

(412) 773-2342  
Boston, MA  
akyurek@bu.edu

# Afra Feyza Akyürek

PhD Student in Computer Science

Google Scholar  
Personal Webpage  
GitHub: feyzaakyurek  
LinkedIn: afrafezyaakyurek

I am a 5<sup>th</sup> year PhD student in Computer Science, with a focus on Natural Language Processing. My research interests include improving language model safety, alignment and factuality using tools such as self-training, multi-agent learning, learning from human and machine feedback.

## EDUCATION

<b>PhD in Computer Science</b> , <i>Boston University</i>	September 2019 — May 2024
<b>MSc in Statistics</b> , <i>Carnegie Mellon University</i>	August 2018 — July 2019
<b>BSc in Computer Engineering</b> , <i>Koc University</i> , <i>Ranked 3rd in graduating class.</i>	September 2013 — July 2019

## TECHNICAL EXPERIENCE

<b>Research Intern</b> <i>Allen Institute for Artificial Intelligence (AI2)</i> <ul style="list-style-type: none"><li>Learning with natural language feedback with reinforcement learning.</li><li>Collaborative language models.</li></ul>	<b>September 2022 — December 2022</b> Seattle, WA
<b>Machine Learning Research Intern</b> <i>Apple</i> <ul style="list-style-type: none"><li>Measuring and tackling social biases in large-scale language models.</li><li>Working on the paper titled “Challenges in Measuring Biases in Open-Ended Language Generation”.</li></ul>	<b>June 2021 — August 2021</b> Cambridge, MA
<b>Biostatistics Intern</b> <i>Novartis Pharmaceuticals</i> <ul style="list-style-type: none"><li>Creating an explainable predictive model for treatment efficacy and safety of CAR-T therapy.</li><li>Revealed outstanding findings and communicated them to stakeholders with different backgrounds.</li></ul>	<b>May 2019 — July 2019</b> East Hanover, NJ

## SELECTED PUBLICATIONS

- Afra Feyza Akyürek**, Ekin Akyürek, Leshem Choshen, Derry Wijaya, Jacob Andreas. *Deductive Closure Training of Language Models for Coherence, Accuracy, and Updatability*. **Preprint**. 2024
- Afra Feyza Akyürek**, Eric Pan, Garry Kuwanto, Derry Wijaya. *DUnE: Dataset for Unified Editing*. **EMNLP 2023**. 2023
- Afra Feyza Akyürek**, Ekin Akyürek, Ashwin Kalyan, Peter Clark, Derry Wijaya, Niket Tandon. *RL4F: Generating Natural Language Feedback with Reinforcement Learning for Repairing Model Outputs*. **ACL 2023**. 2023
- Afra Feyza Akyürek**, Muhammed Yusuf Kocyigit, Sejin Paik, and Derry Wijaya. *Challenges in Measuring Bias via Open-Ended Language Generation*. **4th Workshop on Gender Bias in NLP at NAACL 2022 (Oral)**. 2022
- Afra Feyza Akyürek**, Sejin Paik, Muhammed Yusuf Kocyigit, Seda Akbiyik, Şerife Leman Runyun, and Derry Wijaya. *On Measuring Biases in Prompt-Based Learning*. **Findings of NAACL 2022**. 2022
- Afra Feyza Akyürek**, Ekin Akyürek, Derry Wijaya and Jacob Andreas. *Subspace Regularizers for Few-Shot Class Incremental Learning*. **ICLR 2022**. 2022
- Ekin Akyürek, **Afra Feyza Akyürek** and Jacob Andreas. *Learning to Recombine and Resample Data for Compositional Generalization*. **ICLR 2021**. 2021
- Afra Feyza Akyürek**, Lei Guo, Randa Elanwar, Margrit Betke, Prakash Ishwar and Derry T. Wijaya. *Multi-label and Multilingual News Framing Analysis*. **ACL 2020**. 2020

## SKILLS

<b>Tools and Languages</b>	Python, Git, Docker
<b>Quantitative Research</b>	PyTorch, R, MySQL

## TALKS, HONORS AND AWARDS

<b>Hariri Institute for Computing</b> , Graduate Student Fellow, Boston	May 2021
<b>Best Senior Design</b> , Self-Driving Cars in Unity with Deep Reinforcement Learning	June 2018
<b>UNESCO International Symposium on Girls' Education in STEM</b> , Invited Talk	August 2017
<b>31<sup>st</sup> Place in National University Entrance Exam</b> amid 2 million takers in Turkey	2013