Nathan Bartley

Los Angeles, CA ● 310-773-1145 ● ntbartley@gmail.com

EDUCATION

Ph.D. (Computer Science), **University of Southern California**, 2024
Dissertation: Measuring and Mitigating Exposure Bias in Online Social Networks
Committee: Kristina Lerman (chair), Fred Morstatter, Emilio Ferrara, Mike Ananny

M.Sc. (Computer Science), **University of Chicago**, 2015

B.A. (General Honors) (Biology and Computer Science), **University of Chicago**, 2013

RESEARCH EXPERIENCE

Information Sciences Institute, University of Southern California — Los Angeles, CA
Graduate Researcher — August 2017 – Present

- Led a multi-year project analyzing biases in ranking algorithms on online social networks
  - Measured impact of ranking algorithms on exposure bias on 4 million Twitter users (ASONAM ‘24)
  - Ran a multi-year audit on X/Twitter to analyze exposure to pro-/anti- science (ICWSM ‘24)
  - Used agent-based modeling with 170K agents to test dynamics of exposure under various timeline algorithms (SIGIR ‘24)
- Used distributed word embeddings to analyze 900,000 Pastebin pastes and identify malicious pastes as part of Defence
- Analyzed time series data of Google query trends and geospatial Uber/rideshare adoption in cities
- Analyzed performance of classifier built on top of word embedding models
- Learned embeddings of nodes based on information diffusion cascades

Computation Institute, University of Chicago — Chicago, IL
NLP/Machine Learning Engineer — June 2015 – June 2017
Research Assistant — August 2013 - June 2015

- Analyzed 100,000 scientific articles with distributed word embeddings
- Developed new database schema for scientific publication repositories
- Built a Django website to hook into Amazon MechanicalTurk to run custom experiments
- Managed student research assistants
- Pulled, processed, and analyzed publication data for understanding non-profit research grants and their influence in establishing new areas of research

PROFESSIONAL EXPERIENCE

Office of Safety, Health and Working Conditions, Bureau of Labor Statistics — Remote
Civic Digital Fellow — May – August 2021

- Wrote a Google Colab tutorial for making use of a Huggingface Q&A model
- Automated the extraction of information from death certificate data and matching to an existing database

PROJECT DESCRIPTIONS

Mitigation of Exposure Bias in Online Social Networks *Professional*
- Utilizing agent-based models to assess how different recommendation algorithms affect bias over time

Content Exposure Bias in Online Social Networks *Professional*
- Comparing several metrics in empirical data and simulated social network data to assess impact of recommendation algorithms on different measures of bias

Auditing Bias in Twitter *Professional*
- Using webscraping techniques to gather actual information about how timelines bias exposure to social networks

Learning Node Embeddings from Information Diffusion Cascades *Professional*
- Using the structural information from diffusion cascades to reconstruct underlying social networks

The Network Evolutionary Forces of Online Gaming Communities *Professional*

Nathan Bartley

Los Angeles, CA ● 310-773-1145 ● ntbartley@gmail.com

- Using network features derived from co-play and friendship networks from *World of Tanks* to assess how clans behave and evolve over time

Using Social Role Detection in Online Cyberbullying Coursework
- Using linguistic and network features from Instagram and ASK.fm comments to identify and characterize cyberbullying

Homeless Youth Protective Behaviors Coursework
- Using network features and behavior features derived from homeless youth surveys to understand effects of group norms on risky sexual behavior

Academic Discourse Coursework
- Using topic modeling and function-word analysis to characterize signatures of influential authors with full-text and abstracts of scientific publications from Elsevier and JSTOR.
- Analyzing how interdisciplinary influential authors are over time, and how they relate to other authors within their discipline

Smith Influence *Professional*
- Using word embedding models and computational linguistics to understand the influence of seminal social science works and the dynamics of field development around these seminal works.

