Abror Shopulatov

murodbeck@proton.me murodbecks.github.io github.com/murodbecks Google Scholar

AI Engineering undergraduate at MBZUAI researching AI systems across LLMs and computer architecture: LLM-based assembly transpilation and agentic model porting to ML accelerators (Tenstorrent TT-Metal). Three years of production NLP/ML engineering; interested in AI hardware, compilers, and ML infrastructure.

Research Experience

Undergraduate Researcher — **Abu Dhabi, UAE**
MBZUAI, Computer Architecture Lab (Prof. Abdulrahman Mahmoud) — *10/2025 - present*

- **PortForge-TT (project lead):** Leading an agentic workflow that autonomously ports pretrained models to Tenstorrent Blackhole accelerators: coding agents with scoped tools paired with an independent trusted evaluator on real hardware. All four campaigns (NLLB-200, Chronos-2, ESM-2 650M, Parakeet TDT 0.6B) pass FP32 verification; reduced-precision ports (BF16, BFP8_b) are admitted through NRMSE and task-metric gates, followed by profile-guided optimization passes (2-18× speedups over baselines).
- **Agentic assembly transpilation:** Extended CPU assembly transpiler to O2 optimization across x86-64/AArch64/RISC-V; built the GCP pipeline serving fine-tuned Qwen models (vLLM) with benchmark and eval harnesses (cross-compilation, QEMU). Co-authored an ASPLOS submission; led the pivot to agentic transpilation (Harbor repair harness), showing frontier coding agents reach near-perfect repair rates.
- Help manage lab compute infrastructure: GCP GPU instances and machine access for lab members; co-authored the cluster user guide.

Industry Experience

Machine Learning Engineer — **Tashkent, Uzbekistan**
tilmoch.ai (Tahrirchi) — *02/2023 - present*

- Built Uzbekistan's largest open text corpora - UzBooks (40K OCRed books) and UzCrawl (1.2M web/Telegram sources, ∼35GB total) - via web scraping, Telegram mining, and book OCR; drove their open-source release within the company.
- Adapted Grammarly's GECToR for Uzbek grammar correction: pretrained BERT-family encoders from scratch (tahrirchi-bert-base/small) and fine-tuned them with synthetic and annotated data for the production Tahrirchi editor.
- Built production machine translation: collected and aligned multilingual parallel corpora and scaled them with LLM-based synthetic translation; contributed to every version of the production model - now serving 16 languages - plus an LLM refinement layer (Sayqalchi) that significantly grew adoption.
- Helped manage ML infrastructure and internal tooling: large-scale data and LLM batch pipelines, and compute/API resource administration for model training and serving.

Education

Bachelor of Science in Artificial Intelligence — **Abu Dhabi, UAE**
MBZUAI — *2025-2029*

Relevant coursework: Computers and the Web, Machine Learning, Deep Learning, Algorithms and Data Structures, Data Engineering and Science.

Publications

- **Shopulatov, A.**, Ma, D., & Mahmoud, A. (2026). *PortForge-TT: Enabling Agentic Cross-Platform Model Adaptation using CHIA*. CHIA Hackathon short paper. github.com/murodbecks/portforge-tt
- Mamasaidov, M., Aral, A., **Shopulatov, A.**, & Inomjonov, M. (2025). *Filling the Gap for Uzbek: Creating Translation Resources for Southern Uzbek*. WMT25. aclanthology.org/2025.wmt-1.83
- Mamasaidov, M., & **Shopulatov, A.** (2024). *Open Language Data Initiative: Advancing Low-Resource Machine Translation for Karakalpak*. WMT24. aclanthology.org/2024.wmt-1.48

Skills

- **Programming:** Python (primary), Bash; working knowledge of x86-64/AArch64/RISC-V assembly
- **ML & Agents:** PyTorch, vLLM, HuggingFace ecosystem, fine-tuning, LLM evals; agent harnesses (Harbor, MCP, OpenCode/Claude CLIs)
- **Systems & Accelerators:** Linux, Docker, SLURM, GCP, Clang/QEMU cross-compilation, Tenstorrent TT-Metal/TTNN
- **Languages:** Uzbek (native), English (advanced)