

Shijia Liu

liu.shij@northeastern.edu

Education

Northeastern University, Ph.D. in Computer Science

Boston, MA

Advisor: David A. Smith

August 2019 - May 2025 (Expected)

Specialization: Natural Language Processing, Digital Humanities

GPA: 3.96/4.00

Johns Hopkins University, M.S. in Computer Science

Baltimore, MD

Specialization: Natural Language Processing, Machine Learning

August 2017 - May 2019

GPA: 3.87/4.00

Stanford University, M.S. in Electrical Engineering

Stanford, CA

Specialization: Optimization, Information Theory

September 2012 - June 2014

GPA: 3.72/4.00

UCLA, B.S. in Electrical Engineering (Summa Cum Laude)

UCLA, B.S. in Physics (Summa Cum Laude)

Los Angeles, CA

Specialization: Signal Processing, Communication Systems

September 2008 - June 2012

GPA: 3.90/4.00

Publications

Detecting de minimis Code-Switching in Historical German Books. Shijia Liu, David A. Smith. In the *Proceedings of the 28th International Conference on Computational Linguistics (COLING 2020)* [[ACL](#)]

Measuring the Similarity of Grammatical Gender Systems by Comparing Partitions. Arya D. McCarthy, Adina Williams, Shijia Liu, David Yarowsky, Ryan Cotterell. In the *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP 2020)* [[ACL](#)]

On the Idiosyncrasies of the Mandarin Chinese Classifier System. Shijia Liu, Hongyuan Mei, Adina Williams and Ryan Cotterell. In the *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics (NAACL-HLT 2019)*. [[ACL](#)]

Publications in Preparation

Detecting Reader Annotations in Printed Books with Self-Training and Posterior Adjustment. Shijia Liu, David A. Smith

New Datasets to Benchmark Methods for Semi-Automated Biomedical Abstract Screening. Shijia Liu, Thomas Trikalinos, Iain J. Marshall, Byron C. Wallace

Research In Progress

Language Models for Brain-Computer Interfaces

- Developing word and character level transformer-based language models for typing predictions in BCI systems.
- Using Huggingface API and Pytorch for implementations.

Information Cascades Modeling

- Predict the propagation of news coverage as time-dependent sequences in historical newspapers

Industry Experience

Applied Scientist Intern at Amazon

Alexa AI Group

Cambridge, MA

June 2022 - September 2022

- Explored continual learning with noisy labeled data for natural language understanding tasks.
- Currently preparing a paper submission for the work done.

Financial Software Developer at Bloomberg L.P.

Electronic Order Routing Group

New York, NY

Oct 2014 - June 2017

- Designed and developed backend applications for the company's next-generation real-time electronic trading platform.
- Over ~100,000 trading orders from ~20,000 users flowing through the applications every day.
- Used C++, C and Python in software development. Programmed in multithreading environment.

Software Engineer II Intern at Cisco Systems, Inc

Data Center Group

San Jose, CA

June 2013 - August 2013

- Developed Python scripts to automatically diagnose and extract critical information from core dump files. This helped accelerate the testing process for the QA team.
- Developed Python scripts to automatically generate diagrams of UCS server clusters.

Teaching Experience

Teaching Assistant for Programming in C++ at Northeastern University

Instructor: Dr. Vidoje Mihajlovikj

Boston, MA

Spring 2020

Course Assistant for Artificial Intelligence at JHU

Instructor: Dr. Philipp Koehn

Baltimore, MD

Spring 2019

Course Assistant for Natural Language Processing at JHU

Instructor: Dr. Jason Eisner

Baltimore, MD

Fall 2018

Course Assistant for Intermediate Programming at JHU

Instructors: Dr. Sara More, Dr. Ben Langmead, Dr. Misha Kazhdan

Baltimore, MD

Fall 2017 - Spring 2018

Honors and Awards

Tau Beta Pi Member

2010

Champion Team Member, UCLA Microsoft Programming Contest

2009

Skills

- Programming Languages: Python, C++, C, Java, R
- Data Analysis: Matlab, SQL, Excel
- Operating Systems: Linux, Windows
- Deep Learning Framework: PyTorch, Huggingface