Elnaz Rahmati

erahmati@usc.edu | LinkedIn | Website | GitHub | Google Scholar

SUMMARY

- Ph.D. Student in Computer Science at USC (3.92 GPA), specializing in NLP, LLMs, and Lifelong Learning.
- Experienced in multilingual and low-resource language modeling, continual learning, and cross-lingual evaluation.
- Author of publications in ACL, EMNLP, LREC.
- Skilled in PyTorch, HuggingFace Transformers, and large-scale model development.

EDUCATION

University of Southern California, Los Angeles, CA — Jan. 2024 – Dec. 2027
Thomas Lord Department of Computer Science, USC Viterbi

Ph.D. in Computer Science (GPA: 3.92/4.00)
Selected Courses: Science of LLMs, History of Language and Computing, Text as Data
Research focus: Natural Language Processing (NLP), Large Language Models (LLM), Lifelong Learning, Multilinguality
Advisor: Dr. Morteza Dehghani

Sharif University of Technology, Tehran, Iran — Sep. 2021 – Jun. 2023

M.Sc. in Computer Engineering (GPA: 3.87/4.00)
Selected Courses: Machine Learning (ML), Deep Learning, Natural Language Processing
Thesis: "Context-based Persian Grapheme-to-Phoneme Conversion using Sequence-to-Sequence Models"
Advisor: Dr. Hossein Sameti

Isfahan University of Technology, Isfahan, Iran — Sep. 2017 – Sep. 2021

B.Sc. in Computer Engineering (GPA: 3.96/4.00)
Selected Courses: Computer Vision, Multimedia Systems, Computational Intelligence
Thesis: "Data Mining on Digikala Online Store Datasets"
Advisor: Dr. Mehran Safayani

RESEARCH EXPERIENCE

University of Southern California, Los Angeles, CA
Research Assistant

Latent Communication in Heterogeneous Multi-Agent Systems **(**May. 2026 – Sep. 2026)

- Developed a training-free latent-space generation method that improves performance over text-based communication while reducing latency.
- Developed both training-free and trained methods for cross-model communication, with the trained approach offering higher performance at increased computational cost.

Assessment of Reciprocity and Liking in Conversations (Sep. 2024 – Oct. 2025)

- Led the computational component of the project, focusing on utilizing conversational features to enhance LLMs in supporting individuals experiencing loneliness, presented in **SPSP 2026**.
- Funded by the Center for Computational Language Sciences (CCLS) seed funding ($3,000).

Honor – Retweet Network Analysis for Detecting Polarization (Jul. 2024 – Aug. 2024)

- Engineered an efficient parallel processing framework to construct and cluster a complex retweet network from millions of Persian tweets during social movements.
- Applied keyword extraction and topic modeling to analyze topic evolution and identify temporal cluster changes over two-week windows.
- Executed dialect/language filtering to ensure data quality and geographic specificity (Iran).

Optical Character Recognition (OCR) for Old Persian Newspapers (Aug. 2023 – Dec. 2023)

- Developed an end-to-end pipeline for digitizing historical texts, including image enhancement to repair faded text and remove imperfections from scanned newspapers.
- Created an unsupervised algorithm leveraging classic image processing (erosion, dilation, edge detection) for accurate text and image box segmentation without labeled training data.
- Trained a Byte-level Language Model to significantly denoise and improve the quality of text extracted via OCR.

Sharif University of Technology, Tehran, Iran
Research Assistant

End-to-End Text-to-Speech (TTS) for Low-Resource Languages (Aug. 2022 – Jun. 2023)

- Impact: Developed the third version of the closed-source Ariana TTS system, accessible via this link.
- Authored and released a high-quality, manually labeled, sentence-level grapheme-to-phoneme (G2P) dataset for Persian, alongside a five-million-sentence machine-generated dataset.
- Trained a novel end-to-end G2P model addressing conversion, Kasre-Ezafe detection, and homograph disambiguation simultaneously (**EMNLP 2024** publication).
- Spearheaded the training of the Tacotron model using phonemes as input to achieve state-of-the-art naturalness in low-resource TTS for Persian.

TECHNICAL SKILLS

- **Areas of Expertise**: NLP, LLMs, ML, Data Analysis, Speech Processing, Computational Social Science.
- **Programming Languages**: Python, C++, C, MATLAB, Verilog.
- **Data/Machine Learning**: PyTorch, HuggingFace Transformers, scikit-learn (SKLearn), OpenCV, SQL.
- **Statistics/Modeling**: Probability Theory, Inferential Statistics, Regression Analysis, Bayesian Methods.
- **Tools & Systems**: Webots, Proteus Design Suite, ModelSim.

LEADERSHIP & TEACHING EXPERIENCE

Internship Mentor, Sharif University of Technology | May. 2022 – Sep. 2022

- Mentored three interns in advanced speech and NLP projects, including optimizing Persian speech synthesis and evaluating state-of-the-art Automatic Speech Recognition (ASR) architectures (e.g., Transformers, Conformers).
- One mentored intern successfully joined the production team of the Ariana TTS system.

