

Alexander Loftus

AI researcher & communicator with 7+ years translating deep learning & machine learning research into actionable insights. Kaggle \$100k competition winner, forthcoming CUP textbook author, and organizer of a 200-person mechanistic interpretability conference. Research in code interpretability and evaluation for LLMs. Seeking outward-facing role where deep technical depth, public speaking, and storytelling abilities shine.

Career highlights:

Textbook author: Authored a 524-page technical book on statistical network ML (Cambridge Univ. Press, Nov 2025)

Organizer & Teacher: Organized the New England Mechanistic Interpretability (NEMI) conference; YouTube lecture series creator; taught hundreds of students through meetups, summer camps, and tutorials.

High-impact presentations: Best poster award at NeurIPS 2023 LatinX workshop, first author work in ICLR 2024, delivered 10+ invited talks to 20-300 attendees;

Strategic Advisory Roles: Advisor for cybersecurity/mechanistic interpretability startup; summer mentor for Harvard/MIT students.

Competitive Excellence: Part of a 4-person team that won 1st place in a \$100k Kaggle competition (1,249 teams); featured on the cover of *Scientific American*.

EXPERIENCE

Data Scientist

San Diego, CA

Creyon Bio

2022-2024

Large Protein Models For Splice-Site Prediction: Explored splice site prediction in LLMs trained on protein sequences.

Pre-training, fine-tuning, reinforcement learning, and benchmarking+evals.

ML for Toxicity Prediction: Developed a novel contrastive learning pipeline to predict oligo toxicity from 3-D electrostatic maps; increased classification AUC from 0.73 to 0.88.

Neuron Toxicity Detection: Developed scalable neuron segmentation and toxicology classification pipeline.

All Projects: Helped shape Series B narrative through presentations to C-suite.

Machine Learning Research Engineer

Rockville, MD

Blue Halo

2021-2022

Conditional Image Generation with Generative Adversarial Networks: Replaced GAN pipeline with diffusion-model synthetic data generator. Immediate 10x training run reliability.

Detecting Objects with Enhanced Yolo and Knowledge Graphs: Led knowledge graph effort for object detection project. Delivered live demos to program officers.

Geometric Multi-Resolution Analysis: Led infra for document clustering & analysis method.

Research Assistant

Baltimore, MD

Johns Hopkins University — Dr. Joshua Vogelstein

2018-2021

MRI-to-Graphs: Optimized a diffusion MRI pipeline with Kubernetes, Docker, and AWS Batch. Halved runtime and cut cloud costs by 40%.

Graspologic: Co-authored an open-source graph statistics library. Later adopted by Microsoft Research for large-scale network analysis.

Assistant Director

Seattle, WA

iD Tech Camps — University of Washington

2014-2018 summers

Leader and Manager: Managed 10+ instructors/week and 300+ students.

Curriculum Designer: Authored game development curriculum deployed to 50+ locations, impacting 10k+ students nationwide.

EDUCATION

Northeastern University

Boston, MA

PhD Student, Computer Science

2024-Present

Advisor: Dr. David Bau

Mechanistic interpretability, evaluations, and training dynamics in Code LLMs.

Johns Hopkins University

Baltimore, MD

MSE Biomedical Engineering: Machine Learning & Data Science Focus

2020-2022

Advisor: Dr. Joshua Vogelstein

Thesis: Hands-On Network Machine Learning

dean's list, highest honors, GPA 3.97/4.0.

Western Washington University
BS Behavioral Neuroscience — *Minors*: Chemistry, Philosophy
Founder & President, Computational Neuroscience Club
Vice President, Neuroscience Club

Bellingham, WA
2014-2018

Built computational neuroscience club from scratch, taught weekly seminars.

SKILLS SUMMARY

Languages: Python, Bash, R, Rust, SQL

Tools & Frameworks: pytorch, pytorch-lightning, tensorflow, jax, numpy, scipy, pandas, polars, sklearn, seaborn, matplotlib, docker, AWS, google cloud (GCP), photoshop, SQL, weights & biases, mlflow, kubernetes, linux, cursor, ray, vllm

Areas of Expertise: Linear algebra, probability & statistics, deep learning, information theory, transformers, diffusion models, convolutional autoencoders, GPUs and CUDA, public speaking, leadership & management, teaching, natural language processing, computer vision

Soft Skills: Public Speaking, Technical Writing, Leadership, Mentorship, Community-Building, Confidence & Charisma

LEADERSHIP & COMMUNITY ENGAGEMENT

Conference Organizer

Running 200+ person interpretability conference; Raised \$17,000 grant funding.

