

# Xilin Yang

xy4772@princeton.edu | 734-709-5147

---

## Education

**University of Chicago, Harris School of Public Policy**

2023

Master of Public Policy

- [Survey Research](#) and [Data Analytics](#)
- Courses: Causal Inference, Machine Learning, Survey Research Methodology, Statistics of Sample Surveys, GIS Applications, Linear Algebra, Cost-Benefit Analysis

**University of Michigan, Ann Arbor**

2021

BA in Economics, Philosophy with High Honors

## Working Paper

- **Yang, Xilin.** “Diagnosing Hate Speech Classification: Where Do Humans and Machines Disagree, and Why?”
- **Yang, Xilin.** “A Computational Analysis of Historical Fishery Conflicts in Europe: Evidence from 40 Years of News”
- Kevin Greene, ..., **Xilin Yang.** “Evaluating Biases in Generative AI Depictions of Foreign Political Leaders”
- Institute for Replication, ..., Do Won Kim, Do-Hoon Kim, **Xilin Yang.** “Robustness Checks on Recent Papers from Political Analysis”

## Professional Service

**NeurIPS 2024 Conference Ethics Reviewer**

August 2024

- Conducted ethics reviews for five academic papers by NeurIPS’s standards on treatments involving human subjects, data privacy, copyright, and environmental impact.

## Professional Experiences

**Senior Research Specialist, Princeton University**

September 2023 – Current

### **1. Model Complexity in Political Science Research**

I provided research assistance to the paper “[Model Complexity for Supervised Learning: Why Simple Models Almost Always Work Best, And Why It Matters for Applied.](#)” I replicated 17 quantitative research papers, running simulations via GPU batch code to show why the predictive performance of complex machine learning models is low in social science research.

### **2. Computational Text Analysis of International Conflicts**

I created a text-as-data research design to analyze European countries’ conflicts on fishing quota regulation from the last 40 years of news. I web-scraped 25,000 historical news from Lexis Nexis and used a SoTA open-sourced pre-trained large language model E5-mistral-7b-instruct to generate text embedding and check text similarity. I used semantic role labeling and sentiment analysis to identify interactions among countries.

### **3. Southeast Asian Social Media Study**

I built a data collection pipeline to search and scrape Southeast Asian politicians’ Facebook accounts automatically from Google search results.

**Graduate Researcher, Harris Policy Labs, Harris School of Public Policy**

March –

May 2023

- Served as the leader of a 6-person team of graduate students to leverage open-source data for the Bureau for Policy, Planning, and Learning (PPL) of USAID
- Led the team in the statistical analysis of survey data using R Quantified migration flow and identified factors driving migration in Ethiopia
- Presented findings to USAID and informed its policy decision on regional funding allocation

**Research Intern, United Nations Department of Economic and Social Affairs (UN-DESA)**

January 2022 – July 2022

- Carried out survey research on emerging issues in sustainable development. Contributed visualizations and analysis to the UN Survey Study: [“2022 Development Cooperation Forum \(DCF\) Survey Study: Navigating Covid-19 Recovery and Long-Term Risks | Financing for Sustainable Development Office”](#)
- Served as the focal point for the UN 2022 Youth Forum

**Research Assistant, Department of Political Science, University of Michigan**

2019-2020

- Led literature reviews and data analysis on election data and electoral law documents

**Teaching Experiences****Teaching Assistant**, Ross School of Business, University of Michigan

Fall 2020

- Evaluated student performance for a core MBA course, Strategy 502: Corporate Strategy

**Quantitative Skills****Programming Languages:** R, Python**Tech Literacy:** Web Scraping, PyTorch, JAX, Linux**Data Science:** Text as Data, Text Embedding with LLM, Causal Inference, Geospatial Analysis with R and ArcGIS**Certificates and Awards**

- The 2019-2020 Blue Ribbon Award for Exemplary Research Presentation at the University of Michigan

**Non-Academic Fiction Publication**

- Yang, Xilin. (2022). *Run Away* (逃). Shanghai: Shanghai Literature and Art Publishing House. (ISBN in China: 978-7-5321-8268-8. Verified via <https://pdc.capub.cn/>)
  - Published in March 2022, this novel was collected by Suzhou Library (Tier-1 National Library) and Wuxi Library (Tier-1 National Library). It has sold 1,600 copies since publication.