

# Mohammadamin Shafiei

University of Milan  
Master's in Human-centered AI  
Milan, Italy

Google Scholar  
Homepage  
LinkedIn

**Education**     **University of Milan**     2023 - Present  
MSc in Human-Centered Artificial Intelligence  
Current GPA: 28.27/30

**Shahid Beheshti University**     2018 - 2023  
BSc in Computer Engineering  
Last two years' GPA: 18.3/20

**Publications**     **Shafiei, M.\***, Li, S.S.\*, Tsvetkov, Y. “*Capturing Implicit Demographic Bias in LLMs on Daily Ethical Dilemmas*”, Under Preparation.

Chang, T.\*, Arnett, C.\*, ..., **Shafiei, M.**, ... “*Global PIQA: Evaluating Physical Commonsense Reasoning Across 100+ Languages and Cultures*”, arXiv:2510.24081

**Shafiei, M.\***, Saffari, H.\*, Pilehvar, M.T., Raganato, A. “*TruthTrap: A Bilingual Benchmark for Evaluating Factually Correct Yet Misleading Information in Question Answering*”, Findings of EACL (2026).

**Shafiei, M.\***, Saffari, H.\*, Moosavi, N.S. “*LingAlign: A Multilingual, Multidimensional Benchmark for Evaluating LLM Alignment Variability*”, AAAI (2026) Submission.

**Shafiei, M.\***, Saffari, H.\*, Moosavi, N.S. “*Framing Bias in Arithmetic Reasoning: How Language and Identity Cues Steer LLM Outputs in Objective Tasks*”, arXiv:2506.03923.

Saffari, H.\*, **Shafiei, M.\***, Zhang, H.\*, Harris, L., Moosavi, N.S. “*Beyond Hate Speech: NLP’s Challenges and Opportunities in Uncovering Dehumanizing Language*”, Proceedings of EMNLP (2025).

Saffari, H.\*, **Shafiei, M.\***, Rooein, D., Nozza, D. “*Measuring Gender Bias in Language Models in Farsi*”, GeBNLP @ ACL (2025).

**Shafiei, M.\***, Saffari, H.\*, Moosavi, N.S. “*MultiHoax: A Dataset of Multi-hop False-Premise Questions*”, Findings of ACL (2025).

Saffari, H.\*, **Shafiei, M.\***, Rooein, D.\*, Pierri, F., Nozza, D. “*Can I introduce my boyfriend to my grandmother? Evaluating Large Language Models’ Capabilities on Iranian Social Norm Classification*”, Findings of NAACL (2025).

**Shafiei, M.\***, Saffari, H.\* “*Not All Jokes Land: Evaluating Large Language Models Understanding of Workplace Humor*”, arXiv:2506.01819.

Atari, M., Chen, Y., Li, Y., Li, S., Saffari, H., **Shafiei, M.**, Bolandi, A., Davani, A. “*PsyEmbedding: Fine-tuning Pre-trained Transformer Models for Quantifying Psychological Similarity in Text*”, ACL 2026 Submission.

Bolandi, A.\*, Saffari, H.\*, **Shafiei, M.\***, Vahidi-Asl, M. “*Software Defect Prediction Using Visualization Techniques and CNNs*”, TBA.

## Research Experience

Research Assistant @ {**MPI, Cardiff, University of Milan-Bicocca**} 2025 - Present

- Working on my thesis, which is around Counterfactual Math Reasoning.
- Advisor(s): **Abhilasha Ravichander, MohammadTaher Pilehvar, Alessandro Raganato**

Research Assistant @ {**Polimi, MilaNLP**} 2024 - Present

- Working Computational Social Science and Cultural Analytics, focusing on social norms and demographic biases in LLMs.
- Advisor(s): **Debora Nozza, Francesco Pierri**

Research Assistant @ {**University of Washington**} 2024 - Present

- Advisor(s): **Stella Li**
- Working on Demographics biases of LLMs, and especially, how reasoning affects these biases.

Research Assistant @ {**University of Sheffield**} 2024 - Present

- Advisor(s): **Nafise Sadat Moosavi**
- Working on different evaluation settings of LLMs, exploring safety, alignment, misinformation, hate-speech, multilingual, Math reasoning, and dehumanization.

Research Assistant @ {**Institute for Research in Fundamental Sciences**} 2022 - 2022

- Implemented Transformers using middle-Level languages, like C and HDL.

Research Assistant @ {**Shahid Beheshti University**} 2019 - 2023

- Developed an end-to-end model for Software defect prediction using CNNs
- Worked on link prediction in dynamic graphs via Siamese Networks

## Work Experience

**AI Engineer** @ Guida e Vai 2026 - Present

Developing AI agents that help people with the driving license tests, providing personalized content for each student.

**Student Tutor** @ University of Milan 2024 - 2025

Assisted students with inquiries regarding program procedures, including thesis preparation, internship requirements, and other curricular and extracurricular activities.

**Machine Learning Engineer** @ Wallex Co. 2023 - 2023

Built an AI financial assistant with ChatGPT using prompt engineering, language detection models, and extractors to help users buy/sell coins, get a summary of the news, etc.

**Software Engineer** @ Behsa Co. 2023 - 2023

Implemented an API and pipeline for real-time KPI tracking using Spring

## Teaching Experience

**Teaching Assistant** @ Shahid Beheshti University

- Embedded and Real-time Systems (**Team Leader**) - 2 Semesters

- Computational Intelligence - 1 Semester
- Fundamentals of Robotics (**Team Leader**) - 1 Semester
- Microprocessors and Assembly Language (**Team Leader**) - 3 Semesters
- Computer Architecture (**Team Leader**) - 2 Semesters
- Logic Circuits - 1 Semester
- Principles of Algorithms - 1 Semester

**Skills**      **Programming Languages:** Python, Java, C/C++, Matlab, x86 and ARM Assembly, SystemVerilog, VHDL  
**Machine Learning Libraries:** Keras, TensorFlow, PyTorch, Pandas, NumPy, Scikit-Learn, LangChain  
**Languages:** Persian (native), English (fluent), Italian (A2)  
**Tools:** Git, MongoDB, Webots

**Selected Courses**      **Graduate Courses:** Sociology of AI (4/4), Principles of Social Psychology for AI Design (4/4), Logic for AI (4/4), Math for AI (4/4), NLP (4/4), Machine Learning (4/4), AI, Ethics, and Law (4/4), Deep Learning - during BSc (4/4)  
**Undergraduate Courses:** Fundamentals of Computer Vision (4/4), Fundamentals of Robotics (4/4), Embedded and Real-time Systems (4/4), Internet Engineering (4/4), Artificial Intelligence and Expert Systems (4/4), Microprocessors and Assembly Language (4/4), Computer Architecture (4/4)

**Professional Services**      Reviewer for WASSA workshop @ **EACL** (2026)  
Program Committee @ **AAAI** (2026)  
Ethics reviewer @ **NeurIPS** (2025)  
Reviewer for SWR, WOA, and GeBNLP workshops @ **ACL** (2025)  
Reviewer for Social Sim, SoLaR, and AIA workshops @ **COLM** (2025)