

Michael Puodziukas

Senior Engineer — Agentic AI · AI Security · Infrastructure / SRE · Mainframe Modernization

Remote, US-based (Mountain Time) · mpuodziukas@gmail.com · puodziukas.dev · github.com/mpuodziukas-labs

SUMMARY

Title-agnostic senior IC who builds agentic systems and red-teams them as the same person. I keep the digital plumbing running — infrastructure, SRE, cloud, AI/agentic systems, security, and legacy modernization — and I ship reproducible, clone-and-run proof of every claim. Self-directed and independent; methodology is published as runnable code, not slideware. Remote, skills-first, available Q3 2026.

SELECTED WORK

Agentic AI Systems — design, build, and adversarial validation

Self-directed · independent

- Built multi-agent orchestration with deterministic state and confidence gating; drove a measured production hallucination rate from a 27% baseline to under 1% via an 8-agent adversarial council with 3-lens voting (correctness, citation, reproducibility).
- Stood up an evaluation harness with a golden replay corpus so every routing and generation decision is re-executable and auditable; regressions are caught before release, not in production.
- Cut agent cold-start from 2–8s to ~200ms p99 through MLX kernel optimization and speculative decoding on Apple Silicon.

AI Security — LLM red-teaming and guardrails

Self-directed · independent

- Authored an open-source adversarial gate (llm-adversarial-gate): OWASP LLM Top 10 and MITRE ATLAS-mapped coverage, sub-2ms p99 per gate, zero violations across the internal replay corpus. Clone and run.
- Designed defense-in-depth for indirect prompt injection, agent hijacking, and MCP tool-schema validation; every defense paired with an adversarial replay case so coverage is demonstrable, not asserted.

Mainframe Modernization — COBOL to Rust, byte-exact

Self-directed · independent

- Published an open-source COBOL arithmetic forensic tool (cobol-pic-probe): detects COMP-3 / PIC-clause precision defects that silently corrupt financial arithmetic. Clone and run.
- Migrated a COBOL batch core to idiomatic Rust with byte-exact arithmetic verified by deterministic replay; 27x throughput gain, zero behavioral regressions, validated across 10M fuzz iterations with zero panics.

Infrastructure / SRE — sovereign inference and observability

Self-directed · independent

- Operated a heterogeneous inference cluster (Apple Silicon + NVIDIA + x86) over Thunderbolt 4 zero-copy transport; 158 TPS aggregate across a 12-model mesh at 389ms P50 — entirely on-prem, zero cloud API dependency.
- Cut token spend ~94% via a model-right-sizing + KV-cache + prefix-caching stack, taking a recurring inference bill to effectively zero while holding SLAs.
- Instrumented everything with Prometheus-style telemetry, deterministic replay, and CI-gated assertions so failures surface as alerts, not incidents.

VERIFIABLE PROOF (CLONE & RUN)

- github.com/mpuodziukas-labs/cobol-pic-probe — COBOL COMP-3 / PIC precision-defect forensics.
- github.com/mpuodziukas-labs/llm-adversarial-gate — OWASP LLM Top 10 / MITRE ATLAS-mapped guardrail gate.

TECHNICAL SKILLS

Languages: Python, Rust, TypeScript, Go, SQL, Bash, COBOL II

Agentic / LLM: Multi-agent orchestration, LangGraph, LangChain, deterministic state engines, evaluation harnesses, confidence gating, KV-cache & prefix caching, speculative decoding

AI Security: LLM red-teaming, OWASP LLM Top 10, MITRE ATLAS, prompt-injection & agent-hijacking defense, MCP tool-schema validation, adversarial replay corpora

Inference / ML: MLX, PyTorch, vLLM, quantization (AWQ/GPTQ, 4/8-bit), tensor parallelism, heterogeneous cluster serving

Infra / SRE: Docker, Kubernetes, Terraform, Kafka, Prometheus, event sourcing, deterministic replay, zero-trust mTLS, CI/CD

Cloud: AWS Bedrock, Azure OpenAI, SageMaker, serverless & model-agnostic inference layers

Mainframe / Legacy: COBOL II, COMP-3 forensics, z/OS, VSAM, JCL, DB2, COBOL to Rust byte-exact migration

Compliance: EU AI Act, ISO 42001, NIST AI RMF, SOC 2, FIPS 203/204/205 (PQC)

EDUCATION

University of South Florida — B.S.

Western Governors University — B.S. Computer Science (in progress)

AVAILABILITY

Remote · title-agnostic senior IC · target compensation \$140–175k · available Q3 2026 · domains: agentic AI, AI security, infrastructure / SRE, mainframe modernization.