

LAERDON KIM

(312) 770-0936 | lyk25@cornell.edu | <https://linkedin.com/in/laerdon> | <https://github.com/laerdon> | <https://laerdon.github.io>

EDUCATION

Cornell University

B.A., Computer Science, Mathematics — GPA: 4.13/4.0

Ithaca, NY

Expected May 2027

Activities: **Academic Team Lead** on the Association of Computer Science Undergraduates

Awards: Milstein Scholar in Technology and Humanity, Cosmos Institute Human-AI Interaction Grantee (~8% acceptance)

EXPERIENCE

Machine Learning/NLP Intern

June 2025 – August 2025

Cornell University, Natural Language Processing Group

Ithaca, NY

- Co-authored EMNLP 2025 Main paper exposing critical flaw in SOTA lexical semantic change detection methods, proposing new evaluation metric to reduce false positives.
- Solo-authored publication in NAACL 2025, CLPsych Workshop. Created evaluation pipeline for LLM identification of self-state spans in mental health data. **Won third out of fourteen systems.**
- Researching LLMs for forecasting online toxicity in time-series conversational data. Designed PIVDelay, a novel technique reducing false positive rate by 13%. Presented findings at Cornell BURE Symposium.

Software Developer

September 2024 – Present

ConvoKit (Open-source NLP Package). 588 Github stars.

- Work on a team of 1 of 5 primary maintainers. Developed a modular wrapper class for HuggingFace and custom-trained model use.
- Reviewed and merged PRs for bug fixes, added 4 new datasets, and wrote 6 demo Jupyter notebooks. Modernized backend for forecasting system with Docker containerization. ~6k lines contributed.
- Developed modular data loaders and preprocessing scripts to handle LLM inference on large dialogue datasets (10k+ conversations).

Machine Learning Engineering Intern

June 2024 – August 2024

Illinois Institute of Technology

Chicago, Illinois

- Coded custom CUDA fused attention and LayerNorm kernels for a 1.2x inference speedup on diffusion models.
- Utilized z-score and absolute value threshold outlier masking methods on LLaMA 2, resulting in a 6% reduction in perplexity and 13% reduction in memory usage. Re-implemented SmoothQuant and LLM-AWQ on OPT-8B for text completion inference.

Software Engineering Intern

May 2023 – August 2023

Innocuous AI

New York, NY

- Developed startup website and branded landing page for VC pitches using React and Node.js.
- Designed custom assets in Figma and Photoshop and bundled JS/CSS to reduce requests for ~40% latency reduction.

PUBLICATIONS & PRESENTATIONS

Khonzoda Umarova, Lillian Lee, Laerdon Kim. *Current semantic-change quantification methods struggle with semantic change discovery in the wild*. Accepted at EMNLP 2025 Main Conference.

Laerdon Kim. *A Baseline for Self-state Identification and Classification in Mental Health Data: CLPsych 2025 Task*. Accepted at NAACL 10th Workshop on Computational Linguistics and Clinical Psychology (CLPsych). PI: Cristian Danescu-Niculescu-Mizil. arXiv:2504.14066

Laerdon Kim, Jenny Jin, Ajahla Jefferson, Tina Zeng, Brian Chica-Herrera. "Social Mobility in the Mission Hill/Roxbury Area." Presentation at Opportunity Insights Lab, Harvard University, Boston, MA, May 2024.

PROJECTS

Vocera | AI deepfake detection application

2025

- Built React Native app at UC Berkeley AI Hackathon for real-time voice verification using wav2vec embeddings, feature extraction with OpenSMILE, and speech-to-text with OpenAI Whisper.
- Processed 100+ audio samples, achieving 90%+ accuracy in deepfake detection and average < 3000 ms latency.

EngagingWorlds | Educational technology simulation platform

2023

- Serialized and stored conversation data using JSON and SQLite, enabling efficient analytics retrieval for hackathon-wide demo.
- Developed C# Unity VR "storybook exploration" experience via fine-tuning language models on book data; wrote Flask backend with Claude API to convert conversation transcripts into class analytics for educators. **1st place at MIT AI + Education Hackathon.**

Geospatial Data Analysis and Web Scraping with Python & QGIS | Full-stack processing and visualization

2023

- Using QGIS, Pandas, BeautifulSoup, and SQL, created econometrics data analysis workflow including scraping, regression analysis, and visualization using the Opportunity Atlas API. Presented to Harvard University's Opportunity Insights Laboratory.

TECHNICAL SKILLS

Languages: Python (proficient), Java (proficient), Swift (proficient), C++ (intermediate), HTML/JavaScript/TypeScript/CSS (proficient)

Tools: PyTorch, CUDA, Supabase, AWS, SQL/PostgreSQL, React, Vue, Tailwind, Flask, Docker, GCP

Fluent in common ML packages, i.e. scikit-learn, seaborn, numpy, pandas, huggingface, transformers