

STATEMENT OF PURPOSE

I bring statistics and machine learning together with critical perspectives from social science to consider when, how, and why data and modeling succeed in their aims—and when, how, and why they can fail. I am passionate about improving practice towards more responsible, robust, effective, and just uses of data and modeling, as well as engaging in outreach to help practitioners in policy, government, law, journalism, social science, business, civil society, and elsewhere understand and adopt machine learning and data science.

EDUCATION

School of Computer Science, Carnegie Mellon University, Aug 2013–Aug 2018

PhD in Societal Computing (Institute for Software Research) and MS in Machine Learning (Machine Learning Department). Research on social media and sensor data. ARCS Foundation award. Dissertation: “Bias and beyond in digital trace data.” Committee: Jürgen Pfeffer and Anind K. Dey (Coadvisors), Cosma R. Shalizi (Department of Statistics), and David Lazer (Northeastern University).

Oxford Internet Institute, University of Oxford, Oct 2011–Sep 2012

MSc with distinction in Social Science of the Internet. Master’s thesis on locating the emergence of Internet studies in the 20-year evolution of a large co-authorship network. Advisor: Eric T. Meyer.

Department of the History of Science, Harvard University, Sep 2004–Mar 2009

AB cum laude in History and Science with Music minor. Senior thesis on narratives of the early 20th century South Indian mathematician Srinivasa Ramanujan. Advisor: Lukas Rieppel.

WORK AND RESEARCH EXPERIENCE

Data Science Postdoctoral Fellow, Berkman Klein Center for Internet & Society at Harvard University, Cambridge, MA, Sep 2018–present

Statistical learning and network modeling for data in Media Cloud and other projects; technical advising and tutoring for scholars and practitioners in social science, journalism, law, and policy; building in-house data science capacity; forging stronger university-wide ties between social and data science; and research into conceptual, ethical, theoretical, and practical challenges of deploying data science.

Data Science for Social Good Fellow, Lisbon, Portugal, Summer 2017

Project for Tuscan agencies applying machine learning and network analysis to urban data for sustainable tourism in Florence. Run by the Center for Data Science and Public Policy, University of Chicago, and Nova School of Business and Economics, Universidade NOVA de Lisboa.

Graduate research assistant, UbiComp Lab, Pittsburgh, PA, 2016–2018

Research under Dr. Anind Dey (Human-Computer Interaction Institute and University of Washington) on mobile phone sensor collection of social network data. Ran 3-month 53-subject study, collecting mobile phone sensor data alongside social network survey responses.


Graduate research assistant, Pfeffer Lab, Pittsburgh, PA, 2013–2018

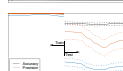
Research under Dr. Jürgen Pfeffer (Institute for Software Research and Technical University of Munich) on biases in social media data (representativeness, data access, effect of platform constraints on behavior), and news media and social media. Use of Twitter 10% sample.


Research assistant, *Big Data: A Revolution That Will Transform How We Live, Work, and Think*, Oxford, UK, 2011–2012

Research for 2013 book by Prof. Viktor Mayer-Schönberger (Oxford Internet Institute) and Kenneth Cukier (*The Economist*). History of statistics and data management, historical and modern case studies, fact-checking, idea development, and editing draft material.

SELECTED PUBLICATIONS

 **Momin M. Malik.** (in submission, 2020). A hierarchy of limitations in machine learning. *Frontiers in Big Data*. Preprint at <https://arxiv.org/abs/2002.05193>.

 Kar-Hai Chu, Jason Colditz, **Momin M. Malik**, Tabitha Yates, and Brian Primack. (2019). Identifying key target audiences for public health campaigns: Leveraging machine learning in the case of hookah tobacco smoking. *J. Med. Internet Res.*, 21 (7), e12443.

 Jürgen Pfeffer and **Momin M. Malik.** (2017). Simulating the dynamics of socio-economic systems. In Betina Hollstein, Wenzel Matiaske, & Kai-Uwe Schnapp (Eds.), *Networked governance: New research perspectives*, pp. 143–161. Cham, Switzerland: Springer.

 **Momin M. Malik** and Jürgen Pfeffer. (2016). Identifying platform effects in social media data. In *Proceedings of the Tenth International AAAI Conference on Web and Social Media (ICWSM-16)*, pp. 241–249.

 **Momin M. Malik** and Jürgen Pfeffer. (2016). A macroscopic analysis of news content on Twitter. *Digital Journalism*, 4(8), 955–979.

 Hemank Lamba, **Momin M. Malik**, and Jürgen Pfeffer. (2015). A tempest in a teacup? Analyzing firestorms on Twitter. In *Proc. of the 2015 Conf. on Advances in Social Networks Analysis and Mining (ASONAM 2015)*, pp. 17–24. **Best Student Paper Award.**

 **Momin M. Malik**, Hemank Lamba, Constantine Nakos, and Jürgen Pfeffer. (2015). Population bias in geotagged tweets. In *Papers from the 2015 ICWSM Workshop on Standards and Practices in Large-Scale Social Media Research (ICWSM-15 SPSM)*, pp. 18–27.

SKILLS

