

Alexis J. Ross

alexisr@allenai.org
<https://alexisjihyeross.github.io/>

EDUCATION	Harvard University , Cambridge, Massachusetts A.B. with Honors in Computer Science and Philosophy (Joint)	Sept 2016 – May 2020
RESEARCH EXPERIENCE	Allen Institute for Artificial Intelligence (AI2), AllenNLP Predoctoral Researcher <i>Mentors:</i> Ana Marasović and Matthew E. Peters	July 2020 – Present
	Harvard University, Computer Science and Philosophy <i>Mentors:</i> Himabindu Lakkaraju and Bernhard Nickel <i>Senior Thesis:</i> Explaining Blackbox Models with Linear Approximations Awarded Thomas T. Hoopes Prize for “outstanding scholarly work” by undergraduates.	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
	Harvard University, NLP Group Research Assistant <i>Mentors:</i> Yonatan Belinkov and Alexander Rush <i>Project:</i> Probing for Neurons that Encode and Exploit Negation in BERT	March 2019 – May 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	<ul style="list-style-type: none">[1] Explaining NLP Models via Minimal Contrastive Editing (MiCE). Alexis Ross, Ana Marasović, and Matthew E. Peters. In submission, 2021.[2] Ensuring Actionable Recourse via Adversarial Training. Alexis Ross, Himabindu Lakkaraju, and Osbert Bastani. In submission, 2021.	
PUBLICATIONS	<ul style="list-style-type: none">[3] How well do NLI models capture verb veridicality? Alexis Ross and Ellie Pavlick. <i>EMNLP 2019</i>. Oral presentation.[4] 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. <i>*SEM, 2019</i>. Best Paper Award.	
TEACHING EXPERIENCE	CS 121: Introduction to the Theory of Computation Teaching Fellow, Harvard School of Engineering and Applied Sciences <i>Instructor:</i> Boaz Barak Awarded a Certificate of Distinction in Teaching for a student evaluation of 4.8/5.0.	Sept 2018 – Dec 2018