

**RESEARCH INTERESTS** My general research interest is in developing methods to learn efficient and robust representations of discrete-valued sequence data (specifically natural language). I'm also interested in understanding and finding ways to fix their pathologies once trained. My research broadly covers topics in representation learning, interpretability, and out-of-domain robustness/generalization.

**EDUCATION** **Massachusetts Institute of Technology** Cambridge, MA  
Visiting Scholar  
Advisor: Prof. Marzyeh Ghassemi  
*Sept 2021 – present*

**University of Toronto** Toronto, Ontario  
Ph.D. Machine Learning  
Advisor: Prof. Marzyeh Ghassemi  
*Sept 2019 – present*

**University of California San Diego** San Diego, California  
BS Computer Science (Summa Cum Laude)  
Advisor: Prof. Zachary Lipton and Prof. Julian McAuley  
*Sep 2014 – Jun 2018*

**PROFESSIONAL EXPERIENCE** **Prescient Design** New York, New York  
Research Intern (Kyunghyun Cho)  
*Blind Biological Sequence Denoising with Self-Supervised Set Learning*  
*Summer 2022*

**Meta** New York, New York (Virtual)  
Research Intern (Naman Goyal)  
*Growing Switch Transformers for Multilinguality*  
*Summer 2021*

**Google** Mountain View, California (Virtual)  
Research Intern (Qi Guo)  
*Improving Dialogue Breakdown Detection with Semi-Supervised Learning*  
*Summer 2020*

**Meta (Full Time)** Menlo Park, California  
Research Engineer (Michael Auli)  
*Sep 2018 – Sep 2019*

**Meta** Menlo Park, California  
Software Engineering Intern  
*Summer 2016 / Summer 2017*

**Qualcomm** San Diego, California  
Software Engineering Intern  
*Summer 2015*

- REFEREED PUBLICATIONS**
1. **N. Ng**, J. W. Park, J. H. Lee, R. Kelly, S. Ra, and K. Cho. “Blind Biological Sequence Denoising with Self-Supervised Set Learning”. In: *TMLR*. 2024.
  2. **N. Ng**, N. Hulkund, K. Cho, and M. Ghassemi. “Predicting Out-of-Domain Generalization with Neighborhood Invariance”. In: *TMLR*. 2023.
  3. J. Bae, **N. Ng**, A. Lo, M. Ghassemi, and R. Grosse. “If Influence Functions are the Question, What is the Answer?” In: *Proc. of NeurIPS*. 2022.
  4. **N. Ng**, K. Cho, and M. Ghassemi. “SSMBA: Self-Supervised Manifold Based Data Augmentation for Improving Out-of-Domain Robustness”. In: *Proc. of EMNLP*. 2020.
  5. T. Lau, **N. Ng**, J. Gingold, N. Desai, J. McAuley, and Z. C. Lipton. “Embryo staging with weakly-supervised region selection and dynamically-decoded predictions”. In: *Proc. of Machine Learning for Healthcare*. 2019.

	6. <b>N. Ng</b> , K. Yee, A. Baevski, M. Ott, M. Auli, and S. Edunov. “Facebook FAIR’s WMT19 News Translation Task Submission”. In: <i>Proc. of WMT</i> . 2019.	
	7. K. Yee, <b>N. Ng</b> , Y. Dauphin, and M. Auli. “Simple and Effective Noisy Channel Modeling for Neural Machine Translation”. In: <i>Proc. of EMNLP</i> . 2019.	
	8. <b>N. Ng</b> , R. Gabriel, J. McAuley, C. Elkan, and Z. Lipton. “Predicting surgery duration with neural heteroscedastic regression”. In: <i>Proc. of Machine Learning for Healthcare</i> . 2017.	
WORKSHOP PUBLICATIONS	1. <b>N. Ng</b> , N. Thangarajan, J. Pan, M. Ghassemi, and Q. Guo. “Improving Dialogue Breakdown Detection with Semi-Supervised Learning”. In: <i>Proc. of Workshop on Human in the Loop Dialogue Systems at NeurIPS</i> . 2020. Oral.	
	2. M. Ott, S. Edunov, A. Baevski, A. Fan, S. Gross, <b>N. Ng</b> , D. Grangier, and M. Auli. “fairseq: A fast, extensible toolkit for sequence modeling”. In: <i>Proc. of NAACL-HLT: Demonstrations</i> . 2019.	
	3. <b>N. Ng</b> , J. McAuley, Z. Lipton, and N. Desai. “Predicting Embryo Morphokinetics in Videos with Late Fusion Nets & Dynamic Decoders”. In: <i>Proc. of ICLR Workshops</i> . 2018.	
SHARED TASKS	<b>1st</b> in Dialogue Breakdown Detection Challenge English task	2020
	<b>1st</b> in WMT News Translation English ↔ German task	2019
	<b>1st</b> in WMT News Translation English ↔ Russian task	2019
PROFESSIONAL ACTIVITIES	<b>Head Organizer</b>	
	Workshop on Robustness in Sequence Modeling at NeurIPS	2022
	<b>Reviewer</b>	
	NeurIPS	2022, 2023
	ICLR	2023
	Machine Learning for Healthcare	2020
TEACHING AND TALKS	<b>University of Toronto</b>	Teaching Assistant
	CSC 2515: Introduction to Machine Learning (Graduate Level)	Fall 2020
	CSC 2541: Topics in Machine Learning: Machine Learning for Health	Winter 2020
	CSC 311: Introduction to Machine Learning	Fall 2019
	<b>Meta</b>	Internal Lecturer
	Special Topics in Deep Learning: NLP and Translation	Feb 2019, Sep 2019
	<b>University of California, San Diego</b>	Teaching Assistant
	CSE 101: Design and Analysis of Algorithms	Winter 2018
	CSE 158: Web Mining and Recommender Systems	Fall 2017
	CSE 21: Mathematics for Algorithms and Systems	Winter 2017
	CSE 11: Introduction to Object-Oriented Programming	Fall 2015
HONORS AND AWARDS	• Jacobs Scholarship, University of California San Diego	2014
	• Regents Scholarship, University of California San Diego	2014