HUMANITY CENTERED ROBOTICS INITIATIVE

BROWN

Sorelle Friedler
Haverford College

"Algorithmic Fairness: Guaranteeing fairness and non-discrimination in machine-learned decision making"

Thursday, March 9, 2017
12-1pm
CIT 477 Lubrano Conference Room

Abstract: Recidivism prediction algorithms incorrectly label more African Americans high risk. Google shows ads for high paying jobs to men more often than to women. Increasingly, we are seeing instances where machine-learning algorithms reflect and reinforce systemic societal biases. How do we know if these decisions are fair or discriminatory - what does it mean for an algorithm to be fair? In order to answer that question, we will examine societal notions of fairness and non-discrimination from philosophy and political science and explain how these notions can be defined using a mathematical framework. This will then allow us to formally reason about fairness in algorithms. I will talk about the growing landscape of research in algorithmic fairness: how we can reason systematically about biases in algorithms, and how we can make our algorithms fair(er).

Sorelle Friedler is an Assistant Professor of Computer Science at Haverford College and an Affiliate at the Data & Society Research Institute. Her research interests include the design and analysis of algorithms, computational geometry, data mining and machine learning, and the application of such algorithms to interdisciplinary data. Sorelle is one of the organizers of the Workshop on Fairness, Accountability, and Transparency in Machine Learning (fatml.org) and has received a Fellowship and recent NSF grant for her work on preventing discrimination in machine learning. Her work on this topic has been featured in IEEE Spectrum, Gizmodo, and NBC News and she has been interviewed about algorithmic fairness by the Guardian, Bloomberg, and NPR. Before Haverford, Sorelle was a software engineer at Alphabet (formerly Google), where she worked in the X lab and in search infrastructure. She holds a Ph.D in computer science from the University of Maryland, College Park.

Host: Seny Kamara/HCRI