

Stephanie Buongiorno

Skills

- 5 years experience collecting, processing, and modeling large amounts of policy and geospatial data (up to ~150 GB) using Python, R, SQL, and Shell.
- 3 years of leadership experience working on multidisciplinary research teams (domestic and international), giving talks to technical and non-technical audiences, and directing student work.
- 7 years of experience writing analytical and creative research papers on historical and policy data.

Relevant Professional Experience

Researcher at the Human Trafficking Project (upcoming) 2022-23

Congressional Funding Bill (H.R. 2471), Southern Methodist University

- In conversation with Congress, the DOJ, and specialist researchers, we will develop tools that respond to the needs of human intel. We will: a) collect data on human trafficking; b) develop human-in-the-loop machine learning techniques to improve analyses of (and responses to) human trafficking; and c) use geospatial patterns to create a cascading failures model.

Teaching Fellow at Foundations and Applications of Humanities Analytics Institute (upcoming) 2022

National Endowment of the Humanities (NEH) Grant (no. HT-272418-20), Santa Fe Institute

- I will plan, teach, and evaluate workshops on text mining and data analytics for humanities data.

Researcher for Global Urbanization and Housing Affordability: Poverty, Property and the City 2018-21

National Science Foundation (NSF) Grant (no. 1520103), Southern Methodist University

- At Democracy Lab I modeled historical policy data for quantitative and qualitative analysis.
- I stored and processed data sets of 100-200 years of political documents in the form of tables, matrices, and n-grams for time-series analysis.
- I developed the first prototype app for data mining British Parliamentary debates to show language change over time. I used natural language processing frameworks like spaCy and gensim.
- I developed an information extraction method in Python and R with ~20% higher precision and recall than existing methods, such as Stanford's *OpenIE*.
- I processed large language models on a supercomputer using spaCy, Gensim, and SciPy.
- I led the development of a data processing pipeline to disambiguate speakers in 100 years of policy data by generating permutations of names and matching them with edit distances.
- I trained and managed a team of 3 to 6 undergraduates.

Research Assistant for "Toward a History of the Associative-Developmental State" 2021-22

- I worked with an ArcGIS specialist to make historical maps of New York and its labor movements. I geocoded and mapped addresses using geopandas, geopy, and folium.

Research Assistant for "A History of Early and Modern China" 2021

- I stored and processed csv data for time series analysis. I topic modeled debates from Hong Kong's legislative council to show change throughout the 20th-century.

Education

PhD, ABD, Applied Science in Engineering, Southern Methodist University Expected 2023

MA, English, West Virginia University 2017

BA, English and Linguistics, The University of Texas at Arlington (Summa Cum Laude) 2014