Samuel Baldwin

Boston, MA | baldwin.sam@northeastern.edu | (860) 817-3628
github.com/samuelbaldwin05 | linkedin.com/in/baldwinsamuel | Available Summer & Fall 2026

EDUCATION

Northeastern University, Boston, MA | Khoury College of Computer Sciences — September 2023 – April 2027
Candidate for B.S. in Data Science — **GPA:** 3.94 / 4.00
Relevant Coursework: Programming with Data, Data Science, Database Design, Linear Algebra, Large Scale Information Storage, Algorithms, Data Structures, Machine Learning, Data Mining, Econometrics, Macro Theory, Micro, Calculus II, Artificial Intelligence, Natural Language Processing, Reinforcement Learning, Probability, Statistics, Software Engineering
Leadership & Awards – Connecticut Seal of Biliteracy in Spanish, Dean's List all Semesters, Eagle Scout

TECHNICAL SKILLS

Languages: Python, Typescript, JavaScript, C++, SQL, R
Tools: Git, GitHub, Tableau, Excel, Figma, pytest, Docker, Linux, Vite, CI/CD, Agile/Scrum
Frameworks & Libraries: TensorFlow, PyTorch, Scikit-learn, Pandas, Numpy, Flask, FastAPI, Django, React, Tailwind, Astro
Cloud & Database: AWS, Terraform, Supabase, Redis, MongoDB, InfluxDB, Postgres, MySQL, SQLite, Chroma

PROFESSIONAL EXPERIENCE

NExT Consulting | Software Engineering Co-op — **July 2025 – December 2025**

Via Separations | Data Engineering Consultant

- Built ETL data pipeline unifying inline sensor data from 1000+ live sensors and offline lab measurements, enabling predictive maintenance for 60+ engineers leading to reduced equipment downtime and performance degradation
- Deployed AWS infrastructure via Terraform managing RDS, InfluxDB, DynamoDB, ECR repos, and ECS tasks
- Improved interactability with data by deploying Grafana, reducing time to decision on “drop everything calls” by 95%

Khoury Advising | Software Engineering Consultant

- Built ML platform rolled out by Khoury Co-op Advisors to engage 2,500+ students, connecting them to experiences that strengthen co-op readiness and resumability while giving advisors data-driven insights to personalize guidance
- Implemented HuggingFace MPNet embeddings and semantic similarity to personalize experience recommendations
- Deployed scalable ECS Fargate backend with RDS database, CloudWatch monitoring, and Google OAuth via Supabase

Water Horizons | Software Data Engineer — **August 2025 – Present**

- Developed desktop app enabling researchers to query paper libraries and discover relevant articles beyond their collection
- Built RAG pipeline leveraging FAISS-based semantic search with OpenAI and Ollama for context-aware research queries
- Built Electron app with FastAPI backend integrating SQLite, Zotero, and Google Scholar scraping for document discovery

Khoury College | Database Design Teaching Assistant — **January 2026 - Present**

- Supporting 200+ student class covering SQL, relational algebra, relational database design, ER modeling, and schema normalization, by hosting weekly office hours and grading assignments and exams

PROJECTS

StrideTrack | Data Tech Lead - Generate, Northeastern’s Product Development Studio — **December 2025 - Present**

- Leading team of 9 for track shoe sensor startup, hired engineers and advised PhD founder on product architecture, MVP scope, and hardware-software integration through weekly stakeholder meetings within 16-week timeline

Northeastern Speech, Hearing, and Communication Lab | ML Research Assistant — **February 2025 – Present**

- Advanced PhD research on speech perception by extracting audiovisual cues from video for fMRI neuroimaging analysis
- Built CV pipeline with YOLOv8, MediaPipe, and a custom RNN to model temporal speech cues across video datasets
- Utilized Discovery HPC cluster for large-scale video processing and model training, doubling throughput with CUDA

Quoridor Game (quoridor.sambaldwin.dev) — **June 2025 – July 2025**

- Independently developed strategy game with different level AI opponents using custom decision algorithms
- Utilized Dijkstra's and custom decision tree for optimal pathfinding with defense logic and strategic fence placements