

# Carolina Zheng

PhD Candidate · Machine Learning · Natural Language Processing  
carozheng@cs.columbia.edu · [www.carolinazheng.com](http://www.carolinazheng.com)

## Education

- 2019 – Present **Columbia University**  
PhD Computer Science  
Advisor: [David Blei](#)  
Selected Courses: Optimization for ML, Foundations of Graphical Models, Representation Learning
- 2017 – 2019 **University of Pennsylvania**  
MSE Data Science  
Selected Courses: Mathematical Statistics, Machine Learning, Software Systems, Advanced Topics in Algorithms & Complexity, Real Analysis
- 2014 – 2019 **University of Pennsylvania**  
BSE Computer Science (SEAS), BS Economics (Wharton School), *summa cum laude*

## Work Experience

- 2021 Research Intern, **Facebook Applied AI Research**  
Explore contrastive learning methods in ads recommendation models for sparse, high-dimensional feature sets.
- 2019 Research Intern, **Columbia Business School**  
Collect data and estimate causal model for viewership utility based on measures of suspense and surprise for a rational Bayesian viewer.

## Publications

- 2023 **Revisiting Topic-Guided Language Models**  
C. Zheng, K. Vafa, D. Blei  
*Transactions on Machine Learning Research*. [\[paper\]](#) [\[code\]](#)
- 2023 **An Invariant Learning Characterization of Controlled Text Generation**  
C. Zheng\*, C. Shi\*, K. Vafa, A. Feder, D. Blei (\*equal contribution)  
*Proceedings of ACL*. [\[paper\]](#) [\[code\]](#)
- 2022 **Do Suspense and Surprise Drive Entertainment Demand? Evidence from Twitch.tv**  
A. Simonov, R. Ursu, C. Zheng  
*Journal of Marketing Research*. [\[paper\]](#)

- 2019 **Complexity-Weighted Loss and Diverse Reranking for Sentence Simplification**  
R. Kriz, J. Sedoc, M. Apidianaki, C. Zheng, G. Kumar, E. Miltsakaki, C. Burch  
*Proceedings of NAACL*. [\[paper\]](#)

## Teaching Experience

- 2021 **Department of Computer Science, Columbia University**  
Teaching Assistant, Analysis of Algorithms I  
Professor: [Christos Papadimitriou](#)
- 2019 **Department of Computer Science, University of Pennsylvania**  
Teaching Assistant, Introduction to Algorithms  
Professor: [Sanjeev Khanna](#)
- 2017 - 2018 **Department of Computer Science, University of Pennsylvania**  
Head Teaching Assistant, Data Structures and Algorithms  
Professor: [Rajiv Gandhi](#)

## Service

### Conference Reviewing

- 2023 Causal Representation Learning Workshop at NeurIPS
- 2023 Spurious Correlations, Invariance, and Stability Workshop at ICML
- 2022 Conference on Artificial Intelligence and Statistics (AISTATS)

### University and Department

- 2021, 2023 Pre-Submission Application Review (PAR), Columbia CS Department

## Languages and Skills

Python (+PyTorch, Pandas, PyTorch Lightning, Transformers), C, C++, C#, SQL, Java