

CODY KICKERTZ

AI Systems Architect · Platform & Infrastructure

Rust · agent runtimes · governed AI infrastructure · distributed systems

Cody.Kickertz@pm.me | ardent.tools | github.com/forkwright | Austin, TX

OPEN SOURCE

Selected projects — all Rust.

github.com/forkwright

Aletheia	Self-hosted multi-agent AI runtime. Embedded Datalog + HNSW memory engine over a knowledge graph; eleven-factor recall; per-tool kernel sandbox (Landlock + seccomp + netns); multi-provider LLM layer with circuit-breaking and loop detection. Single binary, no external services.
Kanon	Standards-and-dispatch control plane for a multi-repo fleet: a standards-as-code lint engine with rust-analyzer HIR type-aware checks, an MCP tool server, and a git forge + CI with native object parsing (no gix/libgit2) and a compare-and-swap work queue. Self-hosting — its own gate blocks its own landings.
Logismos	GPU inference stack for AMD RDNA3 in progress — hand-written HIP FFI and WMMA kernels, a custom GGUF loader, attention from first principles. Verified end-to-end on CPU; GPU cutover scoped and blocked on hardware access.
Hamma	Rust implementation of the Tailscale control protocol built directly from the Go source — Noise IK handshake, control-plane wire types, map streaming — with byte-level protocol tests. No production-grade Rust equivalent exists.
Thumos	From-scratch bare-metal OS for an armv7 smartphone SoC. Own MMU, log-structured filesystem, Ed25519 measured boot, Signal protocol implementation; treats the cellular baseband as untrusted, firewalled at the driver boundary. Cross-compiled and QEMU-tested.
Harmonia	Media platform: custom streaming protocol over QUIC with NTP-style clock sync for multi-room playback, lock-free audio ring buffer, biquad DSP pipeline, headless Raspberry Pi renderer via NixOS aarch64 cross-compile.

EXPERIENCE

Summus Global

2023 – Present

Data Scientist & AI Systems Architect (2025 – Present)

- Architected an agent-native internal analytics platform in Rust — a governed MCP tool server (137 tools) over the data warehouse — with a three-layer write guard (SQL-validator hook → SELECT-only database policy → hard-block on raw-CLI writes) that makes destructive agent operations against production PHI structurally impossible.
- Built a standards-as-code governance layer: 284 machine-enforced rules across 13 registries wired to commit-time gates, plus a self-improving loop (observed tool use → auto-drafted lessons → promoted rules) so a fresh agent performs like an experienced one because the knowledge lives in the infrastructure, not a person.
- Designed a medical-code taxonomy engine unifying 301K ICD-10, CPT, HCPCS, and SNOMED codes into a four-level, 7,771-node hierarchy projected in 12 seconds; custom multi-task LoRA fine-tune (MNRL + Matryoshka retrieval loss + four classification heads) with a verified LoRA-to-candle round-trip requiring zero application-code changes.
- Replaced a 40-cell Hex/Python ROI-scoring notebook with a production Rust engine at exact parity (\$4.18M validation client); found and quantified three compounding legacy scoring bugs inflating ~50% of reported savings, and rebased a flagship client's book onto a third-party-validated basis (independent-validator gap 5.0× → 2.1×).
- Made every client-facing deliverable machine-verifiable: a claims-manifest gate pairs each figure with a runnable, tolerance-bound SQL check before a report can ship — closing a class of hand-copied-number errors across 30+ ROI engagements.

Database, Data, & Business Intelligence Analyst (2023 – 2025)

- Designed a dimensional data model in AWS Redshift (star/snowflake patterns, CTEs, window functions) powering 8 dashboards and 40+ ROI analyses across every business function.
- Built a full-stack medical taxonomy system: React UI, FastAPI backend (50+ endpoints), 6-service architecture (PostgreSQL, Redis, Qdrant, Elasticsearch) serving 192K hierarchical records with semantic search.
- Engineered a physician deduplication system (fuzzy matching, NPI reconciliation) raising NPI coverage from 33% to 92% across 1,200+ profiles, unlocking three product features that depended on clean provider data.

Camp Lejeune, NC

Assistant Disbursing Officer (Dec 2022 – May 2023)

- Second-in-command of the Marine Corps' largest disbursing office: 157 Marines, 12 civilians, 7 finance functions.
- Reengineered core disbursing processes: 16% faster processing, 14% accuracy gain, roughly 375 hours returned annually; eliminated a 3,000+ claim backlog across \$45M+ in military pay and travel claims.

Disbursing Officer, 22d Marine Expeditionary Unit (Aug 2021 – Dec 2022)

- Ran fiscal operations for a 3,000-person MEU through a 7-month, 3-vessel deployment across 15 European and African nations.
- Managed a \$350,000 deployed cash budget with zero discrepancies; consolidated three accounting systems into one – adopted by East Coast MEU deployments that followed.

Travel Section Officer-in-Charge (Dec 2018 – Aug 2021)

- Led a 62-person travel-claims section supporting 60,000+ service members, civilians, and dependents across the eastern United States and Europe.
- Cut claim backlogs 30% through workload triage, process redesign, and production tracking.

Cyber Security Research Intern

- Built a predictive intrusion-detection system for the Air Force Research Laboratory (PCA, k-medoids clustering, SVM classification in R) — 98% accuracy on a 125K-observation network-traffic dataset.
- Selected to present the research at Pentagon Lab Day and the Georgia Institute of Technology.

TECHNICAL

Languages	Rust (primary), Python, SQL, TypeScript, R, Bash/Fish, Nix, Typst
Agent / AI	MCP tool servers, agent orchestration + dispatch (Claude API/SDK), standards-as-code governance, RAG, embeddings (candle, MedEmbed), fine-tuning (LoRA, MNRL, Matryoshka), clinical NLP (scispaCy)
Systems	Tokio async, fjall (LSM), QUIC/Quinn, Landlock/seccomp, WireGuard/Noise, no_std/embedded, HIP/GPU kernels
Data	AWS Redshift, Polars, SQLite, star/snowflake modeling, ETL pipelines, Hex (HIPAA)
Infrastructure	Podman, systemd, Tailscale, NixOS, GitHub Actions, restic, nftables

EDUCATION

The University of Texas at Austin – McCombs School of Business – MBA, May 2026

University of Houston – College of Technology – BS Computer Information Systems