

Amy X. Lu

<http://amyxlu.github.io>
github.com/amyxlu

amyxlu@cs.toronto.edu
+1-647-968-6833
[linkedin.com/in/amyxlu](https://www.linkedin.com/in/amyxlu)

SUMMARY

I am a graduate student at the Vector Institute for AI, applying **computer vision and natural language processing algorithms to biological sequence data**, with an emphasis on **self-supervised representation learning, generalizability, and biological interpretability**. I also serve on SVAI, which organizes research collaborations of genomics/AI researchers and industry partners for understudied rare disease patients.

EDUCATION

- **University of Waterloo** Waterloo, Canada
Bachelors of Science, Honours Science, Bioinformatics Option 2014 – 2019
- **University of Toronto, Vector Institute** Toronto, Canada
Masters, Computer Science 2019 – Present

EXPERIENCE

- **Stanford University** Palo Alto, USA
Visiting Student Researcher — Domain Adaptation in Regulatory Genomics, Dr. Anshul Kundaje 2019 – Present
 - **Domain adaptation, ChIP-seq analyses:** Using domain adaptation to improve classifier performance and learned genomic sequence features for transcription factor binding when evaluating in a different cellular context.
- **University of Toronto/Vector Institute** Toronto, Canada
Masters Student — Representation learning in genomics, Dr. Alan Moses & Dr. Marzyeh Ghassemi 2019 – Present
 - **Natural language processing, self-supervised learning:** Applying contextual word embedding models to learn robust and interpretable representations for genomic and proteomic sequences.
 - **Algorithmic fairness:** Quantitative and qualitative evaluation of bias in “black box” contextual clinical word embeddings, and attempting to debias using adversarial gradient-reversal. (*Submitted to FAT* 2020*)
 - **Generalizability, computer vision:** Benchmarked a new “paired cell in-painting” method for out-of-sample generalization in microscopy against various common computer vision algorithms. (*NeurIPS 2019; ICML 2019 Self-Supervised Learning Workshop*)
- **Harvard Medical School/Boston Children’s Hospital** Boston, USA
Research Intern — Deep Learning in Clinical Genomics, Sliz Lab 2018 – Present
 - **Variant analysis, genotype-phenotype studies:** Applying standard filtering pipelines to a whole exome (WES) dataset of epilepsy patients
 - **Machine learning, dimensionality reduction:** Building machine learning models to classify disease phenotype from high-dimensional exome variant data
- **University of Waterloo** Waterloo, Canada
Undergraduate Thesis Student — Deep Learning in Regulatory Genomics, Doxey Lab 2017 – 2018
 - **Convolutional neural networks in genomics:** Implemented a convolutional neural network which recognizes genomic enhancers regulating femur growth with 91% accuracy.
 - **Interpretability:** Distilled biological knowledge from the “black box” network by reconstructing first-layer features as position-weighted matrices (PWMs) of “motifs” (i.e. genomic patterns). All motifs found statistical matches in JASPAR (an experimentally-verified motifs database).
 - **Self-directed learning:** Self-learned deep learning concepts and Python as a biology student.
- **École polytechnique fédérale de Lausanne** Lausanne, Switzerland
Research Intern — Computational Biology, Laboratory for Biomolecular Modelling 2017
 - **Molecular dynamics:** Contributed three potential findings on enzyme-membrane mechanisms of bacterial “superbugs” using computational simulations based on molecular dynamics (MD) principles.
 - **Automating workflows:** Increased workflow efficiency by writing R, tcl/tk, and Csh scripts.
- **University of Toronto** Toronto, Canada
Research Intern — Data Visualization in Pharmacoepidemiology, Cadarette Lab 2015 – 2017

- **Pharmacoepidemiology:** Took initiative to redesign manual workflows. Contributed two co-authored publications, one poster, and figures for an international epidemiology conference within two months.
- **Data visualization:** Wrangled data using R and SQL to create animated and interactive co-authorship networks.

