

ALIREZA SALKHORDEH ZIABARI

salkhord@usc.edu \diamond alirezaziabari.github.io \diamond +1(213) 294-6159 \diamond Los Angeles, CA

RESEARCH INTEREST

Natural Language Processing (NLP), Computational Social Science, Robustness and Fairness in NLP, Large Language Models (LLMs).

EDUCATION

Ph.D. in Computer Science, University of Southern California, Los Angeles, CA, U.S. 2021 - Expected 2026
PI: **Morteza Dehghani** [Member of Morality and Language Lab](#)

M.Sc. in Computer Science, University of Southern California, Los Angeles, CA, U.S. 2021 - 2024
Related Courses: Advanced Natural Language Processing, Dynamics Of Representation Learning, Advanced Analysis of Algorithms, Multi-level Modeling, Linear System Theory, Random Processes, Probability and Statistics.

B.Sc. in Computer Engineering, Sharif University of Technology, Tehran, Iran. 2016 - 2020
PI: **Mohammad Hossein Rohban**, GPA: 3.74/4

PUBLICATIONS

- [1] Abdurahman, S., **Ziabari, A.**, Moore, D. A., Bartels, D., Dehghani, M. [Evaluating large language models in psychological research: A guide for reviewers - AMPPS](#).
- [2] Golazizian, P., **Ziabari, A.**, Omrani, A., Dehghani, M. [Cost-Efficient Subjective Task Annotation and Modeling through Few-Shot Annotator Adaptation](#) — EMNLP 2024.
- [3] Hejabi P.*, Rahmati E.*, **Ziabari, A.**, Golazizian, P., Thomason, J., Dehghani, M. [Evaluating Creativity and Deception in Large Language Models: A Simulation Framework for Multi-Agent Balderdash](#) — Wordplay at ACL 2024.
- [4] **Ziabari, A.**, Omrani, A., Hejabi, P., Golazizian, P., Kennedy, B., Piray, P., Dehghani, M. [Reinforced Multiple Instance Selection for Speaker Attribute Prediction](#) — NAACL 2024.
- [5] Omrani, A.*, **Ziabari, A.***, Golazizian, P., Sorensen, J., Dehghani, M. [Towards a Unified Framework for Adaptable Problematic Content Detection via Continual Learning](#) — WOAHA at NAACL 2024.
- [6] Omrani, A., **Ziabari, A.**, Yu, C., Golazizian, P., Kennedy, B., Atari, M., Ji, Heng., Dehghani, M. (2023) [Social-Group-Agnostic Bias Mitigation via the Stereotype Content Model](#) — ACL 2023, Nominated for Best Paper Award (1.6% of accepted papers).

IN PREPARATION & UNDER REVIEW

- [1] **Ziabari, A.**, Ghazizadeh, N., Sourati, Z., Ozcan, Karimi-Malekabadi, F., Piray, P., Dehghani, M. [Reasoning on a Spectrum: Aligning LLMs to System 1 and System 2 Thinking](#) Under review at ACL 2025.
- [2] Rahmati, E.*, **Ziabari, A.***, Dehghani, M. [CoCo-CoLa: Evaluating Language Adherence in Multilingual LLMs](#) Under review at ACL 2025.
- [3] Sourati, Z., Ozcan, Karimi-Malekabadi, F., M., McDaniel, C., **Ziabari, A.**, Wen, N., Tak, A., Morstatter, F. and Dehghani, M. [The Shrinking Landscape of Linguistic Diversity in the Age of Large Language Models](#) - Under review at Nature Human Behavior.
- [4] Karimi-Malekabadi, F., Oyserman D., **Ziabari, A.**, Omrani, A. [What Doesn't Kill Us: Harsh Ecologies Beget Difficulty-as-Improvement Norms](#) - arXiv.
- [5] Trager, J.*, **Ziabari, A.***, Davani, A.M., Golazizian, P., Karimi-Malekabadi, F., Omrani, A., Li, Z., Kennedy, B., Karl Reimer, N., Reyes, M., Cheng, K., Wei, M., Merrifield, C., Khosravi, A., Alvarez, E., Dehghani, M. [The Moral Foundations Reddit Corpus](#) — arXiv.

- [6] Sourati, Z., Ozcan, M., McDaniel, C., **Ziabari, A.**, Wen, N., Tak, A., Morstatter, F. and Dehghani, M. [Secret Keepers: The Impact of LLMs on Linguistic Markers of Personal Traits](#) - arXiv.

WORK & RESEARCH EXPERIENCE

Research Assistant at University of Southern California May. 2022 - Present

- PI: Morteza Dehghani. Member of Morality and Language lab.
- Conduct research on continual learning and bias mitigation, applying Reinforcement Learning (RL) for information extraction; actively investigate representation collapse in Large Language Models (LLM).

Research Assistant at University of Southern California May. 2021 - May. 2022

- PI: Maryam Shanechi. Member of Neural Systems Engineering and Information Processing Lab.
- Designed self-supervised transformer model for mood classification.

Machine Learning Engineer at Tapsi Oct. 2020 - May. 2021

- Led online traffic estimation projects, improving to 5% average ETA compared to Google Maps and Waze.
- Deployed a time-series autoencoder pipeline to estimate online traffic using drivers' locations with Pytorch.

Machine Learning Intern at Fanap Advanced Technology Lab (Zlab) July 2020 - Oct. 2020

- Built Object Detection tool, which led to intelligent grocery store carts.
- Maintained a dataset and trained a Yolo-v5 with Tensorflow.

Research Assistant at Sharif University of Technology Sept. 2018 - Aug. 2020

- PI: Mohammad Hossein Rohban. Member of Robust and Interpretable Machine Learning Lab.
- Developed a Meta-Learning algorithm for anomaly detection in images by Pytorch as my bachelor thesis.

SELECTED SKILLS

Programming Language: Python, R, Java, C++, C, MatLab, SQL, bash.

Software: Pytorch, Tensorflow, Keras, HuggingFace, pySpark, NLTK, Pandas, Numpy, Scikit-learn.

Statistics: Multi-level Modeling

Languages: English, Persian

HONORS & AWARDS

Best Poster Award, Society for Personality and Social Psychology (SPSP) 2023

Ranked **9th** among Computer Engineering Students in Sharif University of Technology. 2020

Ranked **213th** among over 200,000 participants in the Nationwide University Entrance Exam. 2016

Member of the National Organization for Development of Exceptional Talents. 2009

TEACHING EXPERIENCE

Teaching Assistant at **University of Southern California** 2024 - Presents

- Graduate Courses: Information Retrieval and Search Engine (Fall 2024), Analysis of Algorithm (Spring 2025).

Teaching Assistant at **Sharif University of Technology** 2017 - 2020

- Graduate Courses: Deep Learning (Spring 2020, Fall 2019).
- Artificial Intelligence(Spring 2020, Fall 2019), Database Design(Spring 2019), Probability and Statistics(Fall 2018), Advance Programming (Spring 2017).