

Alexis J. Ross

alexisro@mit.edu

<https://alexisjihyeross.github.io/>

EDUCATION	Massachusetts Institute of Technology , Cambridge, Massachusetts Ph.D. in EECS <i>Advisor</i> : Jacob Andreas	Sept 2022 – Present
	Harvard University , Cambridge, Massachusetts A.B. with Honors in Computer Science and Philosophy (Joint)	Sept 2016 – May 2020
RESEARCH EXPERIENCE	Allen Institute for AI (AI2), AllenNLP Predoctoral Researcher <i>Mentors</i> : Ana Marasović and Matthew E. Peters	July 2020 – July 2022
	Harvard University, Computer Science and Philosophy <i>Mentors</i> : Himabindu Lakkaraju and Bernhard Nickel <i>Senior Thesis</i> : Using Linear Approximations to Explain Complex, Blackbox Classifiers	Sept 2019 – May 2020
	Microsoft Research Research Intern, Project Hanover <i>Mentor</i> : Tristan Naumann <i>Project</i> : Machine Reading for Precision Medicine	May 2019 – Aug 2019
	Johns Hopkins, Center for Language and Speech Processing Undergraduate Researcher, JSALT Workshop <i>Mentor</i> : Ellie Pavlick <i>Project</i> : General-Purpose Sentence Representation Learning	June 2018 – Aug 2018
	Harvard University, Cognitive Neuropsychology Laboratory Research Assistant <i>Mentors</i> : Gilles Vannuscorps and Alfonso Caramazza <i>Project</i> : Object Orientation Perception	Feb 2017 – Aug 2017
PREPRINTS	[1] Reasoning or Reciting? Exploring the Capabilities and Limitations of Language Models Through Counterfactual Tasks . Zhaofeng Wu, Linlu Qiu, Alexis Ross , Ekin Akyurek, Boyuan Chen, Bailin Wang, Najoung Kim, Jacob Andreas, Yoon Kim. <i>Preprint</i> .	
	[2] ARIES: A Corpus of Scientific Paper Edits Made in Response to Peer Reviews . Mike D'Arcy, Alexis Ross , Erin Bransom, Bailey Kuehl, Jonathan Bragg, Tom Hope, Doug Downey. <i>Preprint</i> .	
PUBLICATIONS	[3] CREST: A Joint Framework for Rationalization and Counterfactual Text Generation . Marcos Treviso, Alexis Ross , Nuno M. Guerreiro, Andre Martins. <i>ACL</i> , 2023.	
	[4] Does Self-Rationalization Improve Robustness to Spurious Correlations? . Alexis Ross , Matthew Peters, Ana Marasović. <i>EMNLP</i> , 2022.	
	[5] Tailor: Generating and Perturbing Text with Semantic Controls . Alexis Ross *, Tongshuang Wu*, Hao Peng, Matthew Peters, Matt Gardner. <i>ACL</i> , 2022.	
	[6] Learning Models for Actionable Recourse . Alexis Ross , Himabindu Lakkaraju, and Osbert Bastani. <i>NeurIPS</i> , 2021.	
	[7] Competency Problems: On Finding and Removing Artifacts in Language Data . Matt Gardner*, William Merrill*, Jesse Dodge, Matthew E. Peters, Alexis Ross , Sameer Singh, Noah Smith. <i>EMNLP</i> , 2021.	

- [8] [Explaining NLP Models via Minimal Contrastive Editing \(MiCE\)](#).
Alexis Ross, Ana Marasović, and Matthew E. Peters. *Findings of ACL, 2021*.
- [9] [How well do NLI models capture verb veridicality?](#)
Alexis Ross and Ellie Pavlick. *EMNLP 2019*. **Oral presentation**.
- [10] [Probing what different NLP tasks teach machines about function word comprehension](#).
Najoung Kim, Roma Patel, Adam Poliak, Alex Wang, Patrick Xia, R Thomas McCoy, Ian Tenney, Alexis Ross, Tal Linzen, Benjamin Van Durme, Samuel R Bowman, and Ellie Pavlick. **SEM 2019*. **Best Paper Award**.

SELECTED HONORS	2023	Diverse Intelligences Summer Institute Fellowship	
	2023	NSF Graduate Research Fellowship	
	2023	Third Prize, Inverse Scaling Competition	
	2022	MIT Shillman Fellowship	
	2020	Thomas T. Hoopes Prize “for extraordinary undergraduate research” (senior thesis)	
	2019	EMNLP Student Travel Scholarship	
	2018	Certificate of Distinction in Teaching for a student evaluation of 4.8/5.0	
INVITED TALKS		Towards Adaptive Teaching with Large Language Models, DI Summit	July 2023
		Towards Model Development with Tailored Supervision, IST & Unbabel Seminars	April 2023
		Generating and Perturbing Text with Semantic Controls, AI2 (All Hands)	August 2021
SERVICE	2022	Co-organizer , cs-sop.org, resource for CS PhD applicants	
	2022	Mentor , MIT EECS Graduate Application Assistance Program (GAAP)	
	2021	Co-organizer and frequent host, NLP Highlights Podcast, over 9000 views in October	
	2021	Co-organizer , AllenNLP Hacks, a hackathon for minoritized students Presented workshop: “My First NLP Project”	
	2021	Secondary Reviewer , ACL	
	2019	Mentor , Harvard Women in Computer Science	
TEACHING EXPERIENCE		CS 121: Introduction to the Theory of Computation	Sept 2018 – Dec 2018
		Teaching Fellow, Harvard School of Engineering and Applied Sciences <i>Instructor</i> : Boaz Barak	
OTHER EXPERIENCE		Solo, Chamber, and Orchestral Violinist	2002 – Present
	2023 –	MIT Chamber Music Society	
	2016 – 2018	Violinist in the Harvard-Radcliffe Orchestra	
	2016 – 2020	Chamber Music Performance at Harvard. Coaches: Parker Quartet	
	2014 – 2016	Juilliard Pre-College, Diploma in Violin Performance. Teacher: Ann Setzer	
	2013 – 2015	Summers at the Meadowmount School of Music	
		Mental Health Advocacy and Counseling, Harvard University	2017 – 2020
	2017 – 2020	Events Co-Chair, Student Mental Health Liaisons	
	2018 – 2020	Peer Mental Health Counselor, Room 13	