

Andy (Dandi) Zhao

(412)-330-8013 ◊ dz352@cornell.edu ◊ <https://www.andyzhao.xyz/>
Gates Hall, 107 Hoy Road, Ithaca, NY 14850 / Bloomberg Center, 2 W Loop Rd, New York, NY, 10044

Education

- Cornell University**, Ithaca/New York, NY, USA
Computing and Information Science
Ph.D. student in Information Science, minor in Government
Advised by Prof. William Hobbs and Prof. Mor Naaman
2020 - 2025 (expected)
- University of Pittsburgh**, Pittsburgh, PA, USA
School of Computing and Information
Master of Science in Information Science
2017 - 2019 ("3+2" program)
- Wuhan University**, Wuhan, Hubei, China
School of Information Management
Bachelor of Management in Library Science
2014 - 2018 ("3+2" program)

Relevant Courses

- Computer Science:**
Machine Learning; Natural Language Processing; Artificial Intelligence; Data Mining; Advanced Database Management; Algorithm Design; Structure of Information Networks; Computer Vision
- Social & Tech:**
Social Computing; Natural Experiment (Causal Inference); Empirical Methods in Research; Large-scale Social Phenomenon; Technology & Law

Academic Works

- [*Chinese astroturf and the limits of computational propaganda on foreign social media*](#) (suspended project, first-authored joint work with Simon DeDeo)
- **Zhao, A., & Chen, Z. (2023).** *Let's report our rivals: how Chinese fandoms game content moderation to restrain opposing voices.* *Journal of Quantitative Description: Digital Media*, 3.
<https://doi.org/10.51685/jqd.2023.006>
 - Poster in IC2S2 2022
- **Zhao, A., & Naaman, M. (2023).** *Unraveling the Potential of Crowds: Insights from a Comparative Study on the Variety, Velocity, Veracity, and Viability of Crowdsourced and Professional Fact-Checking Services.* Forthcoming in the *Journal of Online Trust and Safety*
- [*Platform, jurors, and reporters: understanding how different actors facilitated community-driven content moderation on Weibo*](#) (first-authored joint working paper with Lily Hu)
 - Forthcoming Presentations in ICA 2023 and APSA 2023
- *Moderating Social Media: Misinformation Exposure and the Weibo Jury Community* (work in progress with Alexandra Cirone)
 - Poster in APSA 2022
- *Labor Market of Decentralized Censorship* (work in progress with David Yang, Eddie Yang, and Leo Yang)
- *Filtering Technologies and the Fairness of Natural Language Systems* (working paper, co-author with Eddie Yang, Chad Atalla, Solon Barocas, Su Lin Blodgett, Kate Cook, Kristen Laird, Emily Lawton, Samir Passi, Forough Poursabzi-Sangdeh, Vyoma Raman, Bella Rideau, Emily Sheng, Dan Vann, Hanna Wallach)
 - Poster in TADA 2022

Work Experience

Data & Applied Scientist Intern 5/2023 – 8/2023
Microsoft Redmond, WA

- Synthetic Data Generation and Evaluation for Responsible AI: Assessing the Approach of Using LLM-Generated Synthetic Data

Data & Applied Scientist Intern 6/2022 – 8/2022
Microsoft Bellevue, WA

- Responsible AI: Evaluating Measurement and Mitigation for Representational Fairness Harms in Suggested Reply and Beyond
- Presented my intern paper at Microsoft internal conference MLADS 2022

Research Assistant 7/2019 - 6/2020
Department of Economics, Massachusetts Institute of Technology Cambridge, MA

- Worked on a project about bureaucracy and innovation, which was led by MIT Prof. Daron Acemoglu and Harvard Prof. David Yang
- Collected publication data from academic databases
- Collected leader data and bureaucracy data from official university websites
- Calculated the similarities between papers of different authors based on text and topics
- Analyzed the impact of leadership on topic changes in a difference-in-difference approach

Data Analyst Intern 7/2017 - 8/2017
China Telecom Chongqing, China

- Assisted with user-based modeling for further service
- Evaluated the performance of a crawler and helped with the design process
- Joined department team in data analysis competition and won the champion

Specialized Skills

Programming:
Proficient in Python, R
Experienced in MATLAB, Java, SQL, Cypher

Languages:
Mandarin and southwestern Mandarin Chinese (native); English (proficient)

Project Experience

Coursework: Artificial Intelligence for Bubble Breaker 1/2019 - 4/2019
This is a group project for *Artificial Intelligence*, a grad-level course offered by Prof. Daqing He at the University of Pittsburgh.

- Designed an intelligent agent in Python to play a special game called "Bubble." It is a complicated search problem, but our solution won a full score.

Coursework: A Study on Merit-based Immigration: Prediction of Employer-Sponsored Immigration Certification 2/2018 - 4/2018
This is a group project for *Data Mining*, a grad-level course offered by Prof. Yu-Ru Lin at the University of Pittsburgh.

- Applied data mining skills and algorithms in R to analyze the visa data and predicted whether an application could get certified in the future.

Coursework: Identifying the event topics and participants' identities in collective action events through social media text 2/2018 - /2018

This is a personal course project for *Machine Learning*, a grad-level course offered by Prof. Milos Hauskrecht at the University of Pittsburgh.
Applied deep learning and natural language processing in Python to classify different social posts from a highly skewed distribution and predicted the class label of the collective actions in posts

TAing Experience

INFO 2950: Introduction to Data Science – *Prof. David Mimno at Cornell University*
Teaching assistant in Fall 2020

INFO 5430: Urban Data – *Prof. Emma Pierson at Cornell Tech*
Teaching assistant in Fall 2021 and Fall 2022

INFO 1260 / CS 1340: Choices and Consequences in Computing – *Prof. Jon Kleinberg and Prof. Karen Levy at Cornell University*
Teaching assistant in Spring 2022

INFO 5304 /CS 5304: Data Science in the Wild – *Prof. Raja Rajalakshmi Nandakumar at Cornell Tech*
Teaching assistant in Spring 2023