

Brandon M. Stewart (December 6, 2024)

CONTACT

INFORMATION

Princeton University
149 Wallace Hall
Princeton, NJ 08544

E-mail: bms4@princeton.edu

Web: <http://brandonstewart.org>

RESEARCH INTERESTS

Social Science, Text as Data, Causal Inference, Sociology, Political Science

ACADEMIC

APPOINTMENTS

Princeton University

Associate Professor, Department of Sociology, July 2022–

Faculty Affiliate, Department of Politics 2015–

Faculty Associate, Office of Population Research 2015–

Faculty Associate, Center for Digital Humanities 2015–

Faculty Associate, Center for Statistics and Machine Learning 2015–

Faculty Associate, Center for Information Technology Policy 2016–

Faculty Associate, Institute for Computational Science and Engineering 2016–

Faculty Associate, School of Public and International Affairs 2023–2026

Faculty Associate, Data-Driven Social Science 2024–

Assistant Professor, Department of Sociology, July 2015–June 2022

Arthur H. Scribner Bicentennial Preceptor July 2018–June 2021

CURRENT

EDITORIAL POSITIONS

Co-Editor-in-Chief, *Political Analysis*, August 2024–

joint with Dan Hopkins

Associate Editor, *Sociological Methods & Research*, February 2023–

handling manuscripts on text as data and AI

EDUCATION

Harvard University

Ph.D., Government, May 2015

Committee: Gary King, Beth Simmons, Dustin Tingley

A.M., Statistics, May 2014

College of William & Mary

B.A., Government and International Relations, May 2007

- *summa cum laude*

- Highest Honors in Government

BOOK

Grimmer, Justin, Margaret E. Roberts and Brandon M. Stewart. *Text as Data: A New Framework for Machine Learning and the Social Sciences*. 2022. Princeton: Princeton University Press.

PUBLICATIONS

48. Spirling, Arthur and Brandon M. Stewart. “What good is a regression? Inference to the best explanation and the practice of political science research.” Conditionally Accepted at *Journal of Politics*.
47. Green, Breanna, Will Hobbs, Pedro L. Rodriguez, Arthur Spirling, Brandon M. Stewart. “Measuring Distances in High Dimensional Spaces: Why Average Group Vector Comparisons Exhibit Bias, And What to Do About it” Conditionally Accepted at *Political Analysis*.
46. Rodriguez, Pedro L., Arthur Spirling, Brandon M. Stewart, and Elisa M. Wirsching. ”Multilanguage Word Embeddings for Social Scientists: Estimation, Inference and Validation Resources for 157 Languages.” Conditionally Accepted at *Political Analysis*.
- <https://alcmbeddings.org/>
45. Felton, Chris and Brandon M. Stewart. “Handle with Care: A Sociologist’s Guide to Causal Inference with Instrumental Variables.” *Sociological Methods & Research*. Forthcoming.
44. Saenger, Till Raphael, Musashi Hinck, Justin Grimmer, and Brandon M. Stewart. 2024. AutoPersuade: A Framework for Evaluating and Explaining Persuasive Arguments. In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, pages 16325–16342.
43. Wongkamjan, Wichayaporn, Feng Gu, Yanze Wang, Ulf Hermjakob, Jonathan May, Brandon M. Stewart, Jonathan K. Kummerfeld, Denis Peskoff and Jordan Lee Boyd-Graber. “More Victories, Less Cooperation: Assessing Cicero’s Diplomacy Play” *Proceedings of the Association of Computational Linguistics*, 2024.
42. Kapoor, Sayash, et al. (2024). REFORMS: Consensus-based Recommendations for Machine-learning-based Science. *Science Advances*, 10(18).
- <https://reforms.cs.princeton.edu/>
41. Peskoff, Denis, Adam Visokay, Sander Schulhoff, Benhamin Wachspress, Alan Blinder and Brandon M. Stewart. “GPT Deciphering Fedspeak: Quantifying Dissent Among Hawks and Doves.” *Findings of Empirical Methods in Natural Language Processing*. 2023.
40. Egami, Naoki, Musashi Hinck, Brandon M. Stewart, and Hanying Wei. “Using Imperfect Surrogates for Downstream Inference: Design-based Supervised Learning for Social Science Applications of Large Language Models” *Neural Information Processing Systems*. 2023.
39. Grimmer, Justin, Dean Knox and Brandon M. Stewart. “Naïve regression requires weaker assumptions than factor models to adjust for multiple cause confounding.” *Journal of Machine Learning Research*. 2023, 24(182) 1–70.
38. Rodriguez, Pedro, Arthur Spirling, and Brandon M. Stewart. “Embedding Regression: Models for Context-Specific Description and Inference.” *American Political Science Review*. 2023. 117(40), 1255-1274.
37. Denis Peskoff and Brandon M. Stewart. 2023. “Credible without Credit: Domain Experts Assess Generative Language Models.” In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 427–438, Toronto, Canada. Association for Computational Linguistics.
- Selected for Oral Presentation
- Data and Code: https://denis.ai/bot_thoughts/

