Tri Ho Dinh

Software Engineer (Systems) | Ho Chi Minh City, Vietnam
(+84) 901 364 158 | hodinhtri3010@gmail.com | hodinhtri.me | linkedin.com/in/trihodinh | github.com/HDT-0103

Education

University of Science, VNU-HCM — Ho Chi Minh City, Vietnam
B.Sc. in Information Technology (Information Security) — ***GPA: 3.84 / 4.0*** — *Sep. 2024 – Expected June 2028*
- **Availability: third-year student**; **part-time, 3–4 days/week** (from now); **full-time from Sep. 2027**.
- **Coursework:** Operating Systems, Data Structures & Algorithms, Computer Networks, Distributed Systems, Database Systems.

Projects

xv6 RISC-V Filesystem Extension & Crash Consistency | *C, RISC-V, GDB* — Feb. 2026 – Mar. 2026
- Implemented double-indirect block allocation in MIT xv6, growing max file size from **268 KiB to 64.3 MiB (∼245×)**.
- Handled recursive truncation of up to **65,803 blocks** with leak-free deallocation; wrapped metadata updates in write-ahead logging (`log_write`) and verified crash consistency with a full write/readback stress test.

SmartATS – AI Applicant Tracking Platform | *FastAPI, Next.js, PostgreSQL, pgvector, Azure* — July 2026 – Aug. 2026
- Architected a full-stack ATS of **10 modules** and **54 REST APIs** with asynchronous, high-throughput resume ingestion and parsing.
- Built semantic search on **768-d multilingual embeddings** with **pgvector HNSW** indexes while keeping exact boolean hard filters.
- Added observability for **p95/p99 latency**, token throughput and cost, plus RBAC/ABAC data-access boundaries.

Experience

Meddies — Ho Chi Minh City, Vietnam
AI Engineer (Part-time) – Multilingual PII Detection & Data Platform — *May 2026 – Present*
- Built the **data generation pipeline** producing **1M** synthetic clinical documents in 17 languages, orchestrating a pool of **9 LLM providers** with attribute-conditioned sampling.
- Implemented **13 deterministic validation gates** (regex, character-level offset parser, rule checks, **SHA-256 de-duplication**) with a conservative auto-repair loop that re-validates every fixed record.
- Developing **meddies-faker**, a Python library extending Faker with pluggable **country-aware providers** that sample names and personal attributes from real-world frequency distributions for **20 locales**; internal today, planned open-source.
- Pipeline output powers a PII model reaching **0.827 mean micro-F1** on 15 external benchmarks vs. **0.658** for the strongest baseline (co-first author, arXiv:2609.12544).

IaC Lab, Faculty of IT, HCMUS — Ho Chi Minh City, Vietnam
Undergraduate Researcher – Infrastructure-as-Code (IaC) — *Feb. 2026 – Present*
- Building collection and classification tooling for a large-scale, open **IaC dataset** (Terraform, Kubernetes) aggregated from public sources for future research.
- Benchmarked static-analysis engines across **29 runners** and **1,346 policy rules**; exported results to **SARIF/CycloneDX/JSON** for automated CI/CD gates.

Publications

L. U. Le*, **T. Ho Dinh***, C. Hoang, H. H. Ha. *"Meddies-PII: A Multilingual Framework for PII Extraction in Clinical De-identification."* arXiv:2609.12544, 2026. [arXiv] (*co-first author)

Technical Skills

Languages: C, C++, Python, Bash, SQL
Systems: Linux/POSIX, File Systems & Journaling, Concurrency & Multi-threading, Memory Management, RISC-V (xv6)
Tooling: GDB, CMake, Git, Docker, Unit/Integration Testing, Static Analysis
Backend: FastAPI, REST APIs, PostgreSQL, Async Task Queues
English: C1 (Oxford Placement Test)

Honors & Awards

RMIT Hackathon 2026 – **6th Globally**, **4th in Vietnam**, 3rd at Saigon South Campus (team of 4) — 2026
Academic Merit Scholarship – Faculty of IT, HCMUS (two consecutive years) — Jan. 2025 & Jan. 2026