

Adapting for AI: Complexities of AI Accountability

Sarah Cen, Stanford University

Abstract

As AI becomes increasingly integrated into both the private and public sectors, challenges around AI safety and policy have arisen. There is a growing, compelling body of work around the legal and societal challenges that come with AI, but there is a gap in our rigorous understanding of these problems. In this talk, I dive deep into a few topics in AI safety and policy. We will discuss AI supply chains (the increasingly complex ecosystem of AI actors and components that contribute to AI products) and study how AI supply chains complicate machine learning objectives. We'll then shift our discussion to AI audits and evidentiary burdens in cases involving AI. Using Pareto frontiers as a tool for assessing performance-fairness tradeoffs, we will show how a closed-form expression for performance-fairness Pareto frontiers can help plaintiffs (or auditors) overcome evidentiary burdens or a lack of access in AI contexts.



Sarah Cen is a postdoc at Stanford University and incoming Assistant Professor at Carnegie Mellon University's Departments of ECE & EPP. At Stanford, Sarah works with Prof. Percy Liang in Computer Science and Prof. Daniel Ho in the Stanford Law School. Her research is interdisciplinary and inspired by works in machine learning, economics, law, and policy. She has ongoing work on algorithmic auditing, AI supply chains, due process for AI determinations, risk under the EU AI Act, and formalizing trustworthy algorithms. Previously, Sarah received her BSE in Mechanical Engineering from Princeton University and Master's in Engineering Science (Robotics) from Oxford University, where she worked on autonomous vehicles.

When? Friday, March 28th, 3 pm

Where? InformatiKOM 1 (50.19), seminar room 1