

# The political economy of AI: Who controls the means of prediction?

Maximilian Kasy

Department of Economics, University of Oxford

March 31, 2023

# Introduction

- Concerns about the impact of AI:
  - Fairness, discrimination, and inequality.
  - Privacy, data property rights, and data governance.
  - Value alignment and the impending robot apocalypse.
  - Explainability and accountability.
  - Automation and wage inequality.
- Efforts to regulate AI.
- How can we think systematically about these questions?

*Kasy, M. (2023). The political economy of AI:  
Towards democratic control of the means of prediction.*

# Key arguments

1. AI systems maximize a single, measurable **objective**.
2. In society, different individuals have **different objectives**.  
AI systems generate winners and losers.
3. Society-level assessments of AI  
require trading off individual gains and losses.
4. AI requires democratic control  
of algorithms, data, and computational infrastructure,  
to align **algorithm objectives** and **social welfare**.

# AI is automated decisionmaking

- AI systems maximize measurable objectives:

Russell and Norvig (2016), chapter 2:

*For each possible percept sequence, a rational agent should select an action that is expected to maximize its performance measure, given the evidence provided by the percept sequence and whatever built-in knowledge the agent has.*

- Leading approach: Machine learning (ML):

1. Supervised learning.

$$l(g(X), Y)$$

2. Targeted treatment assignment.

$$h(X) \cdot Y$$

3. Multi-armed bandits.

$$\sum_{t=1}^T Y_t$$

4. Reinforcement learning.

$$Q(X_i, W_i) = E[Y_i + Q(X_{i+1}, W_{i+1}) | X_i, W_i]$$

# Social welfare

Common presumption for many theories of justice:

- Normative statements about society are based on statements about individual welfare
- Formally:
  - Individuals  $i = 1, \dots, n$ .
  - Individual  $i$ 's welfare  $v_i$ .
  - **Social welfare** is a function of individuals' welfare

$$F(v_1, \dots, v_n).$$

# Agents of change

- How do we ensure that the objectives maximized by AI align with maximizing social welfare  $F(v_1, \dots, v_n)$ ?
- Which agents have the interests, the values, and the capacity, to move technology and policy?
- Voluntary ethical behavior by corporate managers and engineers?
- Economics: Corporations are primarily profit maximizing. Profit maximization might not be aligned with social welfare maximization.
- Democratic control is necessary. Those affected by AI decisions need to have effective control over the objectives that are maximized.

# Using this framework to discuss the social impact of AI

## STANDARD PERSPECTIVE

## ALTERNATE PERSPECTIVE

### **Fairness, discrimination, and inequality**

Deviation from profit maximization      Impact on social welfare

### **Privacy, data property rights, and data governance**

Individual property rights      Data externalities and democratic governance

### **Value alignment and the impending robot apocalypse**

Man vs. machine      Corporate vs. social interests

### **Explainability and accountability**

Individual recourse      Public debate over objectives

### **Automation and wage inequality**

AI as production function shifter      AI as automated decisionmaking

Thank you!