

ABDULLAH AZMI

ORCID: 0009-0004-8172-6394 | github.com/aazminu

Lead Researcher, Computational Genomics & Agentic Systems (CGAS) Lab | Northwestern University (CS & NICO)
Evanston, IL & New York, NY | m4u5t4@u.northwestern.edu | erlokix@gmail.com | <https://abdullahazmi.com> | <https://abdullahazmi.com/lab/>

EDUCATION & ACADEMIC AFFILIATION

Northwestern University -- McCormick School of Engineering & NICO
Computer Science & Complex Systems | Computational Genomics & Agentic Systems (CGAS) Lab (Mudd Hall 3510)

- Core Focus: Autonomous Scientific Knowledge Graphs, Multi-Agent Verification, Regulatory Genomics, Probabilistic State-Space Inference, and Distributed Full-Stack Systems.

Evanston, IL
2023 -- Present

RESEARCH & ENGINEERING EXPERIENCE

Computational Genomics & Agentic Systems (CGAS) Lab -- Northwestern CS
Lead Computational Systems Researcher & Principal Software Architect

- Architected PaperPulse ([aazminu/paper-pulse](https://github.com/aazminu/paper-pulse)), an autonomous nightly literature ingestion and knowledge-graph engine indexing ArXiv, PubMed, Semantic Scholar, and OpenAlex into PostgreSQL (pgvector) and Neo4j with deterministic multi-hop claim provenance.
- Developed ChromAlphaGenome ([aazminu/chrom-alpha-genome](https://github.com/aazminu/chrom-alpha-genome)), integrating AlphaGenome cis-regulatory contact tracks, ChromBPNet/DeepLIFT chromatin accessibility attribution, and MaxEntScan donor/acceptor splice-junction Delta-PSI modeling across ClinVar and gnomAD (0.941 AUROC).
- Built FoldEnergetics-xTB ([aazminu/fold-energetics-xtb](https://github.com/aazminu/fold-energetics-xtb)), coupling AlphaFold 3 PAE spectral domain partitioning and Foldseek 3Di structural homology with semiempirical tight-binding quantum mechanics (GFN2-xTB) and ChEMBL bioactivity triage.

2025 -- Present
Evanston, IL

Northwestern Institute on Complex Systems (NICO)
Computational Biology & Probabilistic Inference Researcher

- Engineered BayesKalmanTrace ([aazminu/bayes-kalman-trace](https://github.com/aazminu/bayes-kalman-trace)), a pure NumPy/SciPy recursive Bayesian state-space filter and Rauch-Tung-Striebel (RTS) backward smoother across 500-sample noisy velocity trajectories with analytic missing-data marginalization.
- Calibrated Laplace-smoothed rolling-volatility empirical likelihoods over a 250,000-measurement corpus and fused them via log-linear (Product-of-Experts) pooling, lifting held-out regime classification accuracy from 74.2% to 93.8% (0.041 ECE).
- Automated parallel Python + SQL reporting and multi-omics validation workflows, cutting recurring exhibit preparation latency from over 1 hour to under 8 minutes (87% reduction) while catching upstream schema and reconciliation anomalies.

2024 -- 2025
Evanston, IL

Enterprise Operations & Open-Source Systems Engineering
Full-Stack & Workflow Automation Lead (Airtight Container, LeadAI, Odyssey, Mermaid, FarmData2)

- Built and deployed Airtight Container ([aazminu/airtight-container](https://github.com/aazminu/airtight-container)), a production yard inventory, quoting, and idempotent month-end storage billing platform (React, TypeScript, Node.js/Express, PostgreSQL, AWS EC2; 40K+ LOC, 25 migrations, 275+ automated tests).
- Engineered LeadAI ([aazminu/LeadAI](https://github.com/aazminu/LeadAI)) async batch entity enrichment queues and an event-driven Python/FastAPI/Slack ChatOps bot integrating GitHub Actions and issue trackers, reducing manual change-request triage latency by 95%.
- Delivered reviewed open-source fixes and regression suites to Khoury Odyssey ([aazminu/odyssey](https://github.com/aazminu/odyssey), recursive MDX preview rendering & editor state sync), Mermaid ([aazminu/mermaid](https://github.com/aazminu/mermaid), bar-chart data labels), and FarmData2 ([aazminu/FarmData2](https://github.com/aazminu/FarmData2), transactional APIs & Cypress E2E tests).

2023 -- 2025
New York, NY & Remote

SELECTED OPEN-SOURCE SYSTEMS & SIDEPROJECTS (16 PUBLIC SYSTEMS ON GITHUB.COM/AAZMINU)

- PaperPulse (github.com/aazminu/paper-pulse): FastAPI + PostgreSQL (pgvector) + Neo4j + Celery/Redis + Next.js 14 autonomous literature synthesis, multi-hop citation graph verification, and daily audio briefings.
- ChromAlphaGenome & FoldEnergetics-xTB (github.com/aazminu): Base-resolution chromatin/splice variant prioritization (0.941 AUROC) and AlphaFold 3 PAE + Foldseek 3Di + GFN2-xTB quantum pocket energetics triage with SKILL.md agent protocols.
- BayesKalmanTrace (github.com/aazminu/bayes-kalman-trace): Constant-acceleration Kalman filter + RTS smoother + log-linear likelihood pooling over 250K trajectory measurements (93.8% accuracy, exact dropout marginalization).
- Airtight Container & LeadAI (github.com/aazminu): Production React/Node/PostgreSQL container yard billing platform (275+ tests, idempotent generateShMonthEnd invoicing) and bounded async multi-source entity enrichment queue.
- Systems, Data & ML Architecture Suite: ScrubData (100K+ row Spark/PostgreSQL cleaning & outlier reconciliation), Cycle-Accurate RISC-V RV32I 5-Stage Pipelined Simulator (EX/MEM & MEM/WB forwarding), Dividend Capture Spark/Kafka Streaming (165K+ records), Low-Rank SVD World-Model Compression (LeWM MPC), Adaptive DistilBERT Moderation (<150ms p95), and WebSocket Group Recommender.

TECHNICAL STACK

Languages & Systems: Python (3.11+), TypeScript, JavaScript, SQL (PostgreSQL), C, C++, Rust, RISC-V ISA, Bash, PowerShell 7

Backend, Data & Cloud: FastAPI, Node.js, Express, React 18, Next.js 14, PostgreSQL (pgvector), Neo4j, Redis, Celery, Spark, Kafka, Docker, AWS

Scientific AI & QA: PyTorch, NumPy, SciPy, AlphaFold 3 / AlphaGenome, Foldseek 3Di, GFN2-xTB, Kalman/RTS, Pytest, Jest, Vitest, Cypress