PASQUALE CHAPTERS 5-6

Yanlin "Elaine" Chen

April 8, 2021

Ch5 Machines Judging Humans

- Evaluative AI that aim at measuring people's reliability, trustworthiness, criminality, fit for a job, etc.
- 4 reactions: 1) use the system to one's advantage, 2) improve accuracy, **3) make them fairer via technical methods/ regulation, 4) ban machine judgments altogether**
- Focused on 3 and 4: mend or end evaluative AI
- Tech: Robotic judgments, facial recognition, facial analysis, financial intrusion

Robotic Judgments

- AI judge: Algorithmic sentencing of violations of the laws
- Ubiquitous surveillance, outcome-based policing and replacing guard labor
- Advantage: cheap and efficient; (potentially) more equality in trials
- Danger: effective robotic policing allows gov't abuse
- Example: "House arrest" & home prison; compared to open prison in Scandinavia

Facial Recognition

- Two senses of **Misrecognition**:
  - 1. False matches in identifying people (e.g. criminals); failures to identify minorities, esp. women (“coded gaze”)
  - 2. FR destroys rights to obscurity/ anonymity in a crowd
- Including more **minority faces** in FR databases: good or not
- **Google Glass**: not letting people know they are under watch
- Status of Biometric Faceprints: **“Plutonium of AI”**
- Scientific freedom VS Citizens’ freedom of speech

Discussion: AI Judges and Facial Recognition

- 1. What do you think of the idea of home prison/ "house arrest"?
- 2. As facial recognition is used widely in law enforcement and other areas of life, is it better to include more minority faces and improve the accuracy of identification?
- 3. Has facial recognition already been misused/ used too extensively?
- Other thoughts?

Facial Analysis

- Faces said to reveal sexual orientation, intelligence, criminal features, etc.
  - Accuracy challenged
  - Causal connection VS correlation
- Affective computing: double-edged sword in workplace
  - Scenarios: CEO giving a pre to people around the world; pick out disengaged employees
  - Open to misuse; need clear regulation
  - Ch3: also used in schools

Financial Inclusion: creepy, predatory, subordinating

- Fintech firms scored creditworthiness of borrowers based on political activity, sleeping patterns, etc. Collect info from social media
- Predatory: incentivize machines to keep seeking out desperate borrowers for high interest-rate loans—> borrowers suffer
    - Preventing exploitative educational debt: Certify which training programs have provided a good “return on investment” to students
- Creepy: cellphone tracking, archiving, and data resale
- Subordinating: force people to maintain the same patterns of life that resulted in their desperation in the first place.
    - e.g. penalizing people for getting politically involved and not giving loans

Discussion

- 1. Pasquale claims that identifying criminals according to face features is wrong because it is based on **correlation** rather than **causal connection**. Do you agree, and why?

- 2. Will the predatory and subordinating financial inclusion reinforce social inequality?

- Other thoughts?

Social Credit System & Problem of Judgmental AI

- Social credit system in China
  - Focused on trust: social, commercial, judicial, gov't trust systems
  - Most common: Travel restrictions and bar from high-end consumption for those who filed for bankruptcy; but open to misuse
    - History: moral behavior as a part of politics; the lack of civic org.; society deficient in trust
- Habermas: "systematic colonization of the lifeworld"
  - "The bureaucracies of both governmental and market actors are always in danger of over expanding, juridifying the life-world [family, civil-society institutions] by imposing rules for correct conduct that oversimplify, distort, or outright extant ideals."
  - Non-quantitative judgment is important
- *Thoughts and questions?*

Ch6 autonomous forces

- Is killer robot more humane?
  - More precise targeting, e.g. age group, gender, combatant
  - Take humans out of the loop of targeting decisions, code ethical constraints into robots
    - Cool, calculative attacks by robots: morally worse than emotional?
      - Infinite array of situations, paucity of data: difficult to code general ethical rules and datafy different soldier experiences
- The law of war
  - Rule of distinction (combatant/civilian) — scenario of mother running near a soldier, difficult for robot to distinguish
  - rule of proportionality (excessive injury to civilians)— requires subtle and flexible case-by-case reasoning
  - paradox: makes the war more humane but harder to end

Upping the ante in great-power rivalries

- War robots can make a “million mistakes a second”
  - Malfunctioning or hacked software easier to spark war
- Preemptive, react faster, inclined to revenge
- The logic of nuclear deterrence may apply to autonomous systems, with them widely available around the world
- Barriers to ban
  - Comparison to land mine: could maim and kill non-combatants long after the end of the war
  - US solution: assure that future mines could be turned off; prefer regulation to ban
- Responsibility for autonomous war robots
  - Need to impose responsibility on programmers who cause mistakes
  - Proposal: Autonomous uses only for non-lethal weapons; transition war to peacekeeping

Resistance to military AI and paths to cooperation

- US, China, Saudi Arabia, India, France, Russia= biggest military spending, continue to increase. Investment unlikely to benefit the countries
- Political leadership matters— could choose to focus on economic development rather than military expenses e.g. Taiwan
- Google employees refuse to be involved in military AI tech; Snowden Revelations
- Military and policing AI are not used only for foreign enemies, but also for domestic citizens. e.g. anti-terror tools turned towards criminals and protestors in the US after Sep 11; China's detention of Uyghurs
- Citizens in different countries should all keep in mind AI-driven force can be a tool of oppression
- Development, governance, and humanitarian aid are important to security & stopping the pursuit of zero-sum wars. e.g. investment in infrastructure, public health, climate change

Questions

- 1.How does Pasquale's view compare to Vallor's or Wallach's views on autonomous weapons?
- 2. Should big tech companies such as Google be criticized as unpatriotic if they refuse to work on military tech (for their nation)?
- 3. Are there any other ways to prevent AI military competition? Or how do you imagine its evolution?
- Other thoughts?