

Andrea Wen-Yi Wang

aww66@cornell.edu || <https://github.com/andrea-w-wang>

EDUCATION

- | | |
|---|--------------------|
| PhD Information Science, Cornell University
- Advisors: Allison Koenecke, David Mimno | Aug 2022 - Present |
| MS Data Science, New York University
- Mathematics and Data track | 2017 - 2019 |
| BA Finance, National Taiwan University
- Presidential Award (top 5%) | 2012 - 2016 |

PUBLICATIONS

- **Wang A**, Lan J, Wang M, Yu C The Evolution of Rumors on a Closed Social Networking Platform During COVID-19: Algorithm Development and Content Study. *JMIR Med Inform* 2021;9(11):e30467. doi: 10.2196/30467
- Harvey, Anna, Orion Taylor and **Andrea Wang**. COVID-19, Jails, and Public Safety: December 2020 Update. Washington, D.C.: Council on Criminal Justice, December 2020

RESEARCH EXPERIENCES

- | | |
|--|----------------------|
| Data Scientist <i>Public Safety Lab, New York University</i> | Jan 2019 - June 2022 |
| <ul style="list-style-type: none">• Created and maintained a system collecting individual-level detainee records daily from around 1,000 U.S. county jail rosters in the US with Python requests, selenium, BeautifulSoup.• Designed operational pipeline involving Amazon Web Services, Github, and Airtable that improved the success rate by 80%.• Built API with Node.js and web interfaces with a MEAN (MongoDB, Express, Angular, Node) stack. | |
| Data Researcher <i>Information Operations Research Group</i> | Sep 2020 - Aug 2021 |
| <ul style="list-style-type: none">• Developed an efficient clustering algorithm to identify messages that belong to the same rumor and studied the temporal propagation patterns of false pandemic-related rumors on LINE. | |
| Core Software Developer <i>Open Culture Foundation</i> | Oct 2019 - Jul 2020 |
| <ul style="list-style-type: none">• Developed and managed the <i>0archive</i> project, an open source automated archiving system for Taiwanese news medias websites, forums, and content farms to provide open texts data for civic and academic researchers. | |
| Research Assistant <i>Institute of Statistical Science, Academia Sinica</i> | Sep 2015 - Jul 2016 |
| <ul style="list-style-type: none">• Performed simulations and analyzed network data with R to investigate the minimum change required for a large network to be balanced on Structural Balance on Social Networks. | |

TEACHING EXPERIENCES

- | | |
|--|-------------|
| Probability and Statistics for Data Science (Graduate-level) <i>NYU Center for Data Science</i> | Fall 2018 |
| <ul style="list-style-type: none">• Weekly Lectures. Topics include important distributions, random variables, convergence, law of large numbers, central limit theorem, Markov Chain, simulation methods, parametric inference, hypothesis testing, and Bayesian inference. | |
| Introduction to Data Science (Graduate-level) <i>NYU Center for Data Science</i> | Spring 2019 |
| <ul style="list-style-type: none">• Weekly Lectures. Contents include Data Visualization with Tableau, Machine Learning models in scikit-learn, Auto-Machine Learning, Text Processing with NLTK, etc. | |

INVITED TALKS AND PRESENTATIONS

- COSCUP 2021. Information operations research as a data science research. Virtual. Jul 2021.
- Chung Yuan Christian University Graduate Seminar. From NTU Finance to NYU Center for Data Science: My experiences and learnings. Taoyuan, Taiwan. May 2021.
- g0v summit 2020. 0archive - an open source archiving system and open data for Taiwan online information space. Tainan, Taiwan. Dec 2020.

SKILLS

- **Programming Languages:** Python, SQL, R, Javascript, Matlab, Linux
- **Cloud Platform** AWS
- **Databases** MySQL, MongoDB, Google BigQuery