36. Mathur, Arunesh, Angelina Wang, Carsten Schwemmer, Maia Hamin, Brandon M. Stewart and Arvind Narayanan. “Manipulative Tactics are the Norm in Political Emails: Evidence from 300K emails from the 2020 U.S. election cycle” *Big Data & Society*, 2023 10(1).
- Selected Press Coverage: *Washington Post*, *Vice*, *Fast Company*, *The Atlantic*
- <https://electionemails2020.org/>
35. Egami, Naoki, Christian Fong, Justin Grimmer, Margaret E. Roberts, and Brandon M. Stewart. “How to make causal inferences using texts.”, *Science Advances*, 2022, 8(42), 1–13.
- published with commentary by Dhanya Sridhar and David M. Blei.
34. Feder, Amir, Katherine A. Keith, Emaad Manzoor, Reid Pryzant, Dhanya Sridhar, Zach Wood-Doughty, Jacob Eisenstein, Justin Grimmer, Roi Reichart, Margaret E. Roberts, Brandon M. Stewart, Victor Veitch and Diyi Yang. “Causal Inference in Natural Language Processing: Estimation, Prediction, Interpretation and Beyond” *Transactions of the Association for Computational Linguistics*, 2022.
33. Ying, Luwei, Jacob M. Montgomery, and Brandon M. Stewart. “Topics, Concepts, and Measurement: A Crowdsourced Procedure for Validating Topics as Measures.” *Political Analysis*, 2022, 30 (4).
32. Lundberg, Ian, Rebecca Johnson and Brandon M. Stewart. “What is Your Estimand? Defining the Target Quantity Connects Statistical Evidence to Theory” *American Sociological Review*, 2021, 86(3), 532–565.
- Winner of the ASA Methodology Section Outstanding Publication Award 2022
31. Grimmer, Justin, Margaret E. Roberts and Brandon M. Stewart. “Machine Learning for Social Science: An Agnostic Approach” *Annual Review of Political Science*, 2021, Volume 24.
30. Roberts Margaret E., Brandon M. Stewart, and Richard Nielsen. “Adjusting for Confounding with Text Matching.” *American Journal of Political Science*. 2020, 64 (4): 887-903.
29. Fragile Families Challenge Team. “Measuring the predictability of life outcomes with a scientific mass collaboration.” *Proceedings of the National Academy of Sciences*, March 2020.
- Author contribution statement: “M.J.S., I.L., A.T.K., J.B.-G., B.E.E., M.H., K.L., A.N., B.M.S., D.J.W., and S. McLanahan designed research”
28. Lundberg, Ian and Brandon M. Stewart. “Comment: Summarizing income mobility with multiple smooth quantiles instead of parameterized means.” *Sociological Methodology*, 2020, 50:96-111.
27. de Marchi, Scott and Brandon M. Stewart. “Wrestling with complexity in computational social science: theory, estimation, and representation.” In *SAGE Handbook of Research Methods in Political Science and International Relations*. SAGE, 2020.
26. Roberts, Margaret E., Brandon M. Stewart, Dustin Tingley. “stm: R Package for Structural Topic Models.” *Journal of Statistical Software*. 2019, 91(2), 1–40.
25. Horowitz, Michael, Brandon M. Stewart, Dustin Tingley, Michael Bishop, Laura Resnick Samotin, Margaret E. Roberts, Welton Chang, Barbara Mellers and Phil Tetlock. “What Makes Foreign Policy Teams Tick: Explaining Variation in Group Performance At Geopolitical Forecasting” *Journal of Politics*. 2019, 81(4), 1388-1404.

