## Alexis J. Ross

## alexisro@mit.edu https://alexisjihyeross.github.io/

EDUCATION Massachusetts Institute of Technology, Cambridge, Massachusetts Sept 2022 – Present

Ph.D. in EECS

Advisor: Jacob Andreas

Harvard University, Cambridge, Massachusetts

Sept 2016 – May 2020

A.B. with Honors in Computer Science and Philosophy (Joint)

RESEARCH Allen Institute for AI (AI2), AllenNLP July 2020 – July 2022

EXPERIENCE Predoctoral Researcher

Mentors: Ana Marasović and Matthew E. Peters

Harvard University, Computer Science and Philosophy Sept 2019 – May 2020

Mentors: Himabindu Lakkaraju and Bernhard Nickel

Senior Thesis: Using Linear Approximations to Explain Complex, Blackbox Classifiers

Microsoft Research May 2019 – Aug 2019

Research Intern, Project Hanover

Mentor: Tristan Naumann

Project: Machine Reading for Precision Medicine

Johns Hopkins, Center for Language and Speech Processing

June 2018 – Aug 2018

Undergraduate Researcher, JSALT Workshop

Mentor: Ellie Pavlick

Project: General-Purpose Sentence Representation Learning

Harvard University, Cognitive Neuropsychology Laboratory Feb 2017 – Aug 2017

Research Assistant

Mentors: Gilles Vannuscorps and Alfonso Caramazza

Project: Object Orientation Perception

Publications [1] Does Self-Rationalization Improve Robustness to Spurious Correlations?.

Alexis Ross, Matthew Peters, Ana Marasović. EMNLP, 2022.

[2] Tailor: Generating and Perturbing Text with Semantic Controls.

Alexis Ross\*, Tongshuang Wu\*, Hao Peng, Matthew Peters, Matt Gardner. ACL, 2022.

[3] Learning Models for Actionable Recourse.

Alexis Ross, Himabindu Lakkaraju, and Osbert Bastani. NeurIPS, 2021.

[4] 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. *EMNLP*, 2021.

[5] Explaining NLP Models via Minimal Contrastive Editing (MiCE).

Alexis Ross, Ana Marasović, and Matthew E. Peters. Findings of ACL, 2021.

[6] How well do NLI models capture verb veridicality?

Alexis Ross and Ellie Pavlick. EMNLP 2019. Oral presentation.

[7] 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	NSF Graduate Research Fellowship Third Prize, Inverse Scaling Competition MIT Shillman Fellowship Thomas T. Hoopes Prize "for extraordinary undergraduate research" (senior thesis) EMNLP Student Travel Scholarship Certificate of Distinction in Teaching for a student evaluation of 4.8/5.0
INVITED TALKS	Apr. 2023 Towards Model Development with Tailored Supervision, <b>IST &amp; Unbabel Seminars</b> Aug. 2021 Generating and Perturbing Text with Semantic Controls, <b>AI2 (All Hands)</b>
SERVICE	<ul> <li>Co-organizer, cs-sop.org, resource for CS PhD applicants</li> <li>Mentor, MIT EECS Graduate Application Assistance Program (GAAP)</li> <li>Co-organizer and frequent host, NLP Highlights Podcast, over 9000 views in October</li> <li>Co-organizer, AllenNLP Hacks, a hackathon for minoritized students         Presented workshop: "My First NLP Project"     </li> <li>Secondary Reviewer, ACL</li> <li>Mentor, Harvard Women in Computer Science</li> </ul>
Teaching Experience	CS 121: Introduction to the Theory of Computation  Teaching Fellow, Harvard School of Engineering and Applied Sciences  Instructor: Boaz Barak  Sept 2018 – Dec 2018
OTHER EXPERIENCE	Solo, Chamber, and Orchestral Violinist2002 – Present2023 –MIT Chamber Music Society2016 – 2018Violinist in the Harvard-Radcliffe Orchestra2016 – 2020Chamber Music Performance at Harvard. Coaches: Parker Quartet2014 – 2016Juilliard Pre-College, Diploma in Violin Performance. Teacher: Ann Setzer2013 – 2015Summers at the Meadowmount School of Music
	Mental Health Advocacy and Counseling, Harvard University 2017 – 2020 Events Co-Chair, Student Mental Health Liaisons

Peer Mental Health Counselor, Room 13

2018 - 2020