

# Mukund Srinath

E345 Westgate Building, University Park, PA 16802  
[mukund@psu.edu](mailto:mukund@psu.edu) ♦ +1 814-826-5823 ♦ [mukundsrinath.github.io](https://mukundsrinath.github.io)

## Education

---

<b>Pennsylvania State University</b> PhD, Informatics Co-supervised by C. Lee Giles and Shomir Wilson	<i>Expected: May, 2024</i> GPA: 3.85
<b>Pennsylvania State University</b> MS, Information Sciences and Technology	<i>August, 2017 - May, 2019</i> GPA: 3.88
<b>R. V. College of Engineering, India</b> BE, Electrical and Electronics Engineering	<i>August, 2011 - May, 2015</i> GPA: 8.94

## Work Experience

---

Trustpage <b>Machine Learning Intern</b>	Detroit, MI <i>May, 2022 - August, 2022</i>
<ul style="list-style-type: none"><li>Designed evaluation metrics to track performance of company-wide suite of machine learning models.</li><li>Implemented and <b>deployed</b> a question similarity model on GCP, improving performance by 12%.</li><li>Applied feature engineering and synthetic data generation techniques on a task with few training samples.</li><li>Created result visualizations and communicated them to non-technical teams after aligning on business objectives.</li></ul>	
ClearAvenue <b>Software Engineering Intern</b>	Columbia, MD <i>May, 2018 - August, 2018</i>
<ul style="list-style-type: none"><li>Designed a proof-of-concept to improve medical device security in software defined networks.</li><li>Implemented <b>unsupervised anomaly detection</b> methods to detect potential network security threats.</li></ul>	
McFadyen Digital <b>Software Engineer</b>	Bangalore, India <i>July, 2015 - July, 2017</i>
<ul style="list-style-type: none"><li>Implemented a full-life-cycle webapp of a human resource information system supporting ~500 users.</li><li>Developed backend and frontend online retail features for search, payment and checkout functionality.</li></ul>	

## Research Experience

---

Human Language Technologies Lab <b>Research Assistant</b>	Pennsylvania State University <i>August, 2019 - Present</i>
<ul style="list-style-type: none"><li>Applying <b>information extraction</b> and <b>information retrieval</b> techniques to analyze privacy policies at scale.</li><li>Created a novel crawling technique to collect the largest corpus of privacy policies containing over 3.5M policies.</li><li>Designed, created and maintain, the world's first privacy policy search engine, including, creating custom ranking function, and statistical analysis of results and user engagement.</li><li>Achieved state of the art results on classification, named entity recognition (NER) and question answering tasks on privacy policies by creating PrivBERT, a privacy language model.</li></ul>	
Intelligent Information Systems Lab <b>Research Assistant</b>	Pennsylvania State University <i>August, 2019 - Present</i>
<ul style="list-style-type: none"><li>Applying <b>responsible AI</b> principles in analyzing and mitigating biases in machine learning models.</li><li>Designed recommendations for ethical, transparent, and socially responsible sentiment analysis.</li><li>Discovered bias in the use of generative large language models (LLMs) for synthetic dataset annotation.</li></ul>	
PIKE Group <b>Researcher</b>	Pennsylvania State University <i>August, 2017 - July, 2019</i>
<ul style="list-style-type: none"><li>Designed a novel technique to select best answers from a list of non-fact-based questions by generating questions.</li><li>Implemented a convolutional neural network (CNN) based algorithm using TensorFlow to rank answers achieving a precision@1 score of 0.452 and mean reciprocal rank of 0.59, thereby setting a baseline for a novel dataset.</li></ul>	

## Awards

---

- **Semi-finalist** in Nittany AI Challenge, 2023 for [ChronoNews.ai](#) an application that creates a timeline of news articles.
- **2nd place** in 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** in Nittany AI Challenge, 2018 for [ProFound: A professor search engine](#).
- **3rd place** in 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** at HackPSU, 2017 for [FindViser](#) an application that suggests research advisers to graduate students.

## Additional Projects

---

### News Timeline Summarization

- Designed a novel tool that captures news from curated sources and organizes it in the form of a timeline for improved sense-making as a means to combat fake news.
- Developed the document clustering, timeline creation, title text generation and subjectivity detection modules using PyTorch.

### Personal Information Type Extraction

- Designed a novel scalable technique to extract the types of personal information collected from users from over 1.4 million privacy policies. Trained a BERT based model to identify valid personal information types given inputs of noun phrases and their POS tags.

### ProFound: Professor Search Engine

- Designed an end-to-end system to track professors based on research topics and interests. Crawled, indexed and ranked ~500 academic websites.

## Skills

---

<b>Programming Languages</b>	Python, Java, Javascript, C++
<b>Frameworks and Packages</b>	TensorFlow, PyTorch, scikit-learn, ElasticSearch, Scrappy, Django, SQL

## Publications

---

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**)

**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*  
The 23rd ACM Symposium on Document Engineering (**DocEng, 2023**) [In Press]

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

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**)