24. Yeomans, Mike, Alex Kindel, Kimia Mavon, Justin Reich, Brandon M. Stewart and Dustin Tingley. "Engagement Across Political Differences in Massive Open Online Courses." *International Journal of Artificial Intelligence in Education*. 2018, 28(4), 553-589.
23. Chaney, Allison J.B., Brandon M. Stewart and Barbara E. Engelhardt. "How Algorithmic Confounding in Recommendation Systems Increases Homogeneity and Decreases Utility." In *Twelfth ACM Conference on Recommender Systems (RecSys '18)*. 2018.
22. Khodak, Mikhail, Nikunj Saunshi, Yingyu Liang, Tengyu Ma, Brandon M. Stewart, and Sanjeev Arora. "A La Carte Embedding: Cheap but Effective Induction of Semantic Feature Vectors." In *Proceedings of the Association of Computational Linguistics*. 2018.
- Selected for Oral Presentation
21. Simmons, Beth A., Paulette Lloyd, Brandon M. Stewart. "The Global Diffusion of Law: Transnational Crime and the Case of Human Trafficking." *International Organization*. 2018; 72(2) 249-281.
20. Kindel Alex, Mike Yeomans, Justin Reich, Brandon M. Stewart and Dustin Tingley. "Discourse: MOOC Discussion Forum Analysis at Scale", in *Proceedings of the Fourth (2017) ACM Conference on Learning @ Scale.*; 2017 :141-142.
19. Roberts, Margaret E., Brandon M. Stewart, Edoardo M. Airoidi. "A topic model for experimentation in the social sciences." *Journal of the American Statistical Association*. 2016; 111(515): 988-1003.
18. Reich, Justin, Brandon M. Stewart, Kimia Mavon, and Dustin Tingley. "The Civic Mission of MOOCs: Measuring Engagement across Political Differences in Forums." *Proceedings of the Third (2016) ACM Conference on Learning @ Scale*. 2016: 1-10.
- Best Paper Award for Conference
17. Roberts, Margaret E., Brandon M. Stewart, Dustin Tingley. "Navigating the Local Modes of Big Data: The Case of Topic Models" In *Computational Social Science: Discovery and Prediction*. Cambridge University Press. 2016.
16. Reich, Justin, Dustin Tingley, Jetson Leder-Luis, Margaret E. Roberts, Brandon M. Stewart. "Computer Assisted Reading and Discovery for Student Generated Text in Massive Open Online Courses." *Journal of Learning Analytics*. 2015; 2(1) 156-184
15. Lucas, Christopher, Richard Nielsen, Margaret E. Roberts, Brandon M. Stewart, Dustin Tingley. "Computer assisted text analysis for comparative politics." *Political Analysis*. 2015; 23(2): 254-277.
- Included in *Political Analysis* virtual issue on Online Research Methods
- Reprinted in *Advances in Political Methodology* 2017 Edited by Robert J. Franzese
14. Chuang, Jason, Margaret Roberts, Brandon Stewart, Rebecca Weiss, Dustin Tingley, Justin Grimmer and Jeffrey Heer. "TopicCheck: Interactive Alignment for Assessing Topic Model Stability" *North American Chapter of the Association for Computational Linguistics – Human Language Technologies (NAACL HLT)*. 2015.
13. Romney, David, Brandon M. Stewart, and Dustin Tingley. "Plain Text: Transparency in the Acquisition, Analysis, and Access Stages of the Computer-assisted Analysis of Texts." *Qualitative and Multi-Method Research*. 2015; 13(1): 32-37.
12. Roberts, Margaret E., Brandon M. Stewart, Dustin Tingley, Christopher Lucas, Jetson Leder-Luis, Shana Gadarian, Bethany Albertson and David Rand. "Structural

topic models for open-ended survey responses.” *American Journal of Political Science*. 2014; 58(4): 1064-1082.

- Awarded Gosnell Prize for Excellence in Political Methodology

11. Chuang, Jason, John D. Wilkerson, Rebecca Weiss, Dustin Tingley, Brandon M. Stewart, Margaret E. Roberts, Forough Poursabzi-Sangdeh, Justin Grimmer, Leah Findlater, Jordan Boyd-Graber and Jeff Heer. “Computer-Assisted Content Analysis: Topic Models for Exploring Multiple Subjective Interpretations.” *Advances in Neural Information Processing Systems Workshop on Human-Propelled Machine Learning*. 2014.
10. Roberts, Margaret E, Brandon M. Stewart, Dustin Tingley and Edoardo M. Airoldi “The Structural Topic Model and Applied Social Science.” *Advances in Neural Information Processing Systems Workshop on Topic Models: Computation, Application, and Evaluation*. 2013.
- Selected for Oral Presentation.
9. Zhukov, Yuri M. and Brandon M. Stewart “Choosing Your Neighbors: Networks of Diffusion in International Relations.” *International Studies Quarterly*. 2013; 57(2):271-287.
8. Grimmer, Justin and Brandon M. Stewart. “Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts.” *Political Analysis*. 2013; 21(3):267-297.
- *Political Analysis* Editor’s Choice Award
- Reprinted in *Advances in Political Methodology* 2017 Edited by Robert J. Franzese
7. O’Connor, Brendan, Brandon M. Stewart, and Noah A. Smith. “Learning to Extract International Relations from Political Context.” In *Proceedings of the Association of Computational Linguistics*. 2013.
6. Andersen, Judith Pizarro, Roxane Silver, Brandon M. Stewart, Billie Koperwas, and Clemens Kirschbaum. “Psychological and Physiological Responses following Repeated Peer Death.” *PLoS ONE*. 2013; 8(9): e75881.
5. Lloyd, Paulette, Simmons Beth and Stewart Brandon M. “Combating Transnational Crime: The Role of Learning and Norm Diffusion in the Current Rule of Law Wave.” In *The Dynamics of the Rule of Law*. Editors: Michael Zürn, André Nolkaemper and Randy Peerenboom. Cambridge University Press. 2012. Cambridge: 153-180.
4. Stewart, Brandon M. and Yuri M. Zhukov. “Use of force and civil-military relations in Russia: an automated content analysis.” *Small Wars & Insurgencies*. 2009; 21(2): 319-343.
3. Shellman, Stephen M. and Brandon M. Stewart. “Early Warning Models of Haitian Flight.” *Civil Wars*. 2007; 9(2): 174-99.
2. Shellman, Stephen M. and Brandon M. Stewart. “Political Persecution or Economic Deprivation? A Time-Series Analysis of Haitian Migration to the United States.” *Conflict Management & Peace Science* 2007; 24(2): 121-37.
1. Reeves, Andrew M, Stephen M. Shellman and Brandon M. Stewart. “Fair & Balanced or Fit to Print? The Effects of Media Sources on Statistical Inferences.” *University of Georgia Occasional Paper Series*. 2006. 1-14.
2. Coppola, Antonio and Brandon M. Stewart. “lbfgs: Efficient L-BFGS and OWL-QN Optimization in R.” Software Vignette. 2014.

