Humans and Machines

Adam Waytz
Northwestern University


Anchor @anchor
Elon Musk says artificial intelligence is humanity's greatest threat
36 minutes ago

News

These are the jobs most at risk of automation according to Oxford University: Is yours one of them?

share

The Economist

The Opinion Pa

Bring on th

Probability tha
losses within the next two decades, 2013
(1=certain)

Job Probability

Home

BloombergBusinessweek

June 22, 2017

Is Your Job About To Disappear?: QuickTake

Robot YuMi, who will direct an orchestra as part of the Robotics Festival of Pisa at the Teatro Verdi of Pisa.

ARTIFICIAL INTELLIGENCE

How College Stude for Our Automate

CNBC

Search Quotes, News & Video

HOME U.S. NEWS MARKETS INVESTING TECH MAKE IT VIDEO SHOWS MORE

WORLD ECONOMY

ECONOMY | WORLD ECONOMY | US ECONOMY | THE FED | CENTRAL BANKS

WEF warns of 'vast numbers of jobs' being disrupted by automation and robots

Predictions About Automation

- Frey & Osborne (2013): 47% of jobs at risk of being automated over next 20 years
- Grace et al. (2017): 50% chance of AI outperforming humans in all tasks in 45 years and of automating all human jobs in 120 years

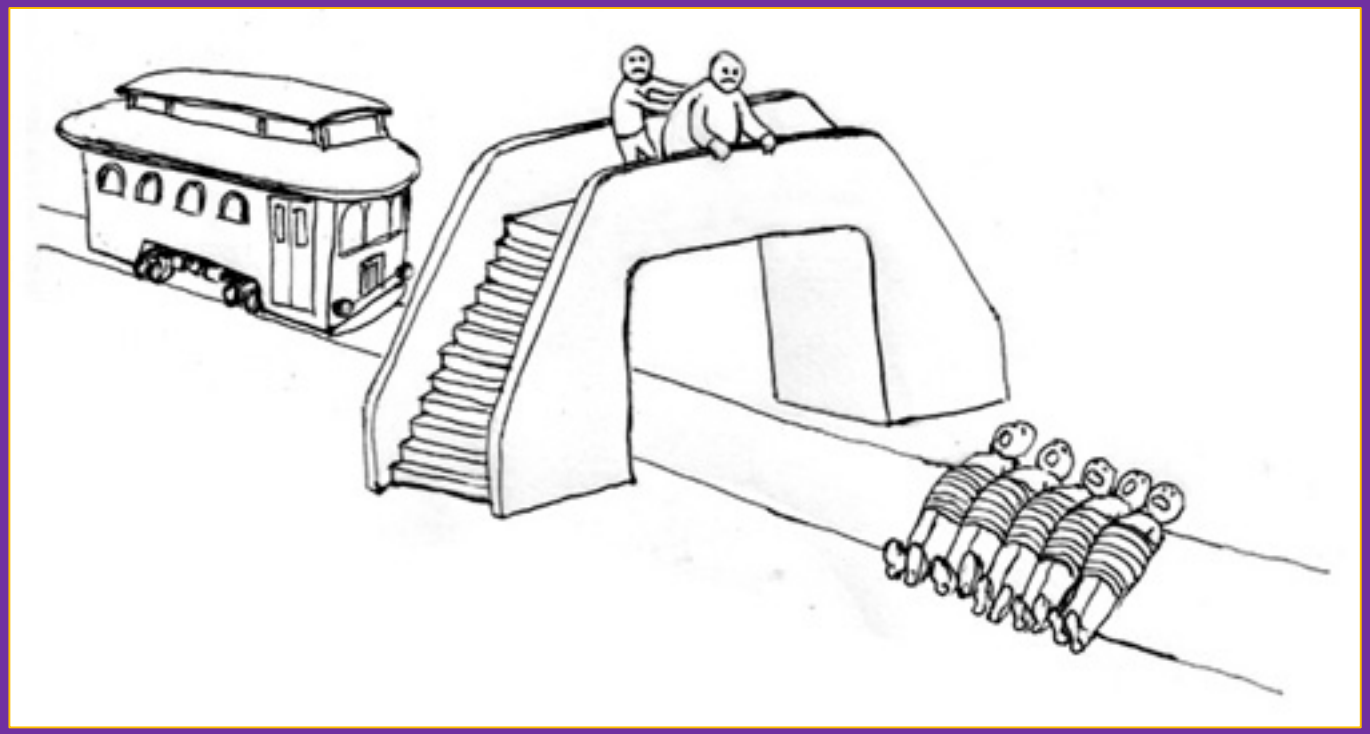
- Acemoglu & Restrepo (2017): For every one robot added, 5.6 workers lose jobs. Predicted by 2025, 0.94-1.76 % lower employment-to-population ratio
- Deming (2012): Jobs that involve social skills less at risk
- Autor (2010); Cortes et al. (2016): Routine jobs most at risk

- Psychological implications ????

