

Ian Stewart

Researcher in natural language processing.

2260 Hayward Street
Ann Arbor, MI 48109
✉ ianbstew@umich.edu
📄 ianbstewart.github.io

I apply natural language processing and statistical models to understand social behavior in online discussions. I work to make NLP systems such as writing assistants more aware of the social context in which they operate.

Education

- 2015–2020 **Georgia Institute of Technology**, *Ph.D. in Human-Centered Computing*.
Thesis: "The laws of LOL: computational approaches to sociolinguistic variation in online discussions."
Advised by Jacob Eisenstein and Diyi Yang.
- 2010–2014 **Dartmouth College**, *A.B. Linguistics (cum laude), minor in Computer Science*.
Thesis: "African American English syntax in Twitter."
Advised by James Stanford and Sravana Reddy.

Skills

- Programming languages Java (proficient), Python (proficient), R (intermediate), \LaTeX (intermediate), Javascript/d3 (beginner)
- Programming libraries sklearn, statsmodels, pytorch, gensim, nltk, Beautiful Soup, matplotlib, pandas
- Natural languages English (native), French (fluent), Japanese (beginner), Spanish (beginner), Māori (beginner)

Work experience

- Sep 2020 – present **Postdoctoral Fellow**, UNIVERSITY OF MICHIGAN.
• Working with Rada Mihalcea in the Language and Information Technologies (LIT) Lab at University of Michigan.
• Developing natural language processing models to adapt to different social needs.
• Investigating emergent bias in NLP systems.
• Mentoring graduate students in computational social science research.
- Aug 2015 – Aug 2020 **Graduate Research Assistant**, GEORGIA INSTITUTE OF TECHNOLOGY.
• Developed computational methods to describe and predict language variation in online discussions.
• Applied linguistic analysis to understand social dynamics in real-life scenarios, e.g. breaking news events.
• Adapted statistical models such as causal inference and survival analysis to test theories of sociolinguistic variation.
- May 2018 – Aug 2018 **PhD Fellow**, Max Planck Institute for Demographic Research.
• Leveraged Facebook advertising API to assess immigrant behavior in the US.
• Collected, organized and analyzed large-scale data with statistical models to test hypotheses on immigrant assimilation.
• Engaged with interdisciplinary research team to set up survey for follow-up analysis.
- May 2016 – Aug 2016 **PhD Intern**, Pacific Northwest National Laboratory.
• PhD intern through the National Security Internship Program.
• Formulated and tested methods to predict word dynamics over time in social media data.
• Implemented unsupervised clustering and smoothing to group semantically related words.
• Corroborated connection between distributional semantic meaning (measured with word vectors) and frequency dynamics.

June 2014 – **Software Engineer Intern**, Charles River Analytics.

- Aug 2015
 - Contributed to agent-based social network generation model (in AnyLogic and Java MASON), in collaboration with two scientists.
 - Designed pipeline for social network synthesis and analysis, implementing metrics such as the network clustering coefficient and community modularity.
 - Extracted parameters from large-scale social media data for network generation model, such as distribution of languages across Twitter users.

Teaching

- Jan - May 2020: Natural language processing (CS 7650)
- Sep - Dec 2015: Introduction to artificial intelligence (CS 6601)

Awards

- Sept 2021 **Research Grant**, SUMMER INSTITUTES IN COMPUTATIONAL SOCIAL SCIENCE.
Funding for an open-source tutorial series about advanced natural language processing methods for computational social science.
- March 2017 **Honorable Mention**, NATIONAL SCIENCE FOUNDATION GRADUATE RESEARCH FELLOWSHIP.
Runner-up in annual Graduate Research Fellowship competition. Award includes increased access to supercomputing resources.
- August 2016 **Tuition Reimbursement Award**, PACIFIC NORTHWEST NATIONAL LABORATORY.
Awarded to PhD interns in the National Security Internship Program who demonstrated outstanding performance in their summer research project. Covers one semester of graduate school tuition.
- August 2015 **Great Promise Award**, CHARLES RIVER ANALYTICS.
Awarded to an intern in each company division who shows significant promise as a software engineer.
- May 2014 **Academic Achievement Award**, DARTMOUTH COLLEGE LINGUISTICS DEPARTMENT.
Awarded to a graduating student who has demonstrated considerable commitment to research in linguistics.

Graduate Coursework

- Social Computing
- Design of Online Communities
- Information Visualization
- Natural Language Understanding
- Nonparametric Statistics
- Time Series Analysis

In preparation.

- I. Stewart, C. Welch, R. Mihalcea, K. Resnicow, L. An, J. Pennebaker. Long-term and short-term effects of expressive interviewing on COVID health behavior.
- I. Stewart, R. Mihalcea. How well do you know your readers? Reader-aware question generation.
- S. Castro, R. Wang, P. Huang, I. Stewart, N. Liu, J. Stroud, R. Mihalcea. Fill-in-the-blank as a Challenging Video Understanding Evaluation Framework.