1. Stewart, Brandon M. “Latent Factor Regressions for the Social Sciences.” Technical Report. 2014.

PUBLIC WORKING PAPERS (See brandonstewart.org or Google Scholar for up to date working papers.)

3. Liu, Naijia, Xinlan Emily Hu, Yasemin Savas, Matthew A. Baum, Adam J. Berinsky, Allison J.B. Chaney, Christopher Lucas, Rei Mariman, Justin de Benedictis-Kessner, Andy Guess, Dean Knox, and Brandon M. Stewart. “‘Filter bubble’ algorithmic recommendations have limited effects on polarization: Naturalistic, short-exposure experiments on YouTube.”
2. Waight, Yuan, Roberts and Stewart. “Speaking with the State’s Voice: The Decade-Long Growth of Government-Authored News Media in China under Xi Jinping”
1. Egami, Hinck, Stewart and Wei. “Using Large Language Model Annotations for the Social Sciences: A General Framework of Using Predicted Variables in Downstream Analyses.”

SOFTWARE

Major Active Packages

stm: R package for estimation of the Structural Topic Model with Margaret Roberts and Dustin Tingley. structuraltopicmodel.com
(2018 Statistical Software Award from Society for Political Methodology)

conText: An R Package for text embedding with Pedro Rodriguez and Arthur Spirling. <https://cran.r-project.org/web/packages/conText/index.html>

validateIt: An R Package for Topic and Label Validation with Luwei Ying and Jacob Montgomery. <https://github.com/Luwei-Ying/validateIt>

Older Software Packages

stmCorrViz: R package for generating an interactive visualization of topic correlations and hierarchy in a Structural Topic Model (STM) with Antonio Coppola, Margaret E. Roberts and Dustin Tingley.
<http://cran.r-project.org/web/packages/stmCorrViz/>

stmBrowser: R package which allows users to interactively explore the relationships between topics and the covariates estimated from the *stm* package in R. With Michael K. Freeman, Jason Chuang, Margaret E. Roberts and Dustin Tingley. structuraltopicmodel.com

stmgui: R package providing a graphical user interface for the ‘stm’ text analysis package with Dan Zangri and Dustin Tingley. <https://CRAN.R-project.org/package=stmgui>

textmatching: An R package for adjustment with text-based confounding with Richard Nielsen and Margaret Roberts.

lbfgs: An R Package for Limited-memory BFGS Optimization with Antonio Coppola. <http://cran.r-project.org/web/packages/lbfgs/index.html>

AWARDS

ASA Methodology Section Leo Goodman Award (2024)
honors an outstanding researcher in sociological methodology within 15 years of PhD

receipt

Political Methodology Emerging Scholar Award (2023, awarded 2024)
honors a young researcher, within ten years of their degree, who is making notable contributions to the field of political methodology.

ASA Methodology Section Outstanding Publication Award (2022)
recognizes an outstanding article or book in sociological methodology published in the prior three calendar years. Awarded for “What is Your Estimand? Defining the Target Quantity Connects Statistical Evidence to Theory” with Lundberg and Johnson

Graduate Mentoring Award (2020)
honors Princeton faculty members (one in each of four divisions) who are exemplary in supporting the development of their graduate students as teachers, scholars, and professionals based on nominations from graduate students.

Outstanding Faculty Advisor Award (2019)
selected by the graduating class of undergraduate students concentrating in Sociology at Princeton University.

Society for Political Methodology Statistical Software Award (2018)
recognizes individuals for developing statistical software that makes a significant research contribution. Awarded to **stm** with Roberts and Tingley.

Arthur H. Scribner Bicentennial Preceptorship (2018-2021)
the bicentennial preceptorship is a three-year honorific appointment held by approximately ten early-career scholars at any one time across all departments at Princeton.

ACM Learning at Scale Conference Best Paper Award (2016)
awarded to “The Civic Mission of MOOCs” with Reich, Mavon and Tingley.

William and Mary Baxter-Ward Fellow (2015)
invited lecture and honor for distinguished alumni of the William & Mary Government department.

Edward M. Chase Dissertation Prize (2015)
for the best essay on a subject relating to the promotion of world peace among any student at Harvard University. Awarded to “Three Papers in Political Methodology.”

Gosnell Prize for Excellence in Political Methodology (2014)
for the best work in political methodology presented at any political science conference during the preceding year. Awarded to “Structural topic models for open-ended survey responses” with Roberts, Tingley and coauthors.

The Political Analysis Outstanding Reviewer Award (2014)
for exemplary assistance to Political Analysis

Political Analysis Editor’s Choice Award (2013)
for an article providing an especially significant contribution to political methodology. Awarded to ‘Text as Data’ with Grimmer.