Portal Project *Professional*
- Using word embedding models and topic models to match potential grants to author profiles, and likewise matching authors to potential collaborators
- Managed the backend design of the process that serves the match information to the web server hosting the site

Keck Grants *Professional*
- Calculating historical h-index and basic bibliometrics to predict/rank likelihood of faculty of winning Keck grants based on historical keck grant data

Anatomy of a Research Paper *Professional*
- Using Word2Vec and similar word embedding algorithms to embed 100,000 scientific articles
- Analyzing distance matrices of these embeddings to track and characterize the conventions of scientific discourse in different disciplines

Tree-based Pipeline Optimization Tool *Open-Sourced Side Project*
- Implemented ability to pass in abstract scoring functions
- Coordinated early experiments for testing on HPC clusters at UPenn

Latin Object Annotator *Professional*
- Stage 1: made use of scikit-learn implementations of Naïve Bayes and Perceptrons as classifers for identifying individual categories of Latin n-grams
- Stage 2: Implementing a GUI in Java and using WEKA implementations of classifier chains to classify multiple targets, each of which is a multi-label class

Agent-based Modeling of *In-Vitro* Microbial Ecology *Professional*
- Built NetLogo agent-based models of bacteria growing in various *in-vitro* environments to compare against existing mathematical models
- Made use of the NetLogo API to write Java programs to automate the simulations and data generation
- Wrote custom visualization scripts in Python for visualizations that are non-trivial in Excel

Bioinformatics (and Big Data) *Class*
- Implemented scripts on Google Compute Engine to download PDB files, parse them into Thrift structures, serialize them into HDFS, and process them using complex Pig and Cassandra queries to hold results in Cassandra column families.

Advanced Data Analysis *Class*
- Implemented a Naïve Bayesian classifier and k-Means clustering algorithm using MRJob scripts on AWS EMR instances

The Topology of GitHub Repositories *Class*
- Built a pipeline to filter and explore characteristics of active public GitHub repositories
- With a subset of active repositories, exploring clustering and NLP techniques to induce the problem that the codebase is addressing

JOURNAL PAPERS

- N. Bartley, K. Burghardt, K. Lerman. Measuring the Echo-chamber Phenomenon Through Exposure Bias. *Edited Book in the Lecture Notes in Social Networks*. 2024.
- Li, Yiqi, Nathan Bartley, Jingyi Sun, and Dmitri Williams. "The larger, the fitter, the better: clans' evolution, social capital and effectiveness." *Internet Research* 33, no. 3 (2023): 1053-1078.

Nathan Bartley

Los Angeles, CA ● 310-773-1145 ● ntbartley@gmail.com

- Palash Goyal, Tozammel Hossein, Ashok Deb, Nazgol Tavabi, Nathan Bartley, Andres Abeliuk, Emilio Ferrara, and Kristina Lerman. Discovering Signals From Web Sources to Predict Cyber Attacks. (*Submitted)*
- Nathan Bartley and James Evans. Anatomy of a Research Article (I*n preparation*)

CONFERENCE PAPERS AND PRESENTATIONS

- Bartley, N., Burghardt, K., & Lerman, K. (2024). Impacts of Personalization on Social Network Exposure. *ASONAM 2024*. (*accepted*)
- Bartley, N., Burghardt, K., & Lerman, K. (2024). Bias Reduction in Social Networks through Agent-Based Simulations. SIGIR. (*accepted*)
- Bartley, N., Burghardt, K., & Lerman, K. (2024). Auditing Exposure Bias on Social Media for a Healthier Online Discourse. ICWSM. https://doi.org/10.36190/2024.12
- Bartley, N., Burghardt, K., & Lerman, K. Evaluating Content Exposure Bias in Social Networks. Ph.D. Symposium at ASONAM 2023.
- Bartley, N., Abeliuk, A., Ferrara, E., & Lerman, K. Auditing Timeline Bias on Twitter. *In 13th ACM Web Science Conference 2021* (WebSci ‘21).
- Tavabi, N., Bartley, N., Abeliuk, A., Soni, S., Ferrara, E., & Lerman, K. (2019, May). Characterizing Activity on the Deep and Dark Web. *In Companion Proceedings of The 2019 World Wide Web Conference* (pp. 206-213). ACM.
- Randal S. Olson, Nathan Bartley, Ryan J. Urbanowicz, and Jason H. Moore (2016). Evaluation of a Tree-based Pipeline Optimization Tool for Automating Data Science. *Proceedings of GECCO 2016*, pages 485-492.
- Nathan Bartley, Ricardo Colasanti, Christopher Henry. Estimating Kinetic Growth Parameters of Microbial Communities with Agent-Based Models. (*In preparation)*