SERVICE AND ACTIVITIES

- **SVAI:** Core Team of SVAI, a non-profit connecting patients of rare genomic diseases to the medical/AI research community and industry partners through collaborative research initiatives.
- **Residence Don, Velocity Incubator:** Leader for the residence program of Canada's most productive start-up incubator; fostered close relations with diverse students to support conduct and mental health crises
- **Tosamaganga Hospital, Tanzania:** Supported operations at a rural Tanzanian hospital and shadowed surgical procedures. Expenses were covered by scholarships, fundraising, and part-time tutoring.
- **Teaching Assistant:** Taught weekly tutorial lectures for a 200-level genetics course.

AWARDS

- **NSERC Michael Smith Foreign Supplement:** Supports high-calibre Canadian graduate students in pursuing research abroad. Valued at \$6,000.
- **Canada Graduate Scholarships-Master's (CGS-M) Award:** NSERC federal research grant. Valued at \$17,500.
- **Scholarship of Excellence in Research:** One of 13 students sponsored for research at EPFL. Valued at CHF 4500.
- **University of Waterloo:** Various entrance awards totaling \$6000.
- **Associate of The Royal Conservatory (ARCT):** Performer's Diploma in Piano (highest academic standing offered).

PUBLICATIONS AND CONFERENCE PROCEEDINGS

- **Accepted:** Lu AX, Lu AXP, Moses A. The Cells Out of Sample (COOS) dataset and benchmarks for measuring out-of-sample generalization of image classifiers. *NeurIPS 2019*.
- **Published:** Ban J, Tadrous M, Lu AX, Cicinelli EA, Cadarette SM. Diffusion of indirect comparison meta-analytic methods to study drugs: a systematic review and co-authorship network analysis. *BMJ Open*.
- **Under revision:** Consiglio GP, Maclure M, Lu AX, Cicinelli EA, Ban JK, McCarthy L, Cadarette SM. Guidance documents for the application of Self-controlled Crossover Observational PharmacoEpidemiology (SCOPE) designs are needed: systematic review and animated co-authorship networks. *Submitted to Pharmcoepi Drug Saf*.

WORKSHOP PAPERS

- Lu AX, Lu AXP, Moses A. Paired Cell Inpainting: A Multiple-Instance Extension of Self-Supervised Learning for Bioimage Analysis. *ICML 2019 Workshop on Self-Supervised Learning*.

POSTERS

- **Presented:** Lu AX, Rockowitz S, Poduri A, Sliz P. From data to precision medicine: predictive machine learning models to uncover disease-associated variants. Poster presentation, November 2019. *Harvard Medical School BCMP Retreat 2019*.
- **Presented:** Lu AX, Consiglio GP, Cadarette SM. Dynamic Visualization in Co-Authorship Network Analysis. Poster presentation, August 2016. *Leslie Dan Faculty of Pharmacy Undergraduate Research Symposium*.
- **Presented:** McIlroy-Young R, Lu AX, Guenther N, Olarnyk A. A Network Analysis: Did the Arab Spring Impact the Academic Network of Tunisia Between 2010 and 2015? Poster presentation, April 2015. *Knowledge Integration Symposium*.
- **Submitted:** Lu AX, Tam ES. Effect of prophylactic brimonidine on subconjunctival hemorrhage in laser-assisted cataract surgery. Poster abstract. *Submitted to American Academy of Ophthalmology 2017*.

HACKATHONS AND PROJECTS

- **NLP-seq:** Used n-grams to extract features for supervised learning of function prediction, trained on Protein Data Bank data.
- **Chatsense:** Mobile chat app using a sentiment analysis machine learning API on voice messages and mapping dominant emotions to emojis, designed for the autistic community. *Team project. Winner at YHacks at Yale*.
- **Alice:** Web login page using a trained facial recognition model to login to online banking without passwords, designed for dementia patients. *Team project. Winner at ConUHacks III*.