

Siyi Liu

Email: siyiliu@seas.upenn.edu

Webpage: <https://liusiyi641.github.io/>

Research Interests

My primary research interest lies within the areas of natural language processing. Currently, I'm interested in understanding the implications of news articles (e.g. arguments/frames), bias in NLP, vision and language, text representations, and NLP for social good. I'm also open to other areas in NLP and deep learning.

Education

| | |
|--|-----------------------------|
| University of Pennsylvania , Philadelphia, PA Ph.D. in Computer and Information Science Advisor: Dan Roth | Fall 2022 - May 2027 |
| University of Pennsylvania , Philadelphia, PA M.S.E. in Computer and Information Science | Fall 2020 - May 2022 |
| Boston University , Boston, MA B.A./M.S. Dual Degree in Computer Science | Fall 2016 - May 2020 |

Publications

- [1] Sihao Chen*, **Siyi Liu***, Xander Uyttendaele, Yi Zhang, William Bruno, Dan Roth
“Design Challenges for a Multi-Perspective Search Engine”. in Findings of NAACL22
[\[Paper\]](#)
- [2] Weiqiu You, **Siyi Liu**, Mark Yatskar
“Zero-shot Image Classification with Pre-trained Text Representations” (In progress)
- [3] **Siyi Liu**, Sihao Chen, Xander Uyttendaele, Dan Roth
“MultiOpEd: A Corpus of Multi-Perspective News Editorials”, in NAACL 2021.
[\[Paper\]](#)[\[Slides/Talk/Poster\]](#)
- [4] **Siyi Liu**, Lei Guo, Kate Mays, Margrit Betke and Derry Tanti Wijaya.
“Detecting Frames in News Headlines and Its Application to Analyzing News Framing Trends Surrounding U.S. Gun Violence.” In *Proceedings of Conference on Computational Natural Language Learning (CoNLL)* 2019.
[\[Paper\]](#) [\[Slides\]](#)
- [5] **Siyi Liu***, Ziang Leng*, Derry Tanti Wijaya.
“Learning To Mirror Speaking Styles Incrementally.” on arXiv.
[\[Paper\]](#)

Demo

A Multi-Perspective Search Engine (In progress)

We build a system that presents search results in clusters of perspectives using Google Search API. Empirically, 63% of the participants in our survey prefer our system's organization to naive Google Search's given the same search results and UI.

[\[Website\]](#)

Research Experience

University of Pennsylvania

Summer 2020 – Present

Advisor: Prof. Dan Roth

Project: Information Pollution Project

- *A dataset with hierarchical structure of query, perspectives, evidence, article*
- *Study and analyze the argumentation structure of news articles*
- *Develop a demo website that searches in real time in a perspective-oriented manner*
- *Argument similarity with sentence compression*

University of Pennsylvania

Fall 2021 – Present

Advisor: Prof. Mark Yatskar

Project: Zero-shot Image Classification with Pre-trained Text Representations

- *Pretrain language models on Wikipedia articles and produce text representations as side-information*
- *Align image and text representations for zero-shot image classification*

Boston University

Spring 2019 – Fall 2020

Advisor: Prof. Derry Wijaya

Projects:

- *Detect frames in news headlines and analyze news framing trends surrounding U.S. gun violence*
- *Capture and transform speaking styles in a low resource setting*
- *Create multimedia summaries using tweets and videos*

City University of Hong Kong

Aug 2019 - Sep 2019

Advisor: Prof. Junhui Wang

Project: Analyze Chinese law documents and cluster similar case descriptions

Awards

Undergraduate Research Opportunity Program Award, Boston University

Fall 2019

Dean's List, College of Arts and Sciences, Boston University

Spring (2018 - 2020)

Teaching

Computational Linguistics (CIS530) at Penn

Teaching Assistant

Skills

Programming Languages: Python (most proficient), Java, C/C++

Tools: Huggingface Transformers, Pytorch, Tensorflow, nltk, sklearn

Services

Reviewers: AAI 2021 (secondary)