Teaching Assistant (Ph.D. & M.Sc. Level) | 2024 – 2026 (USC) / 2022 – 2023 (Sharif) / 2018 – 2019 (Isfahan)

- **USC**: Information Retrieval and Web Search Engines (**Head TA**), Analysis of Algorithms, Introduction to Programming.
- **Sharif**: Natural Language Processing, Speech Processing.
- **Isfahan**: Advanced Programming, Basic Programming.

SELECTED PUBLICATIONS AND PRESENTATIONS

(* indicates equal contribution)

Publications (Preprints & Conference Papers)

1. ***Elnaz Rahmati**, Nona Ghazizadeh, Zhivar Sourati, Nina Rouhani, and Morteza Dehghani. Structural Abstraction as an Inductive Bias for Non-Stationary Language Model Training. Under Review at ICLR 2027.*
2. *Nona Ghazizadeh*, **Elnaz Rahmati***, Alireza Ziabari, Kian Ahrabian, Payam Piray, Morteza Dehghani. DIALECT: Continuous-Space Communication Within and Across Language Models. Under Review at ICLR 2027.*
3. ***Elnaz Rahmati**, Alireza Ziabari, and Morteza Dehghani. CoCo-CoLa: Evaluating and Improving Language Adherence in Multilingual LLMs. In Multilingual Representation Learning (MRL) Workshop at Association for Computational Linguistics: EMNLP 2025.*
4. ***Elnaz Rahmati** and Hossein Sameti. GE2PE: Persian End-to-End Grapheme-to-Phoneme Conversion. In Findings of the Association for Computational Linguistics: EMNLP 2024.*
5. *Preni Golazizian, **Elnaz Rahmati**, ..., Shrikanth Narayanan, Benjamin A.T. Graham, Morteza Dehghani. The Subjectivity of Respect in Police Traffic Stops: Modeling Community Perspectives in Body-Worn Camera Footage. In Proceedings of the 64th Annual Meeting of the Association for Computational Linguistics: ACL 2026.*
6. *Parsa Hejabi*, **Elnaz Rahmati***, Alireza Ziabari, Preni Golazizian, Jesse Thomason, and Morteza Dehghani. Evaluating Creativity and Deception in Large Language Models: A Simulation Framework for Multi-Agent Balderdash. In Wordplay Workshop at Association for Computational Linguistics: ACL 2024.*
7. *Jackson Trager*, Alireza S. Ziabari*, **Elnaz Rahmati**, Aida Mostafazadeh Davani, ..., Morteza Dehghani. The Moral Foundations Reddit Corpus. In Proceedings of the Fifteenth Language Resources and Evaluation Conference (LREC 2026) (pp. 6383–6407). European Language Resources Association (ELRA).*
8. *Parsa Hejabi, **Elnaz Rahmati**, Alireza Ziabari, and Morteza Dehghani. Flip-Flop Consistency: Unsupervised Training for Robustness to Prompt Perturbations in LLMs. In Proceedings of the 64th Annual Meeting of the Association for Computational Linguistics: ACL 2026.*
9. *Sadra Sabouri, **Elnaz Rahmati**, Soroush Gooran, and Hossein Sameti. naab: A ready-to-use plug-and-play corpus for Farsi. In Journal of Artificial Intelligence, Applications and Innovations, 1(2), 1-8.*
10. *Nona Ghazizadeh, **Elnaz Rahmati**, Morteza Dehghani, Payam Piray. Alignment Needs 'Cognitive Control': On The Role of Regularization in LLM Alignment. Accepted at Position Paper Track - Neurips 2026.*

Presentations:

- *Poster presentation: EMNLP 2024, Miami, FL. GE2PE: Persian End-to-End Grapheme-to-Phoneme Conversion.*
- *Poster presentation: EMNLP 2025, Virtual attendance, CoCo-CoLa: Evaluating and Improving Language Adherence in Multilingual LLMs*.
- *Poster presentation: SoCal NLP 2024, San Diego, CA, Challenges in Cross-Lingual Transfer Learning for Offensive Language Detection: Insights from 30 Languages.*
- *Poster presentation: ACL 2024, Virtual attendance, Evaluating Creativity and Deception in Large Language Models: A Simulation Framework for Multi-Agent Balderdash.*

HONORS & AWARDS

- Awarded a $5,000 research grant from the Cognitive Sciences and Technologies Council (Type C Project), Iran, 2023
- Ranked 1st among 66 graduated computer engineering students, Isfahan University of Technology, Iran, 2021
- Ranked 1st among more than 80 computer engineering students, Isfahan University of Technology, Iran, 2020
- Ranked 2nd among more than 20 teams in local ACM-ICPC, Isfahan University of Technology, Iran, 2018
- Fellowship as an Exceptionally Talented student, Isfahan University of Technology, Iran, 2017
- Ranked 193rd among more than 149,000 participants in Iranian Undergraduate Nationwide Entrance Exam, 2017