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, public health, social work, law, government, journalism, social science, industry, 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 and MS in Computation, Organizations and Society (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

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

WORK AND RESEARCH EXPERIENCE

Director of Data Science, Avant-garde Health. Boston, MA, Nov 2020–present
Leading a small team of data scientists at a healthcare data startup, born out of Harvard Business School’s value-based healthcare research and serving about a dozen client hospitals. Research into value-based healthcare; predictive modeling to support bundling payments for Medicare and Medicaid patients; integrating analysis techniques into the product platform.

Analytics support for progressive organizing efforts around the US 2020 General Election. Tracking disinformation and digital voter suppression targeting Black and minority voters; designing and implementing voter mobilization experiments around ballot drop box locations.

Data Science Postdoctoral Fellow, Berkman Klein Center for Internet & Society at Harvard University. Cambridge, MA, Sep 2018–Aug 2020
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, Jun 2016–Aug 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, Aug 2013–Jun 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 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.

Research for Dr. Urs Gasser (Executive Director, Berkman Klein Center) on ‘Youth and Media’ project. Multidisciplinary review of youth information-seeking; policy reports; focus group design; analysis of qualitative data on youth social media usage; theory of ‘news literacy.’

PUBLICATIONS

Momin M. Malik

PUBLICATIONS, continued


REPORTS


AFFILIATIONS

Affiliate, Berkman Klein Center for Internet & Society at Harvard University, 2020-present.

Fellow, Institute in Critical Quantitative, Computational, & Mixed Methodologies (ICQCM), 2021-present.

BLOGGING

TEACHING


“Everything you ever wanted to know about network statistics but were afraid to ask.” 3-hour tutorial. XXXIX Sunbelt Social Networks Conference of the International Network for Social Network Analysis (Sunbelt 2019). June 18, 2019, Montreal, Quebec.

Future Faculty Program (completed), Eberly Center for Teaching Excellence and Educational Innovation, Jan 2014 – Mar 2018
Enrolled since first year of PhD. Two teaching feedback consultations. Nine seminars on topics like student motivation and engagement, discussion facilitation, lecture design and delivery, course and syllabus design, teaching effectiveness monitoring.

Teaching Assistant, Ethics and Policy Issues in Computing (Instructor: Dr. Jim Herbsleb), Spring 2016
Graded weekly response papers and activity on online class forum. Gave a lecture overview of rhetoric and debate in preparation for assigned one-on-one debates. Provided guided and visual design tutorial for scientific poster making. Helped students with course projects (scope, references) and posters. Guest lecture: “Ethical and policy issues in predictive modeling.”

Teaching Assistant, Introduction to Network Science (Instructor: Dr. Jürgen Pfeffer), Fall 2015

PRESENTATIONS


“A hierarchy of limitations in machine learning: Data biases and the social sciences” (invited webinar). Webinar Series: Data Cultures in Higher Education, Faculty of Psychology and Education, Universitat Oberta de Catalunya (Open University of Catalonia). September 29, 2020, Barcelona, Spain [delivered online].


“Revisiting ‘all models are wrong’: Addressing limitations in big data, machine learning, and computational social science” (invited talk). Wednesdays@NICO Seminar Speaker Series, Northwestern Institute for Complex Systems, Northwestern University. February 5, 2020, Evanston, Illinois.


“Interpretability is a red herring: Grasping with ‘prediction policy problems.’ 17th Annual Information Ethics Roundtable: Justice and Fairness in Data Use and Machine Learning. April 5, 2019, Northeastern University, Boston, Massachusetts.

PRESENTATIONS, continued

Reference Library, Toronto, Canada.
““The ethical implications of technical limitations” (invited talk). Fairness, Accountability & Transparency/Asia. Digital Asia Hub and ACM/FAT*. January 12, 2019, Shun Hing College, University of Hong Kong, Hong Kong.
“Machine learning for social scientists.” Fairness, Accountability & Transparency/Asia. Digital Asia Hub and ACM/FAT*. January 11, 2019, Shun Hing College, University of Hong Kong, Hong Kong.
“Theorizing sensors for social network research.” Computational Social Science Institute, UMass Amherst. December 7, 2018, Amherst, Massachusetts.
“What everyone needs to know about ‘prediction’ in machine learning.” Leverhulme Centre for the Future of Intelligence, University of Cambridge. December 3, 2018, Cambridge, UK.
“Anxiety, crisis, and a computational future for journalism.” Philip Merrill College of Journalism / College of Information Studies, University of Maryland. November 27, 2018, College Park, Maryland.
“From the forest to the swamp: Modeling vs. implementation in data science.” Techtopia @ Harvard University. October 2, 2018, Cambridge, Massachusetts.
“Bias busters @ university workshop: A Carnegie Mellon/Google collaboration to address unconscious bias.” With Carol Friese (Carnegie Mellon University, School of Computer Science), Gerry Katilius (Google Pittsburgh), and Diana Marculescu (Carnegie Mellon University, College of Engineering). 2016 ACM Richard Tapia Celebration of Diversity in Computing Conference (Tapia 2016). September 15, 2016, Austin, Texas.
“Bias busting @ university workshop – A Carnegie Mellon/Google collaboration to address unconscious bias.” With Carol Friese (Carnegie Mellon University, School of Computer Science), Gerry Katilius (Google Pittsburgh), and Diana Marculescu (Carnegie Mellon University, College of Engineering). Women in Engineering ProActive Network Change Leader Forum (WEPAN 2016). June 15, 2016, Broomfield, Colorado.
“Identifying platform effects in social media data.” Tenth International AAAI Conference on Web and Social Media (ICWSM-16). Session I: Biases and Inequalities. May 18, 2016, Cologne, Germany.
“Social media data and computational models of mobility: A review for demography.” 2016 ICWSM Workshop on Social Media and Demographic Research (ICWSM-16 SMDR). May 17, 2016, Cologne, Germany.
PRESENTATIONS, continued


ORGANIZING


VOLUNTEERING/COMMUNITY

Facilitator for Bias Busters @ CMU, 2015-2018

Part of 3-person teams leading inclusivity programming for student/faculty/staff groups of 20-30. Led by Dr. Carol Frieze (director, SCS4ALL and Women@SCS), and created in collaboration with Google Pittsburgh.

Senior program committee, International Conference on Web and Social Media (ICWSM), 2020 – present.

Committee Service


Program committee, International Conference on Computational Social Science (IC2S2), 2017 – present.

Other participation


Peer Review

Communications of the ACM (Association for Computing Machinery), 2021.


LANGUAGES


SKILLS

R
igraph
Python
pandas
scikitlearn
MATLAB
PostgreSQL
Gephi
Pajek
LaTeX
Photoshop
Illustrator
InDesign