National Science Foundation Graduate Research Fellowship (2010-2013)

Certificate of Distinction in Teaching (2011)

Richard L. and Faith P. Morningstar Graduate Fellowship (2008-2009)

Warner Moss Prize for Best Undergraduate Thesis in Government (2007)

EXTERNAL
GRANTS

2022–2023 Defense Advance Research Projects Agency SHADE Project (subcontract from UMD; joint with Karthik Narasimhan) \$240,000

2022–2024 NSF/CRA 2021 Computing Innovation Fellows (funding Denis Peskoff) \$352,009

2021 Microsoft Unrestricted Gift “Computational Analysis of Police Body-Worn Camera Footage” (with Dean Knox, Jonathan Mummolo, and Olga Russakovsky) \$200,000 + \$50,000 in cloud computing credit

2020 Facebook Research Award “Do online video recommendation algorithms increase affective polarization?” (with Andy Guess and a large team) \$99,997

2017 National Science Foundation RIDIR Award, “Collaborative Research: Analytical tools for text-based social data integration” (with Amarnath Gupta and Molly Roberts) \$1,516,099

2015 Hewlett Foundation “Computational Measures of Engagement Across Difference in Online Courses” (with Justin Reich and Dustin Tingley), \$50,000

2015 Spencer Foundation “Computational Measures of Engagement Across Difference in Online Courses” (with Justin Reich and Dustin Tingley), \$400,000

TEACHING

(See materials at <https://bstewart.scholar.princeton.edu/teaching>)

Courses

Spring 2025. Princeton University. SPIA 586a Machine Learning for Policy Analysis (Graduate)

Fall 2024. Princeton University Sociology 500: Applied Social Science Statistics (Graduate)

Spring 2024. Princeton University. SPIA 586a Machine Learning for Policy Analysis (Graduate)

Spring 2023. Princeton University. Sociology 306/Statistics and Machine Learning 306: Machine Learning with Social Data: Opportunities and Challenges (Undergraduate)

Fall 2022. Princeton University Sociology 500: Applied Social Science Statistics (Graduate)

Spring 2022. Princeton University. Sociology 306/Statistics and Machine Learning 306: Machine Learning with Social Data: Opportunities and Challenges (Undergraduate)

Fall 2021. Princeton University Sociology 592: Text as Data: Statistical Text Analysis for the Social Sciences (Graduate)

Spring 2021. Princeton University. Sociology 306/Statistics and Machine Learning 306: Machine Learning with Social Data: Opportunities and Challenges (Undergraduate)

Fall 2020. Princeton University. Sociology 400: Applied Social Science Statistics (Undergraduate)

Fall 2020. Princeton University. Sociology 500: Applied Social Science Statistics (Graduate)

Fall 2020. Princeton University. Statistics and Machine Learning 515/ Sociology 515: Machine Learning with Social Data: Opportunities and Challenges (Graduate)

Spring 2019 (Half-Term). Princeton University. Sociology 592: Text as Data: Statistical Text Analysis for the Social Sciences (Graduate)

Fall 2018. Princeton University. Sociology 400: Applied Social Science Statistics (Undergraduate)

Fall 2018. Princeton University. Sociology 500: Applied Social Science Statistics (Graduate)

Spring 2017. Princeton University. Sociology 401/504: Advanced Social Statistics (Undergraduate/Graduate)

Fall 2016. Princeton University. Sociology 400/500: Applied Social Science Statistics (Undergraduate/Graduate)

Fall 2016. Princeton University. Sociology 320: Growing Up Poor Across America: An Introduction to Reasoning With Data co-taught with Maggie Frye (Undergraduate)

Spring 2016 (Half-Term). Princeton University. Sociology 592: Text as Data: Statistical Text Analysis for the Social Sciences (Graduate)

Spring 2016. Princeton University. Sociology 504: Advanced Social Statistics (Graduate)

Fall 2015. Princeton University. Sociology 500: Applied Social Science Statistics. (Graduate)

Princeton Freshman Scholar's Institute

(see more about the organization at <https://fsi.princeton.edu/about>)

Summer 2024. Sociology/Politics 245: Visualizing Data. Course Head. Vikram Ramaswany served as the lead instructor.

Summer 2023. Sociology/Politics 245: Visualizing Data. Course Head. Vikram Ramaswany served as the lead instructor.

Summer 2022. Sociology/Politics 245: Visualizing Data. Course Head. Chris Felton and Vikram Ramaswany served as the lead instructors.

Summer 2021. Sociology/Politics 245: Visualizing Data. Course Head. [Hannah Waight](#) served as the lead instructor.

Summer 2020. Politics 245: Visualizing Data. Course Head. [Hannah Waight](#) served as the lead instructor. NB: this was a condensed and not-for-credit course due to COVID.

Short-Courses

August 2023, Washington University Data Science and Political Science. Text as Data.

August 2021, Statistical Horizons. Text as Data.

December 2020, Statistical Horizons. Text as Data.

April 2019, Yale University Department of Sociology. Text as Data.

July 2018, Bamberg Graduate School of Social Sciences. Text as Data.

June 2018, Data Science Summer School (DS³) École Polytechnique. Text as Data in the Social Sciences.

December 2017, University of Michigan Complex Systems. Text as Data.

May 2015, Florida State University. Text as Data Graduate Methods Workshop.

June 2010, University of Washington, Seattle. Tools for Text Workshop: Automated Topic Models in Political Science: Practical Skills for Document Clustering in R

Fall 2006, College of William & Mary. Political Theory and International Relations (co-taught with Paul Gutierrez and David Dessler, semester long, one-credit course).

