

Conference on Fairness, Accountability, and Transparency: Preface

Sorelle A. Friedler
Haverford College

SORELLE@CS.HAVERFORD.EDU

Christo Wilson
Northeastern University

CBW@CCS.NEU.EDU

We are delighted to introduce these proceedings of the inaugural Conference on Fairness, Accountability, and Transparency (FAT*). FAT* was formed to serve as a venue for the increasing volume of work focusing on the fairness, accountability, and transparency of algorithmic systems. The work published here, and presented at the conference on February 23rd and 24th, 2018, comes from a wide array of disciplines and subfields including machine learning, statistics, measurement and security, theoretical computer science, law, policy, philosophy, sociology, and interdisciplinary work touching on many of these fields. The papers include a legal examination of online discrimination in advertising, new methods for making recommendation systems fairer, a case study showing bias in existing computer vision benchmarks and introducing a new, fairer, dataset, and many other contributions from the areas of interpretability in machine learning, explainability of algorithmic systems, and examinations of fairness in socio-technical systems. These papers include both full-length publications as well as extended abstracts that underwent the same reviewing process. Papers were chosen from a pool of 73 submitted papers.

From the seventeen papers in these proceedings, the program committee has chosen two to recognize with Best Paper awards. The first Best Paper award, chosen for its strong and focused technical contribution, goes to *The Cost of Fairness in Binary Classification* by Aditya Krishna Menon and Robert C. Williamson. This paper studies the trade-off between fairness and accuracy in binary classification, quantifying the trade-off and providing an algorithm to achieve it. The Best Paper chosen for its strong technical contribution with an interdisciplinary lens is *A case study of algorithm-assisted decision making in child maltreatment hotline screening decisions* by Alexandra Chouldechova, Diana Benavides-Prado, Oleksandr Fialko, and Rhema Vaithianathan. This paper gives a careful case study and fairness-focused analysis of the application of risk prediction to child protective services deployed in Allegheny County, PA, USA.

In addition to the contributed publications, FAT* featured keynote talks by Latanya Sweeney and Deborah Hellman. Latanya Sweeney is a Professor of Government and Technology in Residence at Harvard University and the Director of the Data Privacy Lab in the Institute of Quantitative Social Science at Harvard. Her work showing the existence of discrimination in Google advertisements, where black-identifying name searches were found to be more likely to trigger advertisements relating to potential arrest records, is foundational to much of the work on fairness in socio-technical systems. Deborah Hellman is the D. Lurton Massee Professor of Law and the Roy L. and Rosamond Woodruff Morgan Professor of Law at the University of Virginia School of Law. Her work focuses on discrimination and equality in the law and her talk at FAT* examined discrimination through a legal and philosophical lens.

The conference also included separately reviewed tutorial submissions and presentations, including hands-on programming-focused tutorials and translation tutorials aiming to explain computer science concepts for a social science audience and vice versa. Information and resources associated with these tutorials can be found at the conference website: <https://www.fatconference.org/2018/>.

Putting on an inaugural conference requires a lot of dedication and hard work from everyone involved. We thank the steering committee for their leadership of and dedication to this growing field, the track chairs, program committee, and external reviewers for their thoughtful review and consideration of the submitted papers, the publication chair for managing these proceedings, and the general chair, Solon Barocas, and local chairs, Amanda Levendowski and Jason Schultz, for making the conference possible.

We hope you enjoy reading these proceedings as much as we have.

Sorelle A. Friedler and Christo Wilson
 FAT* 2018 Program Committee Chairs

Publication Chair Suresh Venkatasubramanian (University of Utah).

Track Chairs *Theory and Security*: Anupam Datta (Carnegie Mellon University) and Aaron Roth (University of Pennsylvania). *Statistics, Machine Learning, Data Mining, NLP, and Computer Vision*: Alexandra Chouldechova (Carnegie Mellon University), Been Kim (Google), Dirk Hovy (University of Copenhagen), and Michael Kearns (University of Pennsylvania). *Programming Languages, Databases, and other Systems*: Aws Albarghouthi (University of Wisconsin - Madison), Michael Ekstrand (Boise State University), Gerome Miklau (University of Massachusetts - Amherst). *Visualization, Human Computer Interaction, and User Studies*: Carlos Scheidegger (University of Arizona), Karrie Karahalios (University of Illinois Urbana-Champaign). *Measurement and Algorithm Audits*: Christian Sandvig (University of Michigan) and Arvind Narayanan (Princeton University). *Law, Policy, and Social Science*: danah boyd (Microsoft Research and the Data & Society Research Institute) and Ryan Calo (University of Washington).

Program Committee Virgilio Almeida (Universidade Federal de Minas Gerais), Solon Barocas (Cornell University), Bettina Berendt (KU Leuven), Matthew Brehmer (Microsoft Research), Carlos Castillo (Eurecat - Technology Centre of Catalonia), Hal Daum (University of Maryland - College Park), Maarten de Rijke (University of Amsterdam), Nicholas Diakopoulos (Northwestern University), Timnit Gebru (Microsoft Research), Sharad Goel (Stanford University), James Grimmelmann (Cornell University), Krishna Gummadi (Max Planck Institute for Software Systems), Sara Hajian (Eurecat - Technology Centre of Catalonia), Aniko Hannak (Central European University), Brent Hecht (Northwestern University), Bill Howe (University of Washington), Toshihiro Kamishima (National Institute of Advanced Industrial Science and Technology), Jon Kleinberg (Cornell University), Joshua Kroll (University of California - Berkeley), Nikolaos Laoutaris (Eurecat and the Data Transparency Lab), Karen Levy (Cornell University), Zachary Lipton (University of California, San Diego), Kristian Lum (Human Rights Data Analysis Group), Piotr Mardziel (Carnegie Mellon University), Alan Mislove (Northeastern University), Jamie Morgenstern (Microsoft Research), Deidre Mulligan (University of California - Berkeley), Safiya Noble (University of Southern California), Alexandra Olteanu (IBM Research), David Robinson (Upturn), Narseo Vallina Rodriguez (IMDEA and ICSI), Andrea Roth (University of California - Berkeley), Salvatore Ruggieri (University of Pisa), Andrew Selbst (Data & Society Research Institute), Amit Sharma (Microsoft Research), Sameer Singh (University of California - Irvine), Luke Stark (Dartmouth College), Julia Stoyanovich (Drexel University), Katherine Strandburg (New York University), Michael Strube (Heidelberg Institute for Theoretical Studies), Michael Tschantz (International Computer Science Institute - Berkeley), Berk Ustun (Harvard University), Suresh Venkatasubramanian (University of Utah), Sandra Wachter (University of Oxford), Adrian Weller (University of Cambridge), and Tal Zarsky (University of Haifa).

External Reviewers Julius Adebayo (Google) and Piotr Sapiezynski (Northeastern University).