

# Shijia Liu

liu.shij@northeastern.edu

## Education

---

<b>Northeastern University, Ph.D. in Computer Science</b> <i>Advisor: David A. Smith</i> <i>Specialization: Natural Language Processing, Digital Humanities</i> <i>GPA: 3.97/4.00</i>	<b>Boston, MA</b> <i>August 2019 - December 2025 (Expected)</i>
<b>Johns Hopkins University, M.S. in Computer Science</b> <i>Specialization: Natural Language Processing, Machine Learning</i> <i>GPA: 3.87/4.00</i>	<b>Baltimore, MD</b> <i>August 2017 - May 2019</i>
<b>Stanford University, M.S. in Electrical Engineering</b> <i>Specialization: Optimization, Information Theory</i> <i>GPA: 3.72/4.00</i>	<b>Stanford, CA</b> <i>September 2012 - June 2014</i>
<b>UCLA, B.S. in Electrical Engineering (Summa Cum Laude)</b> <b>UCLA, B.S. in Physics (Summa Cum Laude)</b> <i>Specialization: Signal Processing, Communication Systems</i> <i>GPA: 3.90/4.00</i>	<b>Los Angeles, CA</b> <i>September 2008 - June 2012</i>

## Publications

- 
- Through the Lens of History: Methods for Analyzing Temporal Variation in Content and Framing of State-run Chinese Newspapers.** Shijia Liu, David A. Smith. In the *Proceedings of the 2025 Annual Conference of the Nations of the Americas Chapter of the Association for Computational Linguistics (NAACL 2025)*
- Tracing Accounts of Racial Terror in Historical Newspapers (abstract).** Shijia Liu, David A. Smith. In *New Directions in Analyzing Text as Data 2023 (TADA 2023)*
- Adapting Transformer Language Models for Predictive Typing in Brain-Computer Interfaces.** Shijia Liu, David A. Smith. (2023) [[arXiv](#)]
- 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](#)]

## Industry Experience

---

### Applied Scientist Intern at Amazon

Cambridge, MA

*Alexa AI Group*

*June 2022 - September 2022*

- Explored continual learning with noisy labeled data for natural language understanding tasks.
- Explored variations of EWC and self-replay methods.

### Financial Software Developer at Bloomberg L.P.

New York, NY

*Electronic Order Routing Group*

*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

San Jose, CA

*Data Center Group*

*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

Boston, MA

*Instructor: Dr. Vidoje Mihajlovikj*

*Spring 2020*

### Course Assistant for Artificial Intelligence at JHU

Baltimore, MD

*Instructor: Dr. Philipp Koehn*

*Spring 2019*

### Course Assistant for Natural Language Processing at JHU

Baltimore, MD

*Instructor: Dr. Jason Eisner*

*Fall 2018*

### Course Assistant for Intermediate Programming at JHU

Baltimore, MD

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

*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: Pandas, Matlab, SQL, Excel
- Operating Systems: Linux, Windows
- Deep Learning Framework: PyTorch, Huggingface, Scikit-Learn