Mukund Srinath

State College, PA

mukund.srina@gmail.com — mukundsrinath.github.io

Education

Pennsylvania State University	*Expected: May, 2024*
PhD, Informatics	GPA: 3.85
Co-supervised: C. Lee Giles and Shomir Wilson	

Pennsylvania State University	*August, 2017 - May, 2019*
MS, Information Sciences and Technology	GPA: 3.88

R. V. College of Engineering, India	*August, 2011 - May, 2015*
BE, Electrical and Electronics Engineering	GPA: 8.94

Experience

Human Language Technologies Lab — State College, PA
Research Assistant — *August, 2019 - Present*

- Designed, created and maintain, the world's 1st privacy policy search engine with semantic search functionality.
- Pretrained BERT based language model for the privacy policy domain. HuggingFace
- Developed a speciality crawler and collected the largest corpus of privacy policies containing ∼4 million policies.
- Achieved SoTA results on semantic similarity, named entity recognition (NER) and question answering tasks.

Intelligent Information Systems Lab — State College, PA
Machine Learning Researcher — *August, 2019 - Present*

- Applying responsible AI principles in analyzing and mitigating biases in machine learning models.
- Designed ethics sheet for sentiment analysis after surveying 189 papers.
- Designed model agnostic bias detection technique and discovered instances of implicit and explicit bias in a wide range of sociotechnical and AI as a service (AIaaS) models.

Trustpage — Detroit, MI
Machine Learning Intern — *May, 2022 - August, 2022*

- Designed evaluation metrics to track performance of company-wide suite of machine learning models.
- Implemented and deployed a question similarity model on GCP, improving performance by 12%.
- Applied feature engineering and synthetic data generation techniques on low-resource tasks.

McFadyen Digital — Bangalore, India
Software Engineer — *July, 2015 - July, 2017*

- Implemented a full-life-cycle webapp of a human resource information system supporting ∼500 users.
- Developed backend and frontend online retail features for search, payment and checkout functionality.

Awards and Honors

- **Outstanding Paper** @EMNLP'23
- **Best Student Paper** @DocEng'23
- **Best Short Paper** @TrustNLP Workshop (ACL'23)
- **Best Student Paper** @PETS'23
- **Semi-finalist** @Nittany AI Challenge 2023, for ClassCollab: Peer-learning based adaptive learning system.
- **Semi-finalist** @Nittany AI Challenge 2023, for ChronoNews.ai: AI generated timeline of news articles.
- **2nd place** @HackPSU 2019, for *Beyond Tweet*, an application that combines geospatial census data with tweet sentiment to help make market predictions.
- **2nd place and $20,000 in funding** @Nittany AI Challenge 2018, for *ProFound: A professor search engine.*
- **3rd place** @HackPSU 2018, for *WeatherOrNot*, a travel application to recommend vacation destinations based on travel prices, and weather preferences.
- **1st place and honorable mention by IBM** @HackPSU 2017, for *FindViser* an application that suggests research advisers to graduate students.

Additional Projects

News Timeline Summarization [*Python, PyTorch, scikit-learn, Elasticsearch*]
Objective: Given a news article, create an explainable timeline of news story unfolding over time.
Techniques: *Recommendation system, explainable AI, clustering, document similarity, vector search*

Generative AI Language Understanding Benchmark [*Python, TensorFlow*]
Objective: To evaluate large language models (LLMs) and generative AI models (GPT, Llama, Claude) on privacy policy language understanding tasks.
Techniques: *GenAI fine-tuning, zero-shot and few-shot prompt engineering*

Explainable and Fair Model Compression [Python, PyTorch]
Objective: Mitigate biases by creating fair distilled models through counterfactual generation.
Techniques: *XAI, knowledge distillation, statistical analysis, visualization*

Skills

Machine Learning Skills: sklearn, Tensorflow, PyTorch, R. Experience with multi-GPU processing.

Development Skills: Python, Java, C++, ElasticSearch, Django, SQL. Experience with Git and multiple OS.