EXTERNAL
INVITED
PRESENTATIONS

2025: Methods Workshop at UCLA's California Center for Population Research (CCPR), IC2S2 Keynote Address: Swedish Excellence Center for Computational Social Science (SweCSS)

2024: University of Wisconsin Madison: La Follette School seminar, Summer Institute for Computational Social Science: University of Pennsylvania, Columbia University Media Effects Empirical Workshop, University of Chicago Data Science, Washington University at St. Louis Political Science/Sociology/Statistics Joint Seminar

2023: Columbia University Political Methodology Seminar, ETH Zurich Online Seminar in Economics & Data Science, Renmin University of China (Virtual), Washington University at St. Louis TRIADS, Dartmouth Quantitative Social Science, Conference on Digital Experimentation @ MIT Fireside Panel

2022: Yale University Sociology Department, Yale University Data Science, University of Chicago Political Science, Summer Institute for Computational Social Science: University of Pennsylvania, Summer Institute for Computational Social Science: Duke, University of Pennsylvania Sociology Colloquium, University of Michigan Information School: Data Science/Computational Social Science, SUNY Binghamton Statistics Seminar

2021: University of Sheffield "Bridging the Gap: Innovative methods for text analysis in political science and IR" roundtable, Duke University Empirical Implications of Theoretical Models, Watson Institute for International and Public Affairs at Brown University, Columbia University Political Economy Seminar Series

2020: New York University Politics Department, New York University Data Science, Institute for Advanced Studies (Social Science and Machine Learning Conference)

2019: NYU Abu Dhabi, UCLA California Center for Population Research, UC San Diego eScience International Conference, UC San Diego Political Science Methods Seminar, Columbia Data Science Seminar, New York University Data Science, TextXD: Text Analysis Across Domains Keynote at UC Berkeley, Stanford Computational Text Analysis in the Social Sciences

2018: Berkeley Demography, Data Science Summer School (DS³) École Polytechnique, MIT Institute for Data, Systems and Society

2017: University of Minnesota Psychology, Stanford Political Science, New York University Sociology, Association for Computational Linguistics: Workshop on Natural Language Processing and Computational Social Science, Yale University Political Science, Duke University Political Science, University of Washington CSSS, University of Michigan Complex Systems

2016: DARPA ISAT Conference on Causal Inference and Machine Learning, Political Instability Task Force Conference: Machine Learning and Automated Data Collection, Duke University Empirical Implications of Theoretical Models

2015: Microsoft Research New York, New York University: Center for Data Science, William & Mary Government Department

2014: Penn State: New Faces in Political Methodology, University of Maryland Computer Science, UMass Amherst Computer Science, MIT Political Science Methods Workshop, Princeton Politics Methods Workshop

2013: NC State University: Camp Resources Big Data Panel

POSTDOCTORAL
ADVISING

(placements are at time of departure/graduation)

4. Denis Peskoff (2022–, Computing Research Association 2021 Computing Innovation Fellow)
3. Musashi Hinck (né Jacobs-Harukawa), AI Research Scientist at Intel Labs (2022–2024, DDSS Postdoctoral Research Fellow)
2. Michael Zhang, Assistant Professor University of Hong Kong Statistics and Actuarial Science (2018–2020)
1. Allison Chaney, Assistant Professor of Marketing, Duke’s Fuqua School of Business → Lecturer at Fuqua (2016–2019, IC Postdoctoral Research Fellowship)

DISSERTATION
COMMITTEES
(CHAIR)

5. Till Sanger (ORFE, current, co-chair with Ronnie Sircar)
4. Chris Felton, Postdoctoral Scholar at Harvard University School of Education (Sociology/Demography/Social Policy, graduated 2023)
3. Alex Kindel, Assistant Professor at Sciences Po (Sociology, graduated 2023)
2. Hannah Waight, Post-doc at NYU Center for Social Media and Politics → Assistant Professor of Sociology at University of Oregon (Sociology, graduated 2022)
1. Ian Lundberg, NSF Post-doc at UCLA → Assistant Professor at Cornell Information Science → Assistant Professor at UCLA Sociology (Sociology/Demography/Social Policy, graduated 2021)

DISSERTATION
COMMITTEES
(MEMBER)

