+ privilege role IDs** governing a **5,000+ person outsourced workforce** across multiple vendors toward least privilege.
- Built and open-sourced (in Go) a **security-scanning platform** orchestrating SAST, DAST, secret scanning, dependency checks, and container-image validation for teams hosting outside the core environment, relaying results back into central dashboards. Established the org-wide open-source security strategy and governance framework alongside Engineering Efficiency.
- Took over the **bug-bounty and vulnerability-disclosure program** in 2020 and turned it around against hundreds of incoming reports: **response rate 60% → 90%**, average closing time from over a month to **~2 weeks**.
- Led and supported incident response across trust and safety threat domains including account takeover, insider threat, and anomalous transactions.

Bosch — Security Engineer, Bosch Center for AI Research

Nov 2018 – Aug 2019 *Bosch's AI research and enablement division, chartered to help the wider business adopt AI at scale.*

Architected the security framework for the AI research and enablement environment. Threat-modeled **autonomous-driving systems** and their software supply chain, built a **data-integrity verification system** for international ML manufacturing under strict regional data regulation, developed **Kubernetes security policies and RBAC for a GPU cluster**, stood up an early **ML-for-SOC threat-detection capability** with the SOC team ahead of broad vendor adoption, and led **ISO 27001** compliance.

Warner Brothers — Lead Application Security Engineer

Apr 2018 – Oct 2018 Restructured the security-assessment program around risk-based prioritization focused on architecture reviews, strengthened third-party risk processes and security training delivery, and mentored junior engineers.

NBCUniversal — Sr. Application Security Analyst

Sep 2016 – Mar 2018 Led security testing across **broadcast and on-air systems** including a **PCI compliance** program, built a mobile-application review program from the ground up, and developed custom scanning solutions that improved remediation tracking.

Cigital (now Synopsys) — Security Consultant

2015 – 2016 Delivered OWASP Top 10 vulnerability assessments and remediation guidance for client applications, and created consultant training programs on vulnerability discovery, exploitation, and mitigation.

LEADERSHIP & TECHNICAL RANGE

Organizational building functions from zero · hiring and scaling security organizations · executive operations and chief-of-staff work · governance and review boards · mentorship and onboarding at scale

AI security agentic and AI-agent security · MCP tooling · LLM evaluation and LLM-as-a-judge · offensive AI testing and red-teaming · AI risk frameworks · guardrails and content filtering · RAG systems

Security engineering threat modeling and architecture review · incident response · CI/CD and DevSecOps · SAST, DAST, SCA · identity and access governance · Kubernetes and supply-chain security

Governance & compliance ISO 27001 · PCI · GDPR and direct-listing readiness · international data regulation · patent and IP program management

Technical Go · TypeScript · Python · n8n automation · Lua

THOUGHT LEADERSHIP & RECOGNITION

Standards

- **OWASP Top 10 for Agentic Applications (2026)** — Co-Lead, *Rogue Agents* (ASI10) risk area.

Open source

- **Caterpillar** — open-source AI-skill and MCP-tool security scanner (TypeScript, MIT).
- **Interactive Framework Mapper** — web tool mapping AI and security controls across frameworks.

Speaking

- **RSAC 2025** — "Surprise: You Already Know How to Handle AI Risk."
- **GOSIM Paris 2026** — panel, "The Agentic AI Moment: From Models to Intelligent Systems."
- **HumanX 2026** — "What AI Red Teaming Looks Like Outside the Lab."
- **Agentic AI Summit, New York** — "The Burrito Bot That Wrote Code."
- **The Official Cybersecurity Summit 2025** — "Gen AI Application Security & Risk."
- **Responsible AI Summit (UK), 2026 · kidTECH 2026, Los Angeles**
- **"Outsiders & Outliers," Black Hat 2025 week** — master of ceremonies.

Podcast & press

- Host, *Curiouser & Curiouser*, an AI-security podcast. Guests have included Phil Venables, Steve Wilson, Dr. Janet Bastiman, Julie Tsai, Alison Cossette, and Tennisha Martin.
- Guest, **The Application Security Podcast** (S12E1) — "Building an AppSec Program from Scratch."
- Guest, **ShiftNow** (Ep. 6) — "The Rise of AI Red Teaming."
- Named in **Vanity Fair**, "Dario Amodei Has a Cold" (April 2026).

Writing

- *makeSecurityFun*, a personal security and AI blog. "AI Skills Security: A Practitioner's Guide to Emerging Threats" (2026), co-authored. Published vulnerability research on prompt injection in agentic browsers.

AFFILIATIONS & COMMUNITY

Coalition for Secure AI (CoSAI) member · **Advanced AI Society** leadership team · **Polaris Collective** member · **HarmonyWerks** (music non-profit) board member · **Urban Growers Network** advisor and former leadership team

EDUCATION

B.S., Computer & Information Sciences — SUNY College at Fredonia (2011–2015)