# Social foundations for statistics and machine learning

## Opening remarks

Maximilian Kasy

Department of Economics, University of Oxford

May 23, 2023

Motivation for the workshop

Speakers and schedule

Logistics

# Crises of statistics and machine learning in the news

**Science has been in a “replication crisis” for a decade. Have we learned anything?**

**Paperclip-making robots 'wipe out humanity' in killer AI Doomsday experiment**

**Big-Data Algorithms Are Manipulating Us All**

# Single agent decision theory and societal challenges

- Current foundation for both statistics and machine learning (ML):
  Single-agent decision theory.
- This framework cannot address important scientific and societal challenges:
  1. Replication crisis, publication bias, p-hacking, pre-registration,
     reforms of statistics teaching and the publication system.

     *A single agent has no reason to mislead themselves / selectively report!*
  2. The social impact of AI,
     algorithmic discrimination and inequality,
     value alignment of autonomous agents / robots.

     *A single agent has no distributional conflicts / value misalignment!*
- Multiple agents have different objectives and information.

# Science and technology are social (not single agent) activities!

- This is well understood in the philosophy, sociology, and history of science.
- But how to turn this insight into formal, prescriptive recommendations?
- Possible contributions of economics?
  - We share the languages of constrained optimization and probability theory with statistics and ML.
  - But we are also used to considering multiple agents with unequal endowments, conflicting interests, private information.
  - Use the toolkit of mechanism design to characterize optimal statistical decisions subject to constraints of implementability.
  - Use the tooolkit of causal inference and welfare economics to analyze the social impact of algorithmic decisions, and devise alternative algorithm objectives.

Motivation for the workshop

Speakers and schedule

Logistics

# Our speakers

- Isaiah Andrews (Econometrics)
- Celestine Mendler-Dünner (Computer Science)
- Lily Hu (Philosophy)
- Carina Prunkl (Philosophy)
- Jann Spiess (Econometrics)
- Ana-Andreea Stoica (Computer Science)

# Workshop schedule part 1: Tutorial lectures

- Monday, May 22
  - 12:00 Carina Prunkl: *Algorithms and social epistemology*
  - 14:15 Celestine Mendler-Dünner: *Performative Prediction*
  - 16:00 Jann Spiess: *Integrating machine learning into pre-analysis plans*
- Tuesday, May 23
  - 11:30 Lily Hu: *Causal Inference and the Problem of Variable Choice*
  - 14:15 Isaiah Andrews: *Correcting for Selective Publication and Attention*
  - 16:00 Ana-Andreea Stoica: *Diagnosing and mitigating bias in networks*
- Wednesday, May 24
  - 11:30 Scheduled: Nika Haghtalab. Substitute: Maximilian Kasy: *Optimal Pre-Analysis Plans: Statistical Decisions Subject to Implementability*

# Workshop schedule part 2: Frontier talks

- Wednesday, May 24
  - 16:00 Ana-Andreea Stoica: *New models and insights in network interference problems*
- Thursday, May 25
  - 11:30 Carina Prunkl: *Noise - a flaw in algorithmic judgment?*
  - 14:15 Jann Spiess: *Explanations with a purpose: regulating black-box algorithmic decisions*
  - 16:00 Celestine Mendler-Dünner: *Algorithmic Collective Action in ML*
- Friday, May 26
  - 11:30 Scheduled: Nika Haghtalab. Substitute: Maximilian Kasy:

    *Adaptive maximization of social welfare*
  - 14:15 Isaiah Andrews: *A Model of Scientific Communication*
  - 16:00 Lily Hu: *Do Causal Diagrams Assume a Can Opener?*

Motivation for the workshop

Speakers and schedule

Logistics

# Logistics

- For everyone:
  - Concluding panel discussion: Saturday, May 27, 10:00.
  - Coffee and pastries: Outside SR C.
  - Pub evening: Wednesday, May 24, 5:30pm, King's Arms.
- For speakers:
  - Lunch in the department: Common Room.
  - Break room: Room 2126.
  - Dinners at Nuffield: Tuesday & Thursday. Meet there 18:40.
  - Guided tour of Oxford: Wednesday, 13:45, starting at Manor Road.

Thank you!