Selected Publications

Mukund Srinath, Pranav Venkit, Maria Badillo, Florian Schaub, C. Lee Giles, Shomir Wilson
Semantic Similarity Matching for Data Practice Identification in Privacy Policies (Under Review)
European Chapter of the Association for Computational Linguistics (**ACL 2024**)

Mukund Srinath*, Pranav Venkit*, Sanjana Gautam, Vipul Gupta, Rebecca Passonneau, Shomir Wilson
The Sentiment Problem: A Critical Survey towards Deconstructing Sentiment Analysis
[**Outstanding Paper**] Empirical Methods in Natural Language Processing (**EMNLP 2023**)

Mukund Srinath, Lee Matheson, Pranav Venkit, Gabriela Fortuna, Florian Schaub, Lee Giles, Shomir Wilson
Privacy Now or Never: Large-Scale Extraction and Analysis of Dates in Privacy Policy Text
[**Best Paper**] The 23rd ACM Symposium on Document Engineering (**DocEng, 2023**)

Mukund Srinath, Soundarya Sundareswara, C. Lee Giles, Shomir Wilosn
Privacy Lost and Found: An Investigation at Scale of Web Privacy Policy Availability
The 23rd ACM Symposium on Document Engineering (**DocEng, 2023**)

Pranav Venkit, **Mukund Srinath**, Shomir Wilson
Automated Ableism: An Exploration of Explicit Disability Biases in Sentiment and Toxicity Analysis Models
[**Best Paper**] 3rd Workshop on Trustworthy Natural Language Processing (**ACL, 2023**)

Abraham Mhaidli, Selin Fidan, An Doan, **Mukund Srinath**, Lee Matheson, Shomir Wilson, Florian Schaub
Researchers' Experiences in Analyzing Privacy Policies: Challenges and Opportunities
[**Best Paper**] Proceedings on Privacy Enhancing Technologies (**PETS, 2023**)

Pranav Venkit, **Mukund Srinath**, Shomir Wilson
A Study of Implicit Language Model Bias Against People With Disabilities
29th International Conference on Computational Linguistics (**COLING, 2022**)

Mukund Srinath, Shomir Wilson, C. Lee Giles.
Privacy at Scale: Introducing the PrivaSeer Corpus of Web Privacy Policies.
59th Annual Meeting of the Association for Computational Linguistics (**ACL, 2021**)

Mukund Srinath, Soundarya Sundareswara, Shomir Wilson, C. Lee Giles.
PrivaSeer: A Privacy Policy Search Engine.
International Conference on Web Engineering (**ICWE, 2021**)

News and Media Mentions

Penn State News:	Most websites do not publish privacy policies, researchers say	2023
Penn State News:	Trained AI models exhibit learned disability bias, IST researchers say	2023
The Hill:	Common AI language models show bias against people with disabilities: study	2022
El Pais:	Latin texts and unaffordable lengths: the revelations of a privacy policy search engine	2021
Penn State News:	Search engine could help researchers scour internet for privacy documents	2021

The Collegian: HackPSU winning website aims to combat challenge of finding an adviser   2018
Penn State News: 2018 Nittany AI Challenge winners provide EdTech solutions, receive $50,000   2018

Academic Service

Reviewer: ACL 2023, EMNLP 2023, EMNLP 2023 (Industry track), ACL 2022, EMNLP 2022, ACL 2021, EMNLP 2021

Student Volunteer: FAccT 2023, ACL 2021

Vice President, Graduate Student Association, Penn State

Teaching Experience

Guest Instructor, Integration of Privacy and Security	Fall 2023
Guest Instructor, Foundations of Deep Learning	Fall 2023
Guest Instructor, Information Retrieval and Search Engines	Fall 2022
Guest Instructor, Information Retrieval and Search Engines	Fall 2021
Teaching Assistant, Application Development Design Studio II	Spring 2018
Teaching Assistant, Application Development Design Studio II	Fall 2017