

# Yeganeh Kordi

✉ kordi@brown.edu • 🌐 yeganehkordi.github.io • in yeganehkordi  
🐙 yeganehkordi • 🐦 yeganehkordi

## Research

I am broadly interested in artificial intelligence, machine learning, and natural language processing. My research focuses on AI alignment, data-centric machine learning, and weak-to-strong generalization. I study how to improve language models through better data selection, filtering, and curation, with the goal of reducing the human effort required to make these systems useful.

## Education

- **Brown University**, Providence, RI 2023–Present  
Ph.D. in Computer Science  
Advisor: Stephen Bach
- **Brown University**, Providence, RI 2023–2025  
Sc.M. in Computer Science  
Advisor: Stephen Bach
- **Amirkabir University of Technology**, Tehran, Iran  
Bachelor of Science, Dual Degree
  - Electrical Engineering, Control Systems 2016–2020
  - Computer Engineering, AI and Computer Networks 2018–2021

## Publications

- 🔗 AfRIEpsAAAAJ 🔗 2156538832
- **Tur[k]ingBench: A Challenge Benchmark for Web Agents**  
Kevin Xu, **Yeganeh Kordi**, Tanay Nayak, Adi Asija, Yizhong Wang, Kate Sanders, Adam Byerly, Jingyu Zhang, Benjamin Van Durme, Daniel Khashabi  
NAACL, 2025
  - **Self-Instruct: Aligning Language Models with Self-Generated Instructions**  
Yizhong Wang, **Yeganeh Kordi**, Swaroop Mishra, Alisa Liu, Noah A. Smith, Daniel Khashabi, Hannaneh Hajishirzi  
ACL, 2023  
#1 Most Influential Paper at ACL 2023 by Paper Digest
  - **Super-NaturalInstructions: Generalization via Declarative Instructions on 1600+ Tasks**  
🔹 Yizhong Wang, 🔹 Swaroop Mishra, ♣ Pegah Alipoormolabashi, ♣ **Yeganeh Kordi**, ..., Chitta Baral, Yejin Choi, Noah A. Smith, Hannaneh Hajishirzi, Daniel Khashabi  
EMNLP, 2022  
🔹 Co-first authors; ♣ Co-second authors, alphabetical order.  
#2 Most Influential Paper at EMNLP 2022 by Paper Digest
  - **UnifiedQA-v2: Stronger Generalization via Broader Cross-Format Training**  
Daniel Khashabi, **Yeganeh Kordi**, Hannaneh Hajishirzi  
arXiv preprint arXiv:2202.12359

## Research Experience

- **BATS Lab at Brown University** Sep 2023 – Present  
Advisor: Stephen Bach
  - Researching task difficulty estimation and Item Response Theory to guide data selection and curriculum learning for large language models.
- **Allen Institute for AI and H2Lab at University of Washington** Jul 2021 – Sep 2023  
Mentors: Daniel Khashabi, Yizhong Wang
  - **UnifiedQA-v2**: Created a cross-format QA model with broader supervision, improving generalization.
  - **Super-NaturalInstructions**: Introduced a benchmark of 1,600+ NLP tasks with expert-written instructions

and developed *Tk-INSTRUCT*, a transformer trained to follow in-context instructions.

- **Self-Instruct**: Proposed a self-training pipeline that uses a Pre-trained LM to generate and learn from its synthetic instruction-output pairs.

- o **Center for Language and Speech Processing at Johns Hopkins University**

*Aug 2022 – Dec 2023*

Mentor: Daniel Khashabi

- **Tur[k]ingBench**: Built a benchmark and baseline for instruction-following in HTML-based web environments.

## Teaching Experience

---

- o **Teaching Assistant**, Amirkabir University of Technology

- Machine Learning
- Algorithm Design
- Data Mining
- Cloud Computing
- Digital Logic Circuits

*Fall 2021 – Spring 2023*

*Fall 2021 – Winter 2022*

*Fall 2021*

*Spring 2021*

*Fall 2020 – Fall 2021*

## Invited Talks

---

- o University of Tehran

*Spring 2023*

Self-Instruct: Aligning Language Models with Self-Generated Instructions

- o Sharif University of Technology

*Winter 2024*

Self-Instruct: Aligning Language Models with Self-Generated Instructions

## Mentorship

---

- o Ilana Nguyen, BS student at Brown University

*Sep 2024–Present*

## Professional Service

---

### Reviewer

---

- o Neural Information Processing Systems (NeurIPS)

*2023*

- o International Conference on Learning Representations (ICLR)

*2025*

- o ACL Rolling Review (ARR)

*Feb 2025, May 2025*

### Department and University Service

---

- o Co-organizer of CS PhD Mentorship Program, Brown University

*2024–Present*

- o Research Leader, Summer NSF REU program

*Summer 2025*

- o International Orientation Mentor, Brown University

*Summer 2024*

- o Technical Committee Member, IranOpen RoboCup Competition

*2018*

- o Chair of Education, Student Scientific Association of Electrical Engineering, Amirkabir University

*2018–2019*

## Technical Skills

---

- o **Programming**: Python, Bash, Java, C/C++, Shell scripting
- o **Machine Learning and Deep Learning**: PyTorch, TensorFlow, Scikit-learn, Transformers, CUDA
- o **NLP**: Hugging Face Datasets, Tokenizers, NLTK, SpaCy, lm-eval-harness, vLLM
- o **Data Engineering**: Pandas, NumPy, Matplotlib, Seaborn, JSON, CSV, SQL
- o **Web development**: HTML, CSS, JavaScript, Flask, Django
- o **Other**: Git, Linux CLI, Slurm, LaTeX

## Honors

---

- o Admitted to double-degree program by Talented Student Office, Amirkabir University of Technology.
- o Ranked 19th in Iranian M.Sc. entrance exam among over 5,000 participants.
- o Ranked 148th in Iranian B.Sc. entrance exam among over 70,000 participants.

**📄 References, Further information, and Proofs are available upon Request**