Shramay Palta

University of Maryland, College Park

4108, Brendan Iribe Center for Computer Science and Engineering

spalta@umd.edu | shramay-palta.github.io | shramay-palta | shramay-palta | @PaltaShramay | Shramay Palta

Summary

I am broadly interested in the areas of **Natural Language Processing (NLP)** and **Computational Linguistics**, with a focus on **Explainability**, **Commonsense Reasoning** and **Bias and Fairness in NLP.**

Education

University of Maryland, College Park *College Park, Maryland, USA*

Doctor of Philosophy, Computer Science. **GPA: 3.90/4.00** August 2021 - Present

Advisor: **Professor Rachel Rudinger**

University of Maryland, College Park *College Park, Maryland, USA*

Master of Science, Computer Science. **GPA: 3.90/4.00** August 2021 - May 2023

Advisor: **Professor Rachel Rudinger**

Birla Institute of Technology and Science, Pilani (BITS Pilani) *Pilani, India*

Bachelor of Engineering, Electrical and Electronics Engineering **First Division** August 2017 - May 2021

Thesis Supervisor: **Dr. Ashok Agrawala (UMD)** and **Dr. Navneet Gupta (BITS).**

Research and Work Experience

Computational Linguistics and Information Processing (CLIP) Lab *College Park, Maryland, USA*

Graduate Student under Prof. Rachel Rudinger November 2021 - Present

- Using **explainability** techniques to determine the impact of model rationales on humans.
- Studying whether **Large Language Models** can perform **Process of Elimination** on downstream tasks like MCQs for commonsense reasoning.
- Investigating different forms of **biases** in **NLP models and datasets.**

Human-Data Interaction Group, University of Maryland *College Park, Maryland, USA*

Graduate Research Assistant under Prof. Leo Zhicheng Liu September 2021 - March 2022

- Using **Natural Language Processing** Techniques to harvest design feedback from visualization comments on social media platforms like Reddit.

Maryland Information and Network Dynamics (MIND) Lab, University of Maryland *College Park, Maryland, USA*

Research Assistant under Prof. Ashok Agrawala (Undergraduate Research Thesis) May 2020 - February 2021

- Analyzing the spread of **COVID-19 and Flu virus** on campus using location and breathing data collected from **Spire Tags**.

Global Health Centre, Graduate Institute of International and Development Studies *Geneva, Switzerland*

Research Intern under Dr. Amandeep Gill, Executive Director, UNSG's Panel on Digital Cooperation May 2020 - October 2020

- Researched the role of **micro-narratives** as proxy variables to fill in missing data, and to develop **human-centered benchmarks** for **digital health** and used **natural language techniques** to study the social, health, and mental impacts of the **COVID-19** pandemic.

TurnoutNow LLC *Lancaster, Pennsylvania, USA*

Data Science Intern May 2019-July 2019

- Using **real-time location data** from **IoT BLE Beacons** and **natural language generation** tools with live data connections to generate narratives for end users.

Publications

Measuring Style Similarity in Diffusion Models – *Arxiv*

Gowthami Somepalli, Anubhav Gupta, Kamal Gupta, **Shramay Palta**, Micah Goldblum, Jonas Geiping, Abhinav Shrivastava and Tom Goldstein

It's Not Easy Being Wrong: Evaluating Process of Elimination Reasoning in Large Language Models – *Arxiv*

Nishant Balepur, **Shramay Palta** and Rachel Rudinger

Fork: A Bite-Sized Test Set for Probing Culinary Cultural Biases in Commonsense Reasoning Models – *Findings of the 61st Conference of the Association for Computational Linguistics (ACL 2023)*

Shramay Palta and Rachel Rudinger

Investigating Information Inconsistency in Multilingual Open-Domain Question Answering – *Arxiv*

Shramay Palta, Haozhe An, Yifan Yang, Shuaiyi Huang and Maharshi Gor

Activities

Reviewer: ACL 2023, EACL 2023, EMNLP 2022

Member: CS Department Council, Graduate Admissions Committee, Graduate Student Committee (CS GradCo)

Projects

Rationalizing Commonsense Reasoning Questions
College Park, Maryland

Collaboration between UMD and AI2 — January 2023 - Present

- Determining whether the **plausibility ratings** of questions and options in isolation are predictive of the gold label.
- Picking questions from **commonsense reasoning datasets** like Social-IQa and CommonsenseQA and using different **explainability** methods to rationalize such questions.
- Presenting such questions and rationales to human annotators and determining the impact of model-generated rationales on humans to determine cases of **over-trust** and **over-reliance**.

Process of Elimination with Large Language Models
College Park, Maryland

Work with Prof. Rachel Rudinger — August 2023 - December 2023

- Benchmarking multiple **LLMs** to determine whether they can perform **Process of Elimination** to pick incorrect answers in a downstream task like MCQs for Commonsense Reasoning.

Using Explainability Methods for Human-in-the-loop Adversarial Question Dataset Generation
College Park, Maryland

Work with Prof. Jordan Boyd-Graber and Prof. Hal Daumé — May 2023- Present

- Using **Explainability** methods to aid users to write questions and generate **adversarial fact checking datasets**.

Culinary Cultural Biases in Commonsense Reasoning
College Park, Maryland

Work with Prof. Rachel Rudinger — November 2021- January 2023

- Investigated modern-day **commonsense reasoning** NLP Models and datasets like **CommonsenseQA** to determine if they have an implicit or explicit **cultural bias** baked into them.
- Introduced a new dataset, **Fork**, consisting of questions that can be used to stress test models using examples of cultural and social norms, and material and physical differences revolving around different culinary cultural practices and related customs. (Accepted to **Findings of ACL 2023** and presented at the **Third Workshop on Trustworthy Natural Language Processing.**)
- Demonstrated that models have **systematic cultural biases** aligned with **US** over non-US cultures.

Skills

Languages	Python, R, SQL, Linux/Unix shell, Java, C++, C, Assembly Language.
Tools	Pandas, NumPy, NLTK, spacy, Transformers, Keras, TensorFlow, PyTorch, scikit-learn, Matplotlib, Jupyter, Git, LaTeX, MATLAB, MySQL.
Key Courses	Explainable Natural language Processing, CommonSense Reasoning and Natural Language Understanding, Natural Language Processing, How and Why AI Answers Questions, Human AI Interaction, Advanced Numerical Optimization.
OS	MacOS, Linux, Windows, FreeBSD.

Achievements

2022	**Dean's Fellowship Award**: Awarded the Graduate School Dean's Fellowship Award for outstanding academic achievement.
2021	**Dean's Fellowship and Chair's Fellowship Award**: Awarded the Graduate School Dean's Fellowship and the Chair's Fellowship Award for outstanding academic achievement.
2010	**World Robot Olympiad**: Represented India at the World Robot Olympiad held in Manila, Philippines, and secured a **world rank of 31** in my category.
2010	**Indian Robot Olympiad**: Awarded the **1st Runners Up Award** in my category for the north chapter.