Sardor Nodirov

207-616-7032 | sinodi26@colby.edu | linkedin.com/in/sardornodirov | nodirov.com

EDUCATION

COLBY COLLEGE — Waterville, ME
Bachelor of Arts in Computer Science, AI concentration — Expected May 2026

- Presidential Scholar; Program GPA: 3.9/4.0
- Relevant Coursework: Neural Networks, Data Analysis & Visual., Comp. Archit., DSA, AI Ethics
- President of Colby College AI Society

THE UNIVERSITY OF EDINBURGH — Edinburgh, Scotland
Study Abroad Program at The Schools of Electronics & Electrical Engineering and Informatics — Jan 2025 - May 2025
Relevant Coursework: Natural Language Processing, Quantum Computing, Embedded Systems

TEMPLE UNIVERSITY JAPAN — Tokyo, Japan
Study Abroad Program with Coursework in Computer Imaging — Sep 2024 - Dec 2024

RESEARCH PROJECTS

Real-Time Human Motion Generation For VR Characters |*Honor's Thesis* | **Demo:** lapwing.live — May 2025 – Present

- Implemented a kinematic retargeting pipeline to map BVH motion data to VRM humanoids in real-time
- Solved root motion drift and rotation artifacts using inertial blending heuristics, ensuring anatomically correct 60 FPS playback without foot-sliding
- Engineered a low-latency, non-blocking pipeline (asyncio/WebSockets) to synchronize concurrent LLM token streaming with TTS audio and BVH motion synthesis.

WORK EXPERIENCE

EMBER AI (Early-Stage Startup) — San Francisco, CA (remote)
Founding Full-Stack AI Engineer (Contract) — May 2024 – Nov 2024

- Engineered a **multi-modal** AI pipeline, using **state-of-the-art Vision Transformers** and LLMs to generate interactive, narrated memory experiences from user-uploaded raw images
- Architected a fault-tolerant, asynchronous generation backend utilizing **Redis** and **BullMQ** to handle high-concurrency generation queues
- Built a real-time playback engine with UI, achieving near-instant **<50ms** visual-audio sync latency across distributed clients.
- Optimized global asset delivery and storage costs by implementing Cloudflare **R2** with edge caching, reducing projected egress costs by 50% compared to standard S3 architecture.

DAVIS INSTITUTE FOR AI — Waterville, ME
AI Research Assistant – Conversational Architectures — Sep 2023 – May 2024

- Architected **a real-time, full-duplex** voice agent achieving **sub-second** end-to-end latency by orchestrating an asynchronous FastAPI WebSocket pipeline with TTS and Azure Speech STT streams.
- Designed a **hybrid inference engine** optimizing for speed vs. reasoning; implemented **router logic** to dispatch complex queries to reasoning models and latency-sensitive chat to local OSS models, achieving 300+ tokens/sec throughput.
- Built a production-grade Retrieval-Augmented Generation (RAG) system; developed a custom Scrapy spider to vectorize **800+** institutional pages into **Qdrant**, enabling high-fidelity semantic search via Cohere embeddings.
- Engineered a custom audio buffering and barge-in protocol handling raw **mulaw/8000Hz** streams, enabling users to naturally interrupt the AI while maintaining conversation context and state consistency.

SKILLS

Languages: Python (Advanced: asyncio, multiprocessing), Java, C, Rust, SQL, TypeScript, JavaScript,
AI & Machine Learning: PyTorch, Tensorflow, MoMask, MDM, Motion generation, Vector Stores & Embeddings, LLM Orchestration, Vision Transformers, Azure Speech Synthesis, Voice Cloning, Linear Algebra, Quaternions
Real-time Systems: WebSockets, WebRTC, Redis (BullMQ/Caching)
3D Graphics & Standards: Three.js, WebGL, BVH, VRM, Bone Retargeting, Root Motion
Full-Stack Engineering: Next.js, React 18, Node.js, FastAPI, Tailwind CSS, shadcn UI
Infrastructure & Data: Cloudflare R2/Workers, Firebase (Firestore), MongoDB, Docker, Git, CI/CD.
Languages Spoken: English, Russian, Uzbek, Japanese