

Towards democratic control of the means of prediction

Maximilian Kasy

Department of Economics, University of Oxford

May 2024

The political economy of AI

1. AI systems **maximize** measurable objectives.
2. AI ethics & safety:
Typically discussed as **optimization errors**.
3. Key issues instead reflect
conflicts over the **choice of objectives**.
4. AI objectives are determined by those
who control the **means of prediction**.
5. **Democratic control** is necessary
to align AI objectives with social welfare.

1. AI maximizes measurable rewards

Textbook definition of AI (Russell and Norvig, 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.

- AI is more than large language models!
- Many contested applications:
 1. Newsfeed selection.
 2. Online pricing.
 3. Hiring.
 4. School admissions.
 5. Selection of households to bomb.

2. AI problems as optimization errors?

E.g. Algorithmic fairness:

Bias \approx deviation from profit maximization.

“The algorithm hired a man instead of a woman,
even though the woman would have been more productive.”

E.g. Value alignment:

Alignment problem \approx failure to maximize the correct objective.

“The AI maximized paper clip production,
even though humans don’t need ∞ paperclips.”

3. AI problems as conflicts over objectives!

- Distributional conflicts (cake for whom).
- Value conflicts (what cake).

E.g. Algorithmic fairness:

Who wins, who loses?

“What is the impact of hiring algorithms on gender inequality?”

E.g. Value alignment:

Not human vs. machine, but different humans with different objectives.

“The social media feed maximized ad-click profits,
but it harmed democracy & teenage mental health.”

4. The means of prediction

- Who gets to pick the objectives of AI systems?
- Those who control the means of prediction:
 1. Data
 2. Compute
 3. Expertise
 4. Energy
- Ethics of corporate engineers are unlikely to shift AI objectives, given the constraints of profit maximization.
- Instead we need:
Credible pressure by a wide variety of agents of change.

5. Democratic control

Q: How can we align the objectives of AI with the objectives of those impacted / society at large?

A: Let those impacted / society at large determine the objectives.

Q: How do we get there?

A1: Mobilize a variety of agents with the interests, values, and capacity to align profits with social goals.

A2: Create legal frameworks which give those affected by AI decisions effective power over AI objectives.

Further reading

- **How AI works (and for whom)**

Book forthcoming with University of Chicago Press.

- **The political economy of AI:**

Towards democratic control of the means of prediction.

Workingpaper available on my website.

Thank you!