Alexis J. Ross

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

EDUCATION Harvard University, Cambridge, Massachusetts

Sept 2016 – May 2020

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

RESEARCH EXPERIENCE Allen Institute for AI (AI2), AllenNLP

July 2020 - Present

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

Harvard University, NLP Group

 $March\ 2019-May\ 2019$

Research Assistant

Mentors: Yonatan Belinkov and Alexander Rush

Project: Probing for Neurons that Encode and Exploit Negation in BERT

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

PREPRINTS

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

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

PUBLICATIONS

[2] Learning Models for Actionable Recourse.

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

[3] 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.

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

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

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

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

[6] 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	2019 EM	pes Prize "for extraordinary undergraduate research," awarded for senior thesis NLP Student Travel Scholarship ificate of Distinction in Teaching for a student evaluation of 4.8/5.0	
SERVICE	2021 Co- P: 2021 Sec	organizer and frequent host, NLP Highlights Podcast, over 9000 organizer, AllenNLP Hacks, a hackathon for minoritized students resented workshop: "My First NLP Project" ondary Reviewer, ACL ntor, Harvard Women in Computer Science	
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, Chan 2016 - 2018 2016 - 2020 2014 - 2016 2013 - 2015	Chamber Music Performance at Harvard. Coaches: Parker Qua Juilliard Pre-College, Diploma in Violin Performance. Teacher:	
	Mental Health Advocacy and Counseling, Harvard University 2017 – 2020 Events Co-Chair, Student Mental Health Liaisons 2018 – 2020 Peer Mental Health Counselor, Room 13		2017 - 2020