1 Some questions

- Is AI going to make our lives worse?
- Will AI help us understand people better, or will it be used to oppress people better?
- What circumstances characterize AI systems which *service* people versus those that *trade* in people?

2 Approaches

either or...

- We present a framing for making sense of AI systems — especially in the context of the above question — through this lens that answers all these questions in useful ways. We also provide people with a useful foothold to begin "climbing" this otherwise unapproachable wall.
- We suggest a way of thinking about AI systems that may provide a useful perspective, specifically by aligning elements of this topic in such a way that certain answers become, or at least intuitive.

3 Problem areas (and solution approaches?)

- advertising/surveillance state
  - what happens? why do i care?
    1. people exist (online)
    2. online systems make inferences about the people
    3. these systems put ads in front of people
- justice/racism
  - bureaucratic orgs do Y
- streaming/garbage
  - bureaucratic orgs do Z
- ???

4 hot question

- Are AIs necessarily a dehumanizing force?
- Is the future of "*managed* work" one of oppressive burdens?
- Are AIs going to be tools of humans, or will humans be tools of AIs? (extensions?)
- in what ways are AIs making decisions that affect us?
  - legal: bail, sentencing, etc?
  - financial: loans, insurance
  - business: freelance work assignment, guiding workers to places
  - politics: information diet

5 why does it matter?

- The answer to (one of those questions) is going to determine if Neo from the Matrix is going to have to save us.

6 prior work: what's missing?

- Lee et al. [2] looked at this, but didn't take a shot at the deeper issues
- Irani and Silberman [1] got into some of the deeper undercurrents, but ...
- need to come up with other examples — look at notes from CHI review; look at FATML papers
- Amazon warehouse workers
- stitchfix? @mav
- Wikipedia?
- Facebook/Instagram/YouTubers working around the opaque preferences of the system?

7 lay out the metaphor

1. Bureaucracies are comprehensive organizations which try to achieve certain **goals**
2. They achieve those **goals** by doing **tasks**
3. The processes that determine how to do a task are what form the structure of bureaucracies
4. Following the process is deeply important, but people within the bureaucracies — bureaucrats — are empowered to use *discretion* in their own ways.
5. This discretion is exercised at different levels and in different contexts to varying degrees, but always in service to the **goals** of the bureaucracy (e.g. doctors providing medical care, so misclassifying patients as inpatient)

AIs are basically like that

1. Bureaucracies are the AIs
2. They're embedded with certain goals
3. They develop their own tasks to accomplish those goals (with incentives and disincentives to approach those goals)
4. Adherence to the process is important
5. It's difficult to conceptualize AIs which can exercise discretion

8 what does the metaphor tell us?

- AIs lack the capacity to exercise discretion in the current design paradigm
- But this is necessary to deal with humans, because of the "... relatively high degree of uncertainty because of the complexity of subject matter (people) and the frequency or rapidity with which decisions have to be made."

9 how should we rethink algos?

- We need to figure out how to design algorithms with mechanisms for people to intercede and break the rules on behalf of the higher order purpose
- AIs which have some conceptual understanding of that higher order purpose — the **goals** of the org — would be important
- Encoding the nuance of those **goals** will prove to be impossible
- Here are some ideas!
  - Question the *descriptive* vs *prescriptive* tendencies of AI/ML (e.g. *reflecting* societal prejudice vs implicitly *encoding* it into the system)
  - Provide operators with well–documented vectors to mess with levers
  - Do the things from that ethics paper about making clear how a decision was made

10 Refs

[1] Lilly C. Irani and M. Six Silberman. "Turkopticon: Interrupting Worker Invisibility in Amazon Mechanical Turk". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '13. ACM, 2013, pp. 611–620. isbn: 978–1-4503–1899–0. doi: 10.1145/2470654.2470742. url: http://doi.acm.org/10.1145/2470654.2470742.

[2] Min Kyung Lee et al. "Working with Machines: The Impact of Algorithmic and Data–Driven Management on Human Workers". In: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. CHI '15. ACM, 2015, pp. 1603–1612. isbn: 978–1-4503–3145–6. doi: 10.1145/2702123.2702548. url: http://doi.acm.org/10.1145/2702123.2702548.