# Zhivar Sourati

+1 (818) 877 3590 | souratih@usc.edu, zhivarsourati@gmail.com | zhpinkman.github.io | Google Scholar

## Education

**Ph.D. Computer Science | University of Southern California, CA, US | 2022 – Present** (expected: May 2027)
Reasoning in Language Models, Analogical-, Prototype-based Reasoning, Natural Language Processing (NLP), Retrieval-augmented generation (RAG), Mechanistic Interpretability, Computational Social Sciences, and Large Language Models (LLMs) (GPA: 4.0 / 4.0)

**M.Sc. Computer Science | University of Southern California, CA, US | 2022 – 2024**
Focus on Artificial Intelligence, Machine Learning, and Natural Language Processing (GPA: 4.0 / 4.0)

**B.Sc. Computer Engineering | University of Tehran, Tehran, Iran | 2017 - 2021**
Thesis: Comparison of the Performance of Permutation and Randomization Tests on Graphs (GPA: 4.0 / 4.0)

## Research & Work Experience

**Research Assistant | USC / ISI | 2022 – Present**
- Understand, evaluate, and improve LMs' reasoning through cognitively inspired and explainable methods like analogical, prototype-based reasoning, and mechanistic interpretability methods, e.g., utilizing case-based and analogical reasoning for logical fallacy detection, creation of an analogical reasoning benchmark on narratives, improving LMs' robustness using prototype-based reasoning, and using mechanistic interpretability to decipher analogical reasoning in LMs
- Study the effects of LLMs on linguistic and cognitive diversity, including how these models influence personality markers, sociolinguistic variation, and the homogenization of human expression

**GenAI PhD Research Intern | Oracle / Oracle Cloud Infrastructure (OCI) | 2025**
- Worked on Retrieval-augmented generation (RAG) for multimodal visually-rich document understanding

**Research Intern | Genentech | 2024**
- Worked as a knowledge graph and large language models intern on KG-aware biomedical representations

**NLP Research Assistant | Zurich University of Applied Sciences | 2021 – 2022**
- Researched automatic extractive and abstractive summarization techniques using Transformers and knowledge graphs on dialogues and their transcripts
- Conducted hate speech analysis and prediction on Twitter timelines

**Research Assistant | University of Tehran | 2020 – 2021**
- Examined machine learning techniques in social network analysis and non-parametric (i.e., permutation) tests on complex networks (e.g., global trade network) using models like Exponential Random Graph Models (ERGMs)
- Investigated the propagation and prevalence of COVID-related topics on social media (e.g., Twitter) both from a graph and NLP perspective
- Studied different models and concepts in reinforcement learning, such as n-armed bandits, On-/Off-policy methods, and investigated their practical use cases, e.g., analyzing the monetary value of time using Prospect Theory

**NLP Research Intern | TeIAS | 2020**
- Studied SOTA and common models, datasets, and tasks in NLP, such as NER and QA chatbots

- Compiled the [Datasets for Farsi (Persian) Natural Language Processing](#)

# Publications & Presentations

- Sourati, Z., Ziabari, A. S., & Dehghani, M. (2025). The Homogenizing Effect of Large Language Models on Human Expression and Thought. arXiv preprint arXiv:2508.01491. (Under review)

- Sourati, Z., Karimi-Malekabadi, F., Ozcan, M., McDaniel, C., Ziabari, A., Trager, J., ... & Dehghani, M. (2025). The Shrinking Landscape of Linguistic Diversity in the Age of Large Language Models. arXiv preprint arXiv:2502.11266. (Under review)

- Ziabari, A. S., Ghazizadeh, N., Sourati, Z., Karimi-Malekabadi, F., Piray, P., & Dehghani, M. (2025). Reasoning on a Spectrum: Aligning LLMs to System 1 and System 2 Thinking. arXiv preprint arXiv:2502.12470. (Under review)

- Sourati, Z., Ozcan, M., Karimi, F., McDaniel, C., Ziabari, A., Trager, J., Wen, N., Tak, A., Morstatter, F., & Dehghani, M. (2024). Secret Keepers: The Impact of LLMs on Linguistic Markers of Personal Traits. arXiv preprint arXiv:2404.00267.