24. Emily Cantrell (Sociology, current)
23. Marie-Lou Laprise (Sociology, current)
22. Katie Donnelly Moran (Sociology, current)
21. Ha-Joon Chung (Sociology, current)
20. Alejandro Schugurensky (Sociology, current)
19. Dahyun Choi (Politics, current)
18. Gavin Cook (Sociology, graduate 2024), Post-doctoral Fellow
17. Lai Wei (Sociology, graduated 2024), Assistant Professor of Sociology at University of Hong Kong (HKU)
16. Herrissa Lamothe, Post-doctoral fellow Brain and Cognitive Sciences MIT (Sociology, graduated 2023)
15. Hannah Postel, Post-Doctoral Research Fellow at Stanford Economics → Assistant Professor at Duke Sanford School of Public Policy (Office of Population Research, graduated 2022)
14. Will Horne, Post-Doctoral research fellow with the Executive Approval Project, Georgia State University → Assistant Professor of Political Science at Clemson University (Politics, graduated 2022)
13. Gracie Himmelstein, Resident Physician (Office of Population Research, graduated 2022)
12. Simone Zhang, Assistant Professor of Sociology at Notre Dame → Assistant Professor of Sociology at New York University (Sociology, graduated 2021)
11. Gregory Gundersen, Quantitative Researcher, Millennium (Computer Science, graduated 2021)
10. Leslie Huang, AI research engineer at Bloomberg LP (NYU Data Science, graduated 2021)
9. Elsa Voytas, Post-doc Dartmouth College → Assistant Professor at IE University in Madrid, Spain → Assistant Professor Quantitative Social Science, Dartmouth College (Politics, graduated 2021)
8. Arunesh Mathur, Post-doc Princeton University → Competition and Markets Authority, U.K. (Computer Science, graduated 2020)
7. Rebecca Johnson, Assistant Professor Quantitative Social Science, Dartmouth College → Assistant Professor McCourt School of Public Policy, Georgetown University (Sociology/Demography/Social Policy, graduated 2020)
6. Naoki Egami, Assistant Professor of Political Science, Columbia University (Politics, graduated 2020)
5. Clark Bernier, Post-doc NYU Data Science → Lead Data Scientist at Houzz (Sociology; graduated 2020)
4. Han Zhang, Assistant Professor, Division of Social Science, The Hong Kong University of Science and Technology → Assistant Professor of Sociology and International and Public Affairs at Brown University (Sociology; graduated 2020)

3. Kendall Cox Park, Post-doc Vanderbilt Management School → Assistant Professor of Organization Studies at Vanderbilt Management School (Sociology; graduated 2018)
2. Yuki Shiraito, Post-doc Dartmouth → Assistant Professor of Political Science, University of Michigan (Politics; graduated 2017)
1. Allison Chaney, Post-doc Princeton University → Assistant Professor of Marketing, Duke's Fuqua School of Business (Computer Science; graduated 2016)

OTHER PH.D.
ADVISING

Dissertation External Reader (4): Patrick Ishizuka (Sociology, 2016), Leah Gillion (Sociology, 2019), Jason Windawi (Sociology, 2021), Will Schulz (Politics, 2023)

Comprehensive Exams Administered (12): Taylor Winfield (Sociology, 2017), Ian Lundberg (Sociology, 2017), Herrissa Lamothe (Sociology, 2018), Alex Kindel (Sociology, 2019), Chris Felton (Sociology, 2019), Cambria Naslund (Sociology, 2020), Eli Lucherini (Computer Science, 2020), Angelina Wang (Computer Science, 2021), Matt Mleczo (Population Research, 2021), Sayash Kapoor (Computer Science, 2022), Maxwell Fineman (Sociology, 2023), Varun Satish (Population Research, 2024)

First Year Advising: Sofia Avila (2024), Yasemin Savas (2024).

UNDERGRADUATE
ADVISING

*JP means Junior Paper, an independent paper done in junior year at Princeton.

Total Theses: 15, Total JPs: 10

2023–24: Jackson Ford (Sociology, JP), Valerie Neske (Sociology, JP)

2022–23: Ali McCarthy (Sociology), Keeshawn Kellman (Sociology), JD Copeland (Sociology, JP), Mashonica Evans (Sociology, JP)

2021–22: Tara Shawa (Sociology), Summer Crown (Sociology, JP), Tracie Kwon (Sociology, JP)

2020–21: Josh Tebeau (Politics), Ali Skarzynski (Sociology), Zazu Asatiani (Sociology), Matthew Hampson (Sociology, JP), Tommy Ren (Sociology, JP), Sarah Lee (Sociology, JP)

2019–20: Jacob Zimmer (Computer Science), Kitty Moraes (Sociology), Grace Miles (Sociology), Chase Hommeyer (Sociology), Sarah Varghese (Computer Science), Chase Williams (Sociology, JP)

2017–18: Gregory Seifert (Sociology), Haydn Melia (Sociology)

2016–17: Katherine Yerkes (Sociology), Yaa Kumah (Sociology)

DEPARTMENTAL
AND UNIVERSITY
SERVICE

Ongoing

Interdisciplinary Data Science Faculty Search Committee (2024–2025)

Center for Digital Humanities Executive Committee (2024–2027)

Research Computing Planning Group (2024–)

Princeton Council on Teaching and Learning (2023–2025), Chair (2024–2025)

Princeton Language + Intelligence Research Advisory Committee (2023–)

Center for Statistics and Machine Learning Undergraduate Program Executive Committee (2022–)

Executive Committee, Data Driven Social Science (2022–2025)

Research Software Engineering Program Steering Committee (2021–)
Director of the Statistics Core and Member of the Executive Committee, Office of Population Research (2018–)
Founder/Organizer, Sociology Statistics Reading Group (2016–)
Founder/Organizer, Graduate Methods Camp, Department of Sociology (2016–)