NEMI
2025

Research Mentor

Mentoring Harvard/MIT students in Summer 2025

CBAI
2025

Strategic Advisor

Advisor to cybersecurity-focused startup specializing in interpretability tooling for AI systems.

Krnel.ai
2025

Meetup Speaker

Speaker & organizer for San Diego AI Meetups.

SDML
2023–2024

Hackathon Organizer

Helped organize hackathon & workshop to explore statistics for high-dimensional testing.

NeuroData Workshop
2019

TALKS & DEMOS

A Shared Infrastructure for Interpretability: *Tech. Innovations for AI Policy Conf., 2025*

Invited demo for policymakers; showcased live editing of GPT2 internals

State of the Art in Knowledge Editing: *A.R. Loftus, 2023*

Survey talk on LLM knowledge-editing methods.

1st Place Solution - Vesuvius Ink Competition: *R. Chesler, A.R. Loftus, A. Tersol Montserrat, T. Kyi, 2023*

Walkthrough of winning \$100,000 ink-detection model.

ICML Conference Highlights: *A.R. Loftus, 2023*

Selected breakthroughs from ICML. Presented to biotech execs and SDML meetup group.

Working with LLMs: *AI San Diego Conference, 2023.*

Invited talk: Introduction to LLM engineering. 300+ attendees

Linear Algebra, from Dot Products to Neural Networks: *A.R. Loftus, 2023.*

Created a YouTube tutorial series on the fundamentals of linear algebra for machine learning.

TEXTBOOK

Hands-On Network Machine Learning with Python: *Eric Bridgeford, Alexander R. Loftus, Joshua Vogelstein.*

Cambridge University Press, in copy-editing phase. To be printed November 2025.

Spectral representation theory on networks. 524 pages, 147 figures.

SELECTED PUBLICATIONS

* indicates equal contribution.

🏆 indicates best poster.

NNsight and NDIF: Democratizing Access to Open-Weight Foundation Model Internals: *A.R. Loftus**, *J.Fiotto-Kaufman**, et al. ICLR 2025.

Open-source suite for probing & manipulating LLM weights without engineering overhead. Ray GCS Service backend with AWS object storage and VLLM for inference speed.

A Saliency-based Clustering Framework for Identifying Aberrant Predictions 📄: *A. Tersol Montserrat, A.R. Loftus, Y. Daihes*. Paper, **NeurIPS LatinX AI Workshop**, 2023.
Detects spurious feature reliance via saliency embeddings.

A low-resource reliable pipeline to democratize multi-modal connectome estimation and analysis: *J. Chung, R. Lawrence, A.R. Loftus, et al*. Paper, in review at *Nature Methods*, 2024
Transforms diffusion MRI scans into graphs; open-sourced ([code](#))

FELLOWSHIPS & AWARDS

First Place Winner Kaggle Vesuvius Competition, \$100,000.	2023
Khoury Distinguished Fellowship Northeastern University PhD fellowship.	2024
Best Poster Award NeurIPS 2023 LatinX AI Workshop.	2023
Harvard AI Safety Technical Fellowship Harvard fellowship for technical work in AI safety.	2025
AWS Research Grant \$10,000 grant for computational research on cloud services.	2019

TEACHING

Head Teaching Assistant Foundations of Computational Biology and Bioinformatics, <i>EN.BME.410/634</i>	Johns Hopkins University Spring 2021
Teaching Assistant NeuroData Design II, <i>EN.BME.438/638</i>	Johns Hopkins University Spring 2020
Teaching Assistant NeuroData Design I, <i>EN.BME.437/637</i>	Johns Hopkins University Fall 2019
Teaching Assistant Introduction to Behavioral Neuroscience, <i>PSY.220</i>	Western Washington University Winter 2017
Curriculum Designer Built curriculum used across 50 locations in the United States by tens of thousands of students.	iD Tech Camps Spring 2017
Instructor Taught programming and game design to high school students.	iD Tech Camps 2014-2018 summers

FUN

Gaming: Starcraft 2 grandmaster in high school, local tournament winner

Music: Fingerstyle guitarist; performed at open mic nights.

Dancing: Partner dance instructor and competition winner (Fusion, West Coast Swing, Zouk)