George Berry

george.e.berry@gmail.com
https://georgeberry.github.io
January 3, 2021

education

Ph.D., Sociology, Cornell University, 2020

Committee: Michael Macy (chair), Benjamin Cornwell, Steven Strogatz
Dissertation: "Refining the telescope: Addressing two challenges to sociological research in online social networks"

M.A., Sociology, Cornell University, 2015

B.A., English, Oberlin College, 2011

work

Data scientist and machine learning engineer, Civis Analytics political R&D team, October 2019—January 2021

- Built core Bayesian modeling infrastructure on top of Tensorflow Probability, which allowed scalable hierarchical Bayesian models with a GLM flavor. This infrastructure powered Civis' core products (forecasting, ad testing) throughout the 2020 general election.
- Built a widely used automated system for producing Bayesian estimates of political advertising, which tested over 2400 presidential ads and 600 Senate ads in the 2020 general election. Turnaround times were on the order of one day. Guidance given from this infrastructure was impactful, guiding hundreds of millions of dollars of ad spend.
- Built a widely used ad-hoc system for producing Bayesian estimates of political advertising, which was employed across progressive campaigns throughout the 2020 primary and general elections.

Data scientist and tech lead, Twitter, June 2019—October 2019

- Partner with product teams to help them understand how social science theory and methods can benefit their ongoing work, with a focus on platform health (anti-abuse)
- Develop evaluation strategies for complex product changes where experiment suffer from interference

Research scientist, Facebook, Core Data Science: Core Statistics/Interaction Science, Sept 2017—Feb 2019

- **Primary focus**: Public conversations. I developed a framework based on graph theory for better measuring outcomes in large conversations, which is now a frequently used measurement tool at Facebook. I worked with the Public Conversations and News teams to understand their

needs and goals in developing this framework. After building a prototype, I productionized the framework and partnered with UX researchers to conduct surveys and tests to deepen our understanding of conversations.
- **Secondary focus**: Election integrity. Co-built integrity monitoring framework widely used in 2017 and 2018 elections.

Intern, Facebook, CDS—Methods, Summer 2016
- Developed a machine learning method for understanding heterogeneous treatment effects in experiments.
- Presented work at the 2016 Conference on Digital Experimentation.

Intern, Facebook, CDS—Economic Research, Summer 2015
- Conducted research on how Facebook's comment ranking algorithm was affecting online discussions.
- Developed and implemented a field experimentation methodology which allowed answering the question of how ranking was affecting individuals.
- Built a machine learning model based on 100,000 labeled comments to predict the "quality" of a comment.
- Wrote a paper on this research, published at WWW2017.

Intern, Graphika, Summer 2014

publications and preprints

George Berry, Antonio Sirianni, Ingmar Weber, Jisun An, and Michael Macy. "Going beyond accuracy: estimating homophily in social networks using predictions". 2020. arXiv:2001.11171.

George Berry, Christopher J. Cameron, Patrick Park and Michael W. Macy. "The Opacity Problem in Social Contagion". 2019. *Social Networks*. arXiv:1702.02700.

George Berry. "Role action embeddings: scalable representation of network positions". 2018. arXiv:1811.08019.

Michael Genkin, Cheng Wang, **George Berry**, Liuyuan Chen and Matthew E. Brashears. 2018. "Blaunet: An R-based graphical user interface package to analyze Blau space". *PloS One*. *R package*.

George Berry, Antonio Sirianni, Nathan High, Agrippa Kellum, Ingmar Weber, and Michael Macy. "Estimating group properties in online social networks with a classifier". 2018. *SocInfo*. arXiv:1807.09406.

George Berry and Sean J. Taylor. "Discussion quality diffuses in the digital public square". 2017. *WWW*. arXiv:1702.06677.

conference presentations and invited talks

"Going beyond accuracy: estimating homophily in social networks using predictions". *ASA Regular Section on Mathematical Sociology* 2020.

"Discussion quality diffuses in the digital public square". *Invited talk at SUNY Stony Brook.* 2019.

"The Opacity Problem in Social Contagion", with Christopher J. Cameron. *ASA Section on Social Networks: New Methods and Measures* 2017.

"Evolving cooperation in social structures". *Cornell Graduate Research Symposium* 2017, *ASA Section on Rationality & Society* 2017.

"Discussion quality diffuses in the digital public square". *WWW2017.*

"Machine learning for policy recommendations in sociology". *Cornell SGSA Seminar* 2016.

"Two-stage: Find and summarize CATEs in experiments". *Conference on Digital Experimentation* 2016.

"Correctly measuring social contagion". *INSNA Sunbelt* 2016, *Cornell Sociology Symposium* 2016.

"Effects of ranking in online discussions". *Conference on Digital Experimentation* 2015.

honors and awards

Research Assistantship. "A New Infrastructure for Monitoring Social Class Networks". P.I. David Grusky and Michael Macy. 2015-6.

Teaching Assistantship. "Contemporary Sociological Theory" with David Strang. Spring 2014.

Teaching Assistantship. "Controversies about Inequality" with Steven Morgan. Fall 2013.

Sage Fellowship. Cornell University, 2012-3 and 2016-7.

programming languages

General purpose: Python

Statistical: R

Database: SQL

professional service

Reviewer. NetSci 2019-present.

Reviewer. WebSci 2019-present.

Reviewer. *IC2S2* 2019-present.

Reviewer. *SocInfo* 2019-present.

Reviewer. *Rationality & Society* 2017-present.

Reviewer. *WWW* 2017-present.

Reviewer. *ICWSM* 2016-present. Best reviewer 2019 award.

Co-chair. Cornell Sociology Graduate Student Association, 2014.

open source contributions

Natural Language Toolkit (NLTK): Rewrote algorithm for computing Vader Sentiment.

Blaunet: Primary programmer on open source R package