Ahmed Alwahedi

ahmed.t.alwahedi@gmail.com | +1 857 335 7416 | ahmedalwahedi.com | github.com/ATW117 | linkedin.com/in/ahmed-alwahedi

Education

Boston University, *BA in Mathematics and Computer Science* — Boston, MA — May 2027

- GPA: 3.74/4.00 | Dean's List (x4) | ISNR 2026 CTF 1st place winner
- **Hackathons:** MIT iQuHACK (2024, 2025); HackHarvard (2024, 2025); BostonHacks; Yale Quantum Computing Hackathon (2024).

Skills and Technologies

- **Languages:** Python, Java, SQL, Rust, C, JavaScript, Go, R
- **AI/ML:** PyTorch, Scikit-learn, NumPy, Hugging Face, LoRA Fine-tuning, RAG
- **Backend/Infra:** FastAPI, MCP, PostgreSQL, Azure, Git

Professional Experience

7X, *Forward Deployed Engineering Intern* — Dubai, UAE — Jul 2026 – Aug 2026

- Architected a config-driven AI employee platform with Python, FastAPI, and CrewAI; four department agents combined scoped MCP tools, knowledge bases, scheduled duties, role-gated chat, and free-text planning.
- Built LLM execution governance with RBAC, per-agent tool scopes, tool/token/time budgets, human approval gates, audit/replay, and scoped one-hop delegation.
- Integrated MCP/Oracle tools and PostgreSQL telemetry and deployed via GitHub Actions to Azure App Service while iterating directly with business departments.

MBZUAI (UGRIP), *Machine Learning Research Intern* — Abu Dhabi, UAE — Jun 2026

- Built two **ground-truthed multi-agent LLM testbeds** (linear and star pipelines) with single-fault injection, so the responsible agent is known *by construction*.
- Benchmarked **black-box** attribution (leave-one-out, counterfactual repair, LLM-as-judge) against **white-box** signals (gradient×input saliency, contrastive logit ablation, TracIn-style influence); judges localize the failure *site* but not the *cause*, unchanged by an 8× judge scale-up.
- Showed a **blind, trace-only structural rule** recovers first cause on 6/6 fault episodes where a 14B judge scores 1/6, and used ground truth to catch a **label leak** that had inflated an earlier result to 1.000.

ADIA Lab, *Machine Learning Intern* — Abu Dhabi, UAE — Jun 2025 – Aug 2025

- Benchmarked Logistic Regression, KNN, Random Forest, and XGBoost with Optuna; selected an F1-optimized Random Forest for diabetes screening (F1 0.71, recall 0.80, accuracy 0.77).
- Ran an age-subgroup fairness audit with quantile-balanced bins, replacing fixed bins whose 25-person group produced a misleading 1.00 recall, and used SHAP to trace weak young-group recall (0.55 vs. 0.84–0.89) to an age–pregnancy confound.

Personal Projects

Kanto & Indigo, Self-Play RL for Pokémon *Rust · PyTorch · PPO · Transformers*

- Built **Kanto**, a Rust Gen 1 battle simulator that matches Pokémon Showdown byte for byte (**0 divergences** over 81K battles / 6.5M decisions of differential testing). Runs ∼**900× faster** for RL at 1M environment steps/s on one CPU core.
- Trained **Indigo** transformer agents (3–12M parameters) purely by **self-play PPO with a league**; the 12M model scores **60.2%** over 2,001 held-out games against **TaurosV0**, a 4× larger offline-RL agent and one of the strongest open Gen 1 models, and **81.3%** against the search bot Foul Play.
- Added determinized rollout search with paired evaluation (+2.8 points vs. TaurosV0, 95% CI)

Saturnia LoRA *Fine-tuning · Synthetic Data· Python · PyTorch*

- Ran a controlled **LoRA ablation** on FLUX.2 Klein 4B: six adapters, one per mix of reference vs. synthetic data, with rank, LR, steps and seeds fixed; built a compiler with hash-verified manifests and deterministic holdouts.
- Scored 252 seed-matched images **blind with a local VLM judge** against step-zero baselines: style rose from 1.49 to 2.31–2.92 of 5 (95% bootstrap CIs exclude zero), but five of six mixes were indistinguishable.

Family Historian *Multilingual RAG · Python, vLLM, FastAPI, BGE-M3, Vision LLMs*

- Transcribed **9,350 handwritten and printed Arabic/English pages** with a **7-model local OCR ensemble** (vLLM) reconciled by **word-level voting** calibrated on a hand-labeled gold set; an **LLM judge** resolved 314K disputed words (handwritten-Arabic accuracy **50%→76%**).
- Built **hybrid retrieval** (BGE-M3 + BM25 + reranking), raising known-item recall@10 from **0.54 to 0.93**, with streamed answers grounded strictly in cited sources.
- Deployed a bilingual **FastAPI** platform (document viewer, people directory, timeline) via **Cloudflare Tunnel** with device-capped access control.