

Leslie Huang

lesliehuang@nyu.edu

NYU Center for Data Science, 60 Fifth Avenue

leslie-huang.github.io • linkedin.com/in/huangleslie

EDUCATION

New York University | PhD in Data Science, Sept 2018–March 2021 (*expected*)
Transferred PhD programs with full support for my interdisciplinary research.

Dissertation: “Contextualized NLP tools for unsupervised text analysis”

- “Using BERT to identify discriminating words from contextual connotation”
- “Estimating authors’ stylistic distinctiveness with embeddings: enhancements to ‘A General Model of Author Style’ ”
- “Methods for optimizing unsupervised clustering of interrogative texts in political discourse”

Committee: Arthur Spirling (chair), Kyunghyun Cho, Jonathan Nagler

New York University | MA in Politics, 2018

- Fully funded Politics PhD student, Sept 2015–Aug 2018
- PhD Qualifying Paper: “Estimating the distribution of state power at the UN using topic models” (readers: Arthur Spirling, Jennifer M. Larson)
- Masters’ Paper: “Modelling responsiveness and willingness to negotiate using sentiment: a textual analysis of the FARC-Colombia peace talks” (adviser: Arthur Spirling)

Columbia University | BA in English Literature and Political Science, 2011

Graduated *summa cum laude*

- GPA: 3.99; Dean’s List; Phi Beta Kappa honor society

PUBLICATIONS

Huang, L., P. Perry and A. Spirling. “A General Model of Author ‘Style’ with Application to the UK House of Commons, 1935–2018.” Preprint at *Political Analysis*: doi.org/10.1017/pan.2019.49

Huang, L., K. Cho, and S. Bowman. “Can we identify word senses from deep contextualized word embeddings without supervision?” [Link to Medium article.](#)

WORKING PROJECTS

Huang, L., A. Spirling, J. Zhang, and C. Danescu-Niculescu-Mizil. “Methods for optimizing unsupervised clustering of interrogative texts in political discourse.”

SOFTWARE

R packages on CRAN

- **stylest**: Estimates a new stylometric measure for textual similarity, especially designed to enable comparisons across large groups of authors/speakers.
- **zipR**: Implements Pythonic `zip()` with convenient options to repeat, broadcast, and fill arrays.
- **leaflet.providers**: Enables users to get and retrieve data from providers of map tiles for the leading interactive map visualization package `leaflet`.

PROFESSIONAL EXPERIENCE

RStudio | Engineering Intern

June–August 2019

- Developed software to enhance the `shiny` and `leaflet` ecosystems maintained by RStudio, the leading IDE for data science in R. Deliverables include the `leaflet.providers` package and a not-yet-released reactive web app for R users.

NYU, Center for Data Science | “NLP/Text-as-Data Speaker Series” Coordinator

June 2017–present

- Manage logistics and communications for weekly guest talks by cutting-edge NLP researchers from top universities and companies such as Google, Facebook, and Microsoft Research ([website](#)).

NYU, Politics | Research Assistant

June–September 2017

- Wrote/refactored simulations in R for a study on incomplete sampling of networks in experiments.

NYU, Environmental Studies | Research Assistant

January–August 2017

- Implemented unsupervised machine learning models. Cleaned and constructed dataset of webscraped texts.

W. W. Norton & Company | Assistant Managing Editor

February 2013–July 2015, New York, NY

- *Promoted from Project Editor in September 2013.*
- Ensured accuracy of all text and visual elements throughout the production cycle, from manuscript to finished book.

Foreign Policy Association | Assistant Editor

June 2011–February 2013, New York, NY

- Guided leading scholars in developing book chapters on international affairs. Wrote and researched supplementary publications. Designed print/digital publications. Hired/supervised interns.

TEACHING

NYU, Center for Data Science | Teaching Assistant for PhD/MS courses

- Text-as-Data (R), Spring 2020
- Intro to Data Science (Python), Fall 2018
- Text-as-Data (R), Spring 2018

AWARDS AND FELLOWSHIPS

- Center for Data Science Fellowship, 2018–present
- Society for Political Methodology Travel Grant, 2018, 2019
- Center for Data Science Supplementary Fellowship Grant, 2018
- American Political Science Association Travel Grant, 2018
- MacCracken Fellowship, 2015–2018
- NYU Dept. of Politics Summer Research Funding, 2016, 2017, 2018

CONFERENCE PRESENTATIONS

- “STAR” Talk, Natural Language, Dialog and Speech Symposium, 2019
- Poster, Society for Political Methodology Annual Meeting, 2019
- Paper presentation, Conference on Data Science and Law, 2019
- Paper presentation, American Political Science Association Annual Meeting, 2018
- Paper co-presentation, Society for Political Methodology Annual Meeting, 2018
- Poster, Society for Political Methodology Annual Meeting, 2018
- Invited participant, Text Analysis Developers’ Workshop, 2018

SKILLS

Programming

Python (sklearn/pandas, PyTorch)

R

SQL (SQLite, MariaDB)

Spark (PySpark)

Methods

Machine learning

Deep learning

Text mining

Data visualization

Causal inference