

Abhinav (August) Srivastava

atg271@gmail.com | (857) 313-2851 | Santa Clara, CA
abhinavmir.xyz | github.com/abhinavmir | linkedin.com/in/abhinavmir

EDUCATION

Boston University — M.S., Computer Science (Cybersecurity) <i>Boston, MA</i>	2021–2023
VIT University — B.S., Computer Science <i>Vellore, India</i>	2017–2021

RESEARCH INTERESTS

Security, privacy, and formal methods in healthcare systems. Storage engines purpose-built for clinical monitoring; adversarial evaluation of voice agents that handle protected health information; privacy-preserving computation across institutional boundaries; cryptographic proofs over genomic data; formal verification of the data models underpinning health information exchange; and program analysis of legacy clinical software.

RESEARCH PROJECTS

Maize — an open parser and an empirical study of VistA <i>MUMPS program analysis</i>	2025–present
--	--------------

- Open tree-sitter grammar for MUMPS, the language underlying the U.S. Department of Veterans Affairs hospital system, producing a reusable and analyzable concrete syntax tree where no public machine-readable grammar previously existed.
- Parses 98.6% of VistA's 33,951 routines (3.6M lines) with no error nodes; expression semantics verified by differential testing against YottaDB.
- Call-graph and runtime-code-generation measurements over the OSEHRA VistA-M corpus.

EmberDB — FHIR-native time-series storage for critical care <i>Systems / clinical data infrastructure</i>	2025–present
---	--------------

- Time-series database treating FHIR Observations as first-class citizens for ICU monitoring, eliminating the impedance mismatch between clinical data models and storage.
- Chunked storage with FHIR-aware metric encoding; integrated multivariate pattern detection covering seasonal decomposition, change-point detection, and anomaly scoring.
- Benchmarked against InfluxDB, TimescaleDB, and QuestDB on MIMIC-III (330M observations, 38K patients).

Healthcare voice agent security <i>Adversarial evaluation of AI systems handling PHI</i>	2025–present
--	--------------

- Security evaluation of commercial and open-source voice AI platforms deployed in clinical settings.
- Testing framework built around a FHIR-compliant EHR system for realistic attack simulation, covering authentication bypass, prompt injection, PHI extraction, and synthetic voice attacks.
- Proposes a tiered security framework that reduces attack success rates while preserving clinical workflow.

Privacy-preserving healthcare cost prediction <i>Secure multiparty computation</i>	2025
--	------

- Custom SMPC protocol enabling joint model training across hospitals, insurers, and public health agencies without pooling raw data.

- Additive secret sharing with Beaver-triple-based multiplication under an honest-majority assumption, plus a commitment-based verification layer for detecting malicious participants.
- Optimised for gradient-based workloads, requiring fewer communication rounds than general-purpose frameworks.

zkDNA — zero-knowledge proofs for negative STR DNA match verification

2025

Applied cryptography

- Interactive zero-knowledge protocol proving DNA non-match without revealing genetic data, operating over committed STR allele values with locus-level inequality proofs.
- Formal security analysis establishing completeness, soundness, and the zero-knowledge property.
- Targeted at forensic investigation, where a suspect requires privacy-preserving exclusion.

Formal verification of FHIR resources

2025

Lightweight formal methods for EHR data models

- Framework applying formal verification to the HL7 FHIR Patient resource, formalising structure, cardinality constraints, choice-type invariants, and cross-field dependencies in Alloy.
- Resource-centric verification boundary modelling external references as opaque types to keep analysis tractable.
- Bounded model checking to exhaustively verify structural integrity, semantic consistency, and safety invariants; compositionally extensible to other resources and cross-resource workflows.

Hospital Gym — reproducible simulation for benchmarking clinical decision policies

2025

github.com/AbhinavMir/hospital-gym

- Emergency-department simulator modelling intake, triage, resource scarcity, and inventory under a randomised but clinically plausible event schedule.
- Every event logged as JSONL; a given scenario, seed, and action sequence replays exactly, so two policies can be compared under identical conditions.
- Exposed over MCP so language-model agents can be benchmarked as decision policies against a fixed environment.

OPEN-SOURCE CONTRIBUTIONS

18 fixes merged or maintainer-approved upstream, each reproduced, root-caused, and landed with a regression test in the project's own suite.

- **Compilers and optimisers:** Go `cmd/compile` — lowered the floating-point min/max idiom to a single instruction (MINSD/MAXSD/FCSEL) in place of a branch; reviewed and approved by Keith Randall and Russ Cox. DuckDB — optimiser eliminated a top- n window it should have preserved, erroring a valid QUALIFY clause. Cython — crash in constant folding. Ruff — FURB131 emitted a fix that raised a runtime `KeyError`.
- **Parsers and input hardening:** jq (raw NUL bytes truncated tokens), findutils and coreutils (panics on non-UTF-8 and missing input), tmc/x12 (panic on a nil ISA header).
- **Systems and memory safety:** Apache NuttX (reference leak and double free in the Bluetooth HCI ACL path), Zephyr RTOS (POSIX `putmsg()`).
- **Additional:** Pulumi, Reth, Zed, StackExchange.Redis, Kombu, htsjdk.

PROFESSIONAL EXPERIENCE

Senior Founding Engineer, Ambra

Dec 2025 – present

Venture-backed revenue-cycle automation for emergency medical services; San Francisco / remote

- Own all engineering — architecture, delivery, infrastructure, and roadmap — and lead two engineers and one intern.
- Architected and operate a claims pipeline processing \$1M+ per month across 100+ agencies: ingestion, automated denial detection and classification, and resolution routing over X12 276/277, with ingestion adapters for ten or more payer, clearinghouse, and EMS systems, each implementing the EDI standard differently.
- Led the organisation through HIPAA and SOC 2 from zero: control design, evidence collection, and audit readiness.

Founding Engineer, Simbie AI

May 2025 – Dec 2025

Conversational-AI platform for medical practices; remote

- Built core platform infrastructure; grew the engineering team from one to three while continuing to ship.
- Delivered insurance verification, FHIR integration, robotic process automation, outbound calling, and a whitelabel interface, adopted by four practices in three months.
- Established systematic model benchmarking, prompt iteration, and observability to bring live patient-facing conversation quality to production standard.

Software Engineer, Commure (Scribe)

Jun 2024 – May 2025

Ambient clinical documentation used by thousands of clinicians; Mountain View, CA

- Integrated 15+ electronic health record platforms (athenahealth, eClinicalWorks, NextGen, DrChrono, ModMed, Elation, and others), each with distinct identity models, data schemas, and write-back semantics.
- Owned the clinician-facing mobile application and browser extension; maintained production infrastructure and databases for the flagship product.

TECHNICAL SKILLS

Languages: Python, TypeScript, Go, Rust, C, MUMPS (analysis)

Systems and infrastructure: PostgreSQL, Redis, RabbitMQ, Docker, Kubernetes, Terraform, GCP, AWS

Methods: formal specification (Alloy), grammar and parser construction (tree-sitter), LLM evaluation and benchmarking design, reproducible simulation

Domain standards: HL7 FHIR, X12 EDI (276/277/837/835), NEMESIS; HIPAA and SOC 2 programs