- Ahrabian, K., Sourati, Z., Sun, K., Zhang, J., Jiang, Y., Morstatter, F., & Pujara, J. (2024, October). The curious case of nonverbal abstract reasoning with multi-modal large language models. In Proceedings of the First Conference on Language Modeling (COLM 2024), Philadelphia, PA, USA.

- Jiang, Y., Sun, K., Sourati, Z., Ahrabian, K., Ma, K., Ilievski, F., & Pujara, J. (2024). Marvel: Multidimensional abstraction and reasoning through visual evaluation and learning. Conference on Neural Information Processing Systems, 37, 46567-46592.

- Sourati, Z., Ilievski, F., Sommerauer, P., & Jiang, Y. (2024). ARN: Analogical Reasoning on Narratives. *Transactions of the Association for Computational Linguistics*, *12*, 1063-1086.

- Deshpande, D., Sourati, Z., Ilievski, F., & Morstatter, F. (2024, June). Contextualizing Argument Quality Assessment with Relevant Knowledge. In Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 2: Short Papers) (pp. 316-326).

- Sourati, Z., Deshpande, D., Ilievski, F., Gashteovski, K., & Saralajew, S. (2024, November). Robust Text Classification: Analyzing Prototype-Based Networks. In Findings of the Association for Computational Linguistics: EMNLP 2024 (pp. 12736-12757).

- Jiang, Y., Ilievski, F., Ma, K., & Sourati, Z. (2023, December). BRAINTEASER: Lateral Thinking Puzzles for Large Language Models. In Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing (pp. 14317-14332).

- Thakur, A. K., Ilievski, F., Sandlin, H. Â., Sourati, Z., Luceri, L., Tommasini, R., & Mermoud, A. (2023). Explainable Classification of Internet Memes. In 17th International Workshop on Neural-Symbolic Learning and Reasoning, NeSy 2023.

- Sourati, Z., Ilievski, F., Sandlin, H. Â., & Mermoud, A. (2023, August). Case-based reasoning with language models for classification of logical fallacies. In Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence (pp. 5188-5196).

- Sourati, Z., Venkatesh, V. P. P., Deshpande, D., Rawlani, H., Ilievski, F., Sandlin, H. Â., & Mermoud, A. (2023). Robust and explainable identification of logical fallacies in natural language arguments. Knowledge-Based Systems, 266, 110418.

- Thakur, A., Ilievski, F., Sandlin, H., Mermoud, A., Sourati, Z., Luceri, L., & Tommasini, R. (2023). Multimodal and Explainable Internet Meme Classification. In AI for Social Good (AI4SG) at AAAI-23. Retrieved from https://amulyayadav.github.io/AI4SG2023/.

- ShabaniMirzaei, T., Chamani, H., Abaskohi, A., Zadeh, Z. S. H., & Bahrak, B. (2023). A large-scale analysis of Persian tweets regarding covid-19 vaccination. Social Network Analysis and Mining, 13(1), 148.

- Sourati, Z., Sabri, N., Chamani, H., & Bahrak, B. (2022). Quantitative analysis of fanfictions' popularity. Social Network Analysis and Mining, 12(1), 42.

- Setayesh, A., Sourati, Z., & Bahrak, B. (2022). Analysis of the global trade network using exponential random graph models. Applied Network Science, 7(1), 38.
- Sourati, Z. (2025, March 8–9). LLMs and the erosion of human variance [Paper presentation]. Program on Negotiation AI Summit, Samberg Conference Center, Massachusetts Institute of Technology, Cambridge, MA, United States. https://www.pon.harvard.edu/teaching-materials-publications/program-on-negotiation-ai-summit/
- Chamani, H., Sourati, Z., & Bahrak, B. (2021, October). An Overview of Regression Methods in Early Prediction of Movie Ratings. In *2021 11th International Conference on Computer Engineering and Knowledge (ICCKE)* (pp. 1-6). IEEE.

# Coursework

- Econometrics
- Neural Networks
- Advanced Natural Language Processing
- Machine Learning
- Advanced Analysis of Algorithms
- Text as Data
- Affective Computing
- The Science of Large Language Models
- Seminar in Computer Science Research
- Seminar in Quantitative Psychology
- Oral Communication Skills for Engineering PhD Students

# Skills

- Python, R, C, C++, SQL, Java, Typescript, SPARQL
- PyTorch, Hugging Face, Weights & Biases, Keras, Stata, Linux, Spring, Angular, Git, MongoDB, Node js., MySQL, Neo4j