ABSTRACTS AND POSTERS

- Bartley, N., Burghardt, K., & Lerman, K. Mitigating Exposure Biases in Personalized Timelines through Agent-based Models. *ICML 2024 Workshop on Humans, Algorithmic Decision-Making and Society: Modeling Interactions and Impact*. Poster.
- Bartley, N., Burghardt, K., & Lerman, K. Measuring Impact of Personalized Timelines in Agent-based Simulations of Social Networks. *IC2S2 2024*. Poster.
- Bartley, N., Burghardt, K., & Lerman, K. Majority Illusion in Partially-observed Personalized Timelines. Conference on Network Science 2024. Presentation.
- Nathan Bartley. Proxy Servers for Facilitating Data Donations. *Post-API Conference II*. 2024
- Nathan Bartley. Social Network Timeline Bias and Amplification. *User AI Auditing Workshop at CSCW 2023*.
- Nathan Bartley, Keith Burghardt, Kristina Lerman. Social Network Timeline Exposure and Attention Bias. *Politics & Computational Social Science pre-conference at APSA 2023*.
- Nathan Bartley, Rachel Moran, ASM Rizvi, Sarah Myers West. Technical Approaches to Combating “Fake News”: Assessing Current Typologies of Information Disorder. *USC Annenberg Graduate Fellowship Research and Creative Project Symposium 2018*. Poster.
- Andres Abeliuk, Nathan Bartley, Kristina Lerman. Uber's Impact on Flu Infections: An Exploratory Analysis. *International Conference on Computational Social Science 2018*. Poster.
- Matt Baucum, Nathan Bartley, Mohammad Atari, Andreas Aristridou (2018). Immortal words of the Prominent Researchers: Semantic Analysis of Academic Discourse. *Society for Personality and Social Psychology 2018*. Poster.

AWARDS

USC Viterbi School Best Viterbi Mentor Award Spring 2019

- Chosen by faculty as best mentor for incoming PhD and MS students in the Viterbi School of Engineering

Nathan Bartley

Los Angeles, CA ● 310-773-1145 ● ntbartley@gmail.com

USC Annenberg Graduate Fellowship Research and Creative Project Symposium Spring, 2018 & 2019

- Granted to students performing creative work pertaining to communications, within the School of Cinematic Arts, Annenberg School for Communication & Journalism, and the Viterbi School of Engineering.

Computer Science Graduate Student Teaching Prize

Computer Science Department Spring, 2015

- Granted to teaching assistants who were nominated by the students taking the course and accepted by the faculty within the Computer Science Department
- Nominated for the Physical Sciences Division-wide best Annual Teaching Assistant award

TEACHING EXPERIENCE

University of Southern California

Teaching Assistant, Fundamentals of Computation Fall 2019, Spring 2020, Fall 2021, 2023

Teaching Assistant, Machine Learning for Data Science Summer 2024

Teaching Assistant, Introduction to Programming Fall 2022, Spring 2023

Teaching Assistant, Data Structures and Object-Oriented Design Summer 2020, 2022, 2023, Spring 2024

- Taught a discussion section, held office hours, managed undergraduate graders and administrivia, proctored and helped grade exams.

University of Chicago

Teaching Assistant, Networks and Distributed Systems Fall 2014

- Taught a discussion section, held office hours, managed undergraduate graders and administrivia, proctored and helped grade exams.

OTHER EXPERIENCE

ACM WSDM 2018, University of Southern California Los Angeles, CA

Student Volunteer February 5 – 9 2018

Checked attendees into the conference, and directed speakers to their appropriate locations

SKILLS

Proficient in: C, Python, Java SE/EE, MySQL
Familiar with: C++, MATLAB, R, HTML & CSS, Hive, Jcascalog, Scala, Haskell, Pig
Computational: Classification, Regression, MapReduce, NLP, Text Mining, Big Data, Visualization
Analytical: Basic Numerical Analysis, Exploratory & Predictive Modeling, Experimental Design
Languages: Spanish (proficient), Japanese (intermediate)