

# FAROUQ OGUNTOYE

farouqoguntoye05@gmail.com · [github.com/Herr-Professor](https://github.com/Herr-Professor)

## SUMMARY

---

ML infrastructure and evaluation engineer with ~4 years of experience building reproducible evaluation harnesses, verifiable data pipelines, and performance-critical systems. Built deterministic RL evaluation environments with sandboxed verifiers and multi-model benchmarking, verifiable data pipelines with leakage audits and drift gating, and custom C++ systems tooling (allocators, compression, SIMD) benchmarked against production baselines. Looking to bring systems-level rigour to ML infrastructure, evaluation, and research engineering roles at AI-focused companies.

## SKILLS

---

**Languages:** C++20, Python, JavaScript/TypeScript, SQL

**ML Infrastructure:** RL environments (Gymnasium), deterministic verifiers, sandboxed execution, pass@k evaluation with bootstrap CIs, offline RL (FQI), experiment tracking, multi-model evaluation registries, reproducible reporting

**Data & Pipelines:** Parquet/Arrow, versioned dataset artifacts, data contracts, leakage audits, drift gating, provenance tracking, PostgreSQL, SQLite, async pipelines

**Systems:** Custom allocators (pool/segregated/fixed-size), compression algorithms (Huffman/LZ77/DEFLATE), SIMD (AVX2), thread safety, memory management, profiling, observability

**Backend & Infra:** Flask, FastAPI, REST APIs, Docker, Linux, Git, CI/CD (GitHub Actions), CMake

## SELECTED PROJECTS

---

**TinyCodeTest** (Python, Vercel)

[github.com/Herr-Professor/TinyCode](https://github.com/Herr-Professor/TinyCode)

- ▶ Deterministic verifiers-style RL environment for code generation with sandboxed subprocess execution, per-attempt timeouts, and memory caps.
- ▶ Built full evaluation stack: difficulty-bucketed datasets (easy/medium/hard) with adversarial splits, pass@k scoring with 95% bootstrap confidence intervals, and leaderboard export (JSON/Markdown/HTML).
- ▶ Implemented multi-model registry supporting OpenAI, Anthropic, and Gemini APIs with configurable prompt strategies (direct, plan-then-code).
- ▶ Shipped web UI on Vercel for interactive dataset upload, model selection, API key management, and live eval runs with verification trace output.

**Market Data Pipeline** (Python, Parquet, YAML)

[github.com/Herr-Professor/market\\_data\\_pipeline](https://github.com/Herr-Professor/market_data_pipeline)

- ▶ Verifiable data-quality framework producing versioned Parquet datasets with manifests, checksums, and run provenance (git commit, config hash, dependency versions).
- ▶ Implemented deterministic leakage audits covering forward-looking joins, split leakage, survivorship bias, and label misalignment, with automated drift gating.
- ▶ Built baseline evaluation harness with configurable cost/latency assumptions and sensitivity analysis quantifying lookahead signal inflation.

**Compression Bench** (C++20, CMake, SIMD)

[github.com/Herr-Professor/compression-bench](https://github.com/Herr-Professor/compression-bench)

- ▶ Lossless compression benchmark suite implementing 5 algorithms from scratch (Huffman, LZ77, DEFLATE, RLE, LZW) with streaming API and self-describing container format (CRC32 verification).
- ▶ Multi-threaded DEFLATE with AVX2-accelerated byte matching; benchmark harness with warmup iterations, median-of-N measurements, and terminal/HTML/CSV reporting across 14-file corpus.

**Custom Memory Allocator** (C++, pthreads)

[github.com/Herr-Professor/memory\\_allocator](https://github.com/Herr-Professor/memory_allocator)

- ▶ Thread-safe allocator with 4 strategies (best-fit, fixed-size, pool, segregated) and AVX2 alignment, benchmarked against system malloc, jemalloc, and tcmalloc on RL/inference workloads.
- ▶ Reproducible benchmark harness measuring throughput, scaling, and overhead ratio with CSV outputs, automated plots, and pinned toolchain for reproducibility.

**F1 Strategy RL Environment** (Python, Prime Intellect)

[github.com/Herr-Professor/f1\\_rl](https://github.com/Herr-Professor/f1_rl)

- ▶ RL environment for F1 race strategy with OpenF1-derived scenarios, deterministic verifiers, tool-use rubrics, and deep-reasoning mode.
- ▶ Full evaluation protocol with baselines (95% CIs), ablations, and stress tests. Trained on Prime Intellect Lab hosted RL beta.

**Heath** (Python/Flask, Parquet)

[github.com/Herr-Professor/heath](https://github.com/Herr-Professor/heath)

- ▶ F1 mini-season telemetry simulator with race reconstruction pipeline (ingest, track reference, SC/VSC inference, validation).
- ▶ Offline RL baseline comparison (FQI vs. behaviour cloning vs. heuristic) with Flask-based replay viewer.

## EXPERIENCE

---

**Tonfans** — Remote

*Sep 2024 – Nov 2025*

*Lead Developer*

- ▶ Owned end-to-end delivery of a production Telegram bot integrated with TON blockchain for community payments and access control.
- ▶ Implemented purchase and verification flows for NFTs/Jettons with automated on-chain ownership gating.
- ▶ Built secure Python backend workflows and Telegram Bot API integrations handling high-volume user interactions.

**Lifestores Pharmacy** — Lagos, NG

*Sep 2022 – Aug 2024*

*Systems Engineer*

- ▶ Built a real-time C++ order-processing system sustaining 500+ orders/hour with sub-second latency.
- ▶ Developed a thread-safe C++ memory management library with pooling and free-lists; reduced memory usage and improved throughput by ~25%.
- ▶ Shipped structured logging and telemetry for critical services, reducing time-to-detect/resolve incidents by ~40%.

**Ultra Cloud Technologies** — Remote

*Aug 2021 – Jun 2022*

*Backend Developer*

- ▶ Developed RESTful APIs in Python/Flask for authentication and customer-management services (1,000+ daily requests).
- ▶ Optimized PostgreSQL queries and containerized services with Docker, improving performance and deployment reliability.

## EDUCATION

---

**Obafemi Awolowo University** — B.Sc. Applied Geophysics

*Jul 2025*

**Federal Polytechnic, Ado Ekiti** — National Diploma, Mechanical Engineering

*Jun 2019*