Completed

Princeton AI in Research Advisory Group (2023–2024)
SPIA Review of Data Science and Program Evaluation (2023–2024)
Interdisciplinary Data Science Faculty Search Committee (2022–2023)
Co-organizer Notestein Seminar Series, Office of Population Research (2022–2023)
Provost's Cloud Computing Resources Working Group (2021–2023)
Steering Committee for the Computing Architecture Strategy Initiative (2021–2022)
Humanities Computing Curriculum Committee (2020–2023)
Graduate Mentoring Award Committee (2020)
Center for Digital Humanities Executive Committee (2018–2021)
Organizer, Quantitative Social Science Colloquium, Departments of Sociology, Politics, Psychology (2018–2019)
Co-Organizer, Quantitative Social Science Colloquium, Department of Sociology and Politics (2016–2018)
Member, Graduate Committee, Department of Sociology (2015–16, 2016–17, 2018–2019, 2020–2022)
Member, Diversity Committee, Department of Sociology (2015–17)
Organizer, Theorodology Workshop (2015–16)

SERVICE TO THE
FIELD

Ongoing

Co-Editor-in-Chief, *Political Analysis*, August (2024–)
Associate Editor, *Sociological Methods & Research* (2023–)
Associate Editor, *ACM/IMS Journal of Data Science* (2023–present)
Co-lead Biostatistics, Epidemiology, and Research Design Core of the New Jersey Alliance for Clinical and Translational Science (2022–)
Board Member of Text as Data Association (2021–)
American Sociological Association Methodology Section Council (2021–2024)

Completed

Editorial Board of *Sociological Methodology* (2022–2024)
American Sociological Association Methodology Awards Committee, Chair (2024)
Political Methodology Conference Program Committee (2024)
American Sociological Association Methodology Awards Committee (2023)
American Sociological Association Methodology Awards Committee, Chair (2022)
Steering Committee First Workshop on Causal Inference & Natural Language Processing, *Empirical Methods and Natural Language Processing* Conference (2021)
New Directions in Text Analysis Program Committee, Stanford University (2019)
Mathematical Sociology Award for American Sociological Association Committee (2018)
Political Methodology Society Gosnell Prize Committee (2016–17)

Conferences Organized:

Co-organizer *New Directions in Text Analysis* Conference, Princeton University (2017)

Discussant:

Polmeth: 2015, 2016, 2017, 2018, 2019, 2020, 2021, 2022, 2023, 2024
Text as Data: 2015, 2016, 2018, 2023
American Sociological Association: 2017
Midwest Political Science Association: 2022

Reviewing (by area):

Interdisciplinary: *Science Advances*, *Proceedings of the National Academy of Sciences*, *Nature Communications*

Political Science: *American Journal of Political Science*, *Conflict Management & Peace Science*, *International Studies Quarterly*, *Political Analysis*, *Political Behavior*, *Research & Politics*, *Public Opinion Quarterly*, *Journal of Peace Research*, *Journal of Politics*, *Political Science Research and Methods*, *Journal of Conflict Resolution*, *British Journal of Political Science*, *International Organization*, *American Political Science Review*, *Political Communication*, *Political Research Quarterly*, *World Politics*, *Political Studies Review*

Sociology and Social Science: *Econometrica*, *American Sociological Review*, *Sociological Methodology*, *Communication Methods and Measures*, *Sociological Methods and Research*, *International Journal of Drug Policy*, *Journal of Policy Analysis and Management*, *Mobilization*, *Journal of Law and Courts*, *European Sociological Review*, Nuttfield Foundation, *Journal of Research on Educational Effectiveness*, *Sociological Science*, National Science Foundation SBE, *Advances in Methods and Practices in Psychological Science*, Swiss National Science Foundation, *Demography*, *Journal of Empirical Legal Studies*, *Data and Policy Review*

Statistics/Computer Science Journals: *Journal of the American Statistical Association*, *Annals of Applied Statistics* *Journal of Machine Learning Research*, *Statistics and Medicine*, *Journal of Artificial Intelligence Research*, National Science Foundation Methodology, Measurement, and Statistics (MMS) section *ad hoc* review, *Journal of the Royal Statistical Society: Series B*, *Journal of the Royal Statistical Society: Series C*, *Journal of Business & Economic Statistics*, Panel review for the National Science Foundation, *Artificial Intelligence Review*, *Journal of Open Source Software*

Computer Science Conferences: *Neural Information Processing Systems* “Topic Models: Computation, Application and Evaluation” Workshop, 2013, *North American Association for Computational Linguistics (NAACL-HLT)* 2015, *Empirical Methods on Natural Language Processing* 2015, *Association of Computational Linguistics* 2016, *Conference on Empirical Methods in Natural Language Processing* 2016, *EMNLP 2016 Workshop on NLP+CSS*, *Ethics in Natural Language Processing* 2017, *EMNLP 2017 Workshop on NLP+CSS*, *Association of Computational Linguistics* 2017, *Empirical Methods on Natural Language Processing* 2017, *Association of Computational Linguistics* 2018, *North American Chapter of the Association for Computational Linguistics - Human Language Technologies (NAACL-HLT)* 2018, *Neural Information Processing Systems* 2018, *Empirical Methods on Natural Language Processing* 2020, *Empirical Methods on Natural Language Processing* 2021, *IC2S2* 2024, *Empirical Methods on Natural Language Processing* 2024

Book Presses: SAGE, Princeton University Press, Chapman & Hall CRC, Oxford University Press, Routledge