Shahriar Shayesteh

PhD Student, Pennsylvania State University, State College, USA

RESEARCH INTERESTS

My broad research focus is on **Responsible AI:**

- **Privacy Policy Analysis:** I develop AI tools that extract and analyze privacy policy documents to empower policy-makers in safeguarding user privacy rights (PrivaSeer Project).
- **Tool Use Safety in LLM Agents and LLMs:** Investigating failure modes and robustness challenges in tool-calling and execution behaviour of LLMs and AI agents.

EDUCATION

Pennsylvania State University — Fall 2023 — Fall 2026 (Expected)
Ph.D. In Informatics — GPA: 3.97/4.0
Supervisor: Dr. Shomir Wilson, Dr. Lee Giles
Concentration: PrivaSeer Project

University of Ottawa — Jan 2021 — June 2023
Masters of Computer Science — GPA: 4.0/4.0
Supervisor: Dr. Inkpen
Concentration: Applied Artificial Intelligence

The University of British Columbia — Sep 2017 — May 2020
Non-Degree Program — GPA: 3.73/4.0
Concentration: Computer Science and Statistics

Ferdowsi University of Mashhad — Sep 2010 — Sep 2015
Bachelor of Science, Electrical Engineering — GPA: 14.46/20
Concentration: Telecommunications Engineering

ACADEMIC EXPERIENCE

The Human Language Technologies Lab at Penn State — Fall 2023 — Fall 2026 (Expected)

- Focus: PrivaSeer (Privacy Policy Analysis at Scale)

Mila Institute, Summer School on Responsible AI and Human Rights — June 2023 — June 2023

- Focus: Analyzing Challenges and Potentials in Responsible AI design.

NLP Laboratory University of Ottawa — Jan 2021 — June 2023

- Focus: Mitigating Social Bias in Semi-Supervised Text Classification.

RESEARCH INTERNSHIP EXPERIENCE

Department of Canadian Heritage — Feb 2022 — April 2022

- Research Focus: AI-Driven Qualitative Analysis of Unstructured Text on the Challenges Faced by Canadian Artists During COVID-19.

PUBLICATIONS

Generative Adversarial Learning with Negative Data Augmentation for Semi-Supervised Text Classification.
Shahriar Shayesteh and Diana Inkpen.
Proceedings of the 35th International Florida Artificial Intelligence Research Society Conference (FLAIRS-35), 2022.

SoAC and SoACer: A Sector-Based Corpus and LLM-Based Framework for Sectoral Website Classification.
Shahriar Shayesteh, Mukund Srinath, Lee Matheson, Lu Xian, Sinjoy Saha, Lee Giles, and Shomir Wilson.
Proceedings of the 25th ACM Symposium on Document Engineering (DocEng), 2025.

The PrivaSeer Project: Large-Scale Resources for Analysis of Privacy Policy Text.
Shomir Wilson, Florian Schaub, Lee Matheson, Shahriar Shayesteh, and Lu Xian.
2025 Symposium on Usable Privacy and Security (SOUPS), USENIX.

THESES

Master's Thesis: Social Fairness in Semi-Supervised Toxicity Text Classification
2023, University of Ottawa

Academic Service & Outreach

Reviewer. 5th Workshop on Trustworthy NLP (TrustNLP 2025), NAACL
Association for Computational Linguistics (ACL), May 2025

Graduate Student Panel. IST 197 (Introduction to Research in IST), Penn State University
De, A., Lin, Y., Lucas, J., Schulman, J., & **Shayesth, S.** (2025)

AWARDS

Full Assistantship from the University of Ottawa Jan 2021 — June 2023
For academic excellence and research potential.

Graduate Research Assistantship (GRA) at Penn State University Fall 2023 — Fall 2026 (Expected)
For outstanding research capabilities.

TEACHING ASSISTANT EXPERIENCE

CSI5389 Electronic Commerce Technologies Winter 2023
The University of Ottawa
Professor: Thomas Tran

ITI1121 Introduction To Computing II Fall 2022
The University of Ottawa
Professor: Abdorrahim Bahrami

ITI1121 Introduction To Computing II Winter 2022
The University of Ottawa
Professor: Wail Mardini

ITI1120 Introduction To Computing I Fall 2021
The University of Ottawa
Professor: Mohammad Aljaafreh

CPSC 340 Machine Learning and Data Mining Summer 2021
The University of British Columbia
Professor: Nam Hee Gordon Kim

CPSC 540/440 Advanced Machine Learning Winter 2021
The University of British Columbia
Professor: Mark Schmidt

CPSC 340 Machine Learning and Data Mining (Head TA) Fall 2020
The University of British Columbia
Professor: Frank Wood

CPSC 340 Machine Learning and Data Mining Summer 2020
The University of British Columbia
Professor: Ali Reza Shafaei

CPSC 340 Machine Learning and Data Mining Fall 2019
The University of British Columbia
Professor: Mark Schmidt

SKILLS

- **Programming:** Python, Java, C++, SQL
- **Software:** PyTorch, LangChain, AutoGen, TensorFlow, Scikit-learn

REFERENCES

Prof. Shomir Wilson
Associate Professor, Department of Information Sciences and Technology, Pennsylvania State University, State College, USA
E-mail: shomirwilson@psu.edu