Publications

- 2021 M. Dong, X. Xu, Y. Zhang, I. Stewart, R. Mihalcea. Room to Grow: Understanding Personal Characteristics Behind Self Improvement Using Social Media. SocialNLP Workshop.
- 2021 I. Stewart, D. Yang, J. Eisenstein. Tuitemos o pongamos un tuit? Investigating the Social Constraints of Loanword Integration in Spanish Social Media. SCiL.
- 2020 I. Stewart, D. Yang, J. Eisenstein. Characterizing Collective Attention via Descriptor Context: A Case Study of Public Discussions of Crisis Events. ICWSM.
- 2019 I. Stewart, R. Flores, T. Riffe, I. Weber, E. Zagheni. Rock, rap, or reggaeton? Assessing Mexican immigrants' cultural assimilation using Facebook data. WebConf.

- 2018 I. Stewart, J. Eisenstein. Making “fetch” happen: the influence of social and linguistic context on nonstandard word growth and decline. EMNLP.
- 2018 I. Stewart, Y. Pinter, J. Eisenstein. Si o no, que penses? Catalanian independence and linguistic identity on social media. NAACL.
- 2017 I. Stewart, S. Chancellor, M. De Choudhury, J. Eisenstein. #anorexia, #anarexia, #anarexyia: Characterizing Online Community Practices with Orthographic Variation. IEEE Big Data, SocialNLP Workshop.
- 2017 F. Hohman, S. Soni, I. Stewart, J. Stasko. A Viz of Ice and Fire: Exploring Entertainment Video Using Color and Dialogue. Workshop on Visualization for the Digital Humanities.
- 2017 I. Stewart, D. Arendt, E. Bell, S. Volkova. Measuring, Predicting and Visualizing Short-Term Change in Word Representation and Usage in VKontakte Social Network. ICWSM.
- 2015 E. Stickgold, B. Skarin, I. Stewart, C. Lofdahl. Extending generative models of large scale networks. 24th Conference on Behavior Representation in Modeling and Simulation (BRiMS).
- 2014 I. Stewart. Now we stronger than ever: African American syntax on Twitter. EACL.

Non-archival presentations

- 2021 I. Stewart. Assessing same-gender relationship bias in machine translation. Text as Data Conference.
- 2019 I. Stewart. What natural language processing should do for LGBTQ people. Queer in AI Workshop at NeurIPS. Vancouver, Canada.
- 2017 I. Stewart, J. Eisenstein. Social and Distributional Predictors of the Success of Lexical Innovations in Online Writing. New Ways of Analyzing Variation (NWAV). Madison, WI.
- 2017 I. Stewart, J. Eisenstein. #thighgap to #thyghgapp: Incrementation of Orthographic Variation on Instagram. Diversity and Variation in Language Conference (DiVar1). Atlanta, GA.
- 2015 I. Stewart. We some young kings: Communities, age, and African American English online. 2015 Annual Meeting of the American Dialect Society. Portland, Oregon.

Invited talks

Slides and materials available online.

- 2021 “Topic modeling 101.” Presented for Information Retrieval class at University of Michigan.
- 2020 “Exploratory text analysis for computational social science.” Presented at CORE Congress 2020.
- 2020 “Through the language glass: What NLP can reveal about sociolinguistic variation.” NLP@GT talk.
- 2019 “Language change tutorial.” Presented at IC2S2 in Amsterdam.
- 2019 “Why do you talk like that? Minority languages and politics.” Guest lecture for “Applications of Linguistics” course at Georgia Tech.
- 2018 “Is it social or linguistic? Examining internal factors in language change.” Workshop on New Methods in Computational Sociolinguistics. Leiden, Netherlands.

Reviewing

* indicates outstanding reviewer.

- 2022 ICWSM.
- 2021 CHI, ICWSM, EACL, NAACL, EMNLP, W-NUT.
- 2020 ICWSM, ACL, EMNLP*, W-NUT.
- 2019 ICWSM, NAACL*, ACL, Language Change workshop, W-NUT.
- 2018 ICWSM, NAACL, EMNLP, Workshop on Noisy User-generated Text (W-NUT).

Service

- 2021 Co-organizer for computational social science tutorial series (“NLP+CSS 201”).
- 2021 Judge for Undergraduate Research Opportunity Program Symposium.
- 2019 Co-leader for Human-Centered Computing seminar (weekly discussion for PhD students).
- 2018-2019 Co-Chair of Graduate Student Council Travel Fund Committee.
- 2018 Reviewer for President’s Undergraduate Research Awards (PURA) applicants.

Mentoring

- 2021 Aman Choudhury: Undergraduate (University of Michigan); social media analysis.
- 2021 Aylin Gunal: Undergraduate (University of Michigan); counseling message analysis.
- 2021 Dillan Pribak: Undergraduate (University of Michigan); story generation.
- 2021 Yiting Shen: Master’s student (University of Michigan); analysis of bias in text.
- 2018 Xiaochuang Han: Undergraduate (Georgia Institute of Technology); deep learning for entity resolution.

Interests

- Music making
- Language learning
- Cookie baking
- Letter kerning