

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

Afra Feyza Akyürek

PhD Student in Computer Science

Google Scholar
Personal Webpage
GitHub: feyzaakyurek
LinkedIn: afrafeyzaakyurek

My research focuses on both pre- and post-training aspects of language model development. I work on improving model alignment and reducing hallucinations in LLMs through synthetic data generation, multi-agent learning, and learning from human and machine feedback. Recently, I have explored efficient pre-training of large models using smaller language models. I collaborate closely with Dr. Jacob Andreas at MIT CSAIL and my advisor Dr. Derry Wijaya.

EDUCATION

PhD in Computer Science, *Boston University* September 2019 —
MSc in Statistics, *Carnegie Mellon University* August 2018 — July 2019
BSc in Computer Engineering, *Koc University*, Ranked 3rd in graduating class. September 2014 — July 2018

SELECTED PUBLICATIONS

- **Afra Feyza Akyürek**, Ekin Akyürek, Leshem Choshen, Derry Wijaya, Jacob Andreas. *Deductive Closure Training of Language Models for Coherence, Accuracy, and Updatability*. **Findings of ACL 2024**. 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

TECHNICAL EXPERIENCE

Research Intern September 2022 — December 2022
Allen Institute for Artificial Intelligence (AI2) Seattle, WA

- Designed and implemented a multi-agent language model system with reinforcement learning to optimize model responses based on machine-generated feedback signals (ACL 2023)

Machine Learning Research Intern June 2021 — August 2021
Apple Cambridge, MA

- Evaluating and aligning large language models for social biases (Findings of NAACL 2022)

Statistical Learning Intern May 2019 — July 2019
Novartis Pharmaceuticals East Hanover, NJ

- Developed machine learning model to predict CAR-T therapy outcomes, accurately forecasting treatment efficacy and safety risks; findings presented to VP leadership

SKILLS

Tools and Languages Python, PyTorch, Git, Docker, MLOps tools

TALKS, HONORS AND AWARDS

Hariri Institute for Computing, Graduate Student Fellow, Boston May 2021 - December 2024
Best Senior Design, Self-Driving Cars in Unity with Deep Reinforcement Learning June 2018
UNESCO International Symposium on Girls' Education in STEM, Invited Talk August 2017
31st Place in National University Entrance Exam amid 2 million takers in Turkey 2013