Two Fundamental Issues

- Understanding reasons why humans distrust machines
- How to optimize human-machine partnerships to maximize trust

A dilemma



Two Key Findings:

-About ¼ of people are willing to push

-People do not trust those willing to push

(Everett, Crockett, & Pizarro, 2016)

Distrust in Technology: Causes

- Too cost-benefit oriented (Everett, Pizarro, Crockett, 2016)
  - Distrust in agents that do not follow moral rules
- Algorithmic process is too opaque (Yeomans, Shah, & Mullainathan, Kleinberg, 2017)
  - People prefer a human recommender to an algorithmic recommender (for jokes), even though the algorithm does a better job selecting funny jokes
- Belief that algorithms are less capable of learning (Dietvorst, Simmons, & Massey, 2015)
  - People bet on human versus algorithmic forecasters (for predictions of personnel success, airline statistics) even when they see algorithms outperform humans
- Stereotype of robots as incapable of handling social/emotional tasks (Waytz & Norton, 2014)
  - People express discomfort with robots taking on human work perceived as requiring emotional/social skill

Trust in Technology: Remedies

- Giving people simple explanations for how machine learning works increases trust in algorithm to perform (Yeomans et al., 2017)
  - Explaining collaborative filtering to people increases trust in algorithm in joke task
- People overcome algorithm aversion when allowed to modify them (Dietvorst, Simmons, & Massey, 2016)
  - Giving people opportunities to slightly modify forecaster algorithms increases trust in them
- People trust algorithms for more objective decisions and when expertise is lacking (Logg, 2017)
  - National security experts are least likely to trust forecasting algorithms for political events (cyberattacks, Brexit). Suggests potential need to reduce overconfidence
- People trust robots for social/emotional tasks when robot is designed to appear emotional (Waytz & Norton, 2014; Waytz, Heafner & Epley, 2014)
  - Giving an autonomous car a name/gender/voice attenuates blame for and increases trust in the context of an accident
  - Designing a robot with “emotional” facial features increase trust in that robot for a social task (e.g., therapy, social work)

Optimizing human-machine partnerships

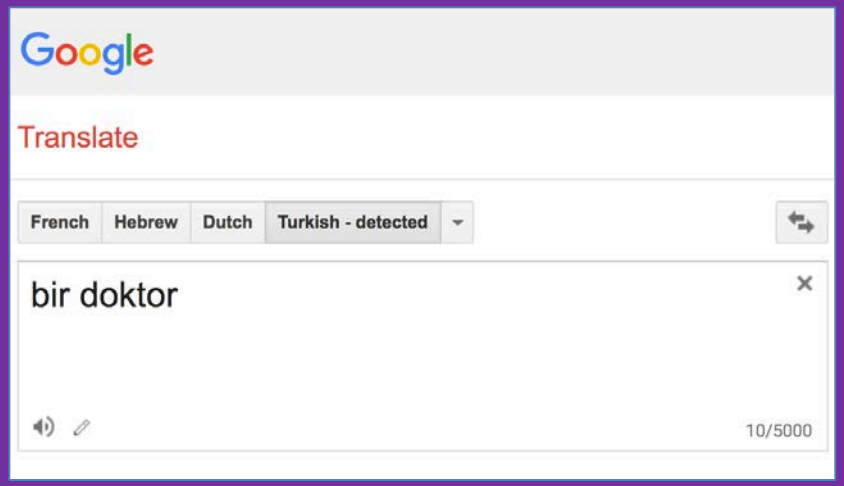
(1) Let robots do things that humans cannot, let humans implement “moral” rules

(2) Let robots handle the dull, rote, mechanical (i.e., robotic) work

(3) Let robots attempt to reduce the emotional burdens that humans face in their jobs

Let Computers Compute / Let Humans Implement Moral Rules

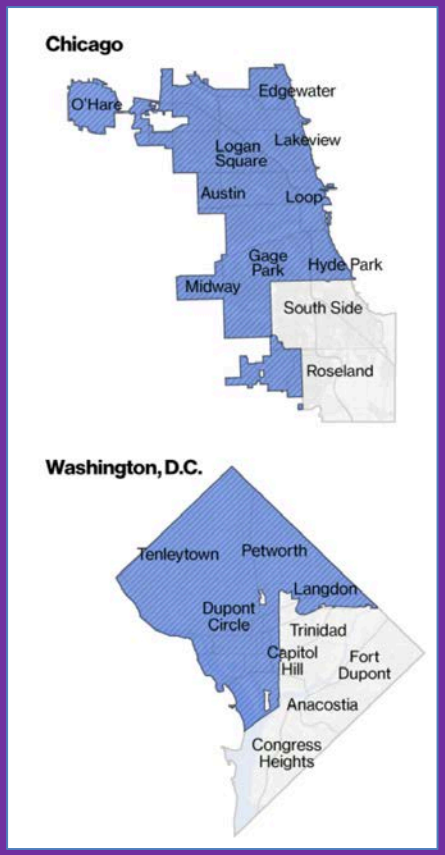
Caliskan, Bryson, Narayanan, 2017



Translates to “he is a doctor.”

Adjusted to “a doctor” (Fixed by Google)

Amazon Doesn’t Consider the Race of Its Customers. Should It?



Once Bloomberg identified this, Amazon corrected it

Let Robots Do the Robotic Work

- A MIT team developed a platform to detect 85% of cyberattacks and reduce false positives by 5%
  - Platform sifts through data, reports a sample of its findings to human analysts.
  - Analysts scan these results for false positives, provide that input to the platform.
  - Platform uses human feedback when performing its next search for cyberattacks. Platform improves
- An otherwise unsupervised machine would produce closer to a 20-25% false positive rate

(Veeramachaneni, Arnaldo, Korrapati, Bassias, & Li, 2016)

Let Machines Handle Emotional Labor

- Example: Privacy Authentication in Customer Service
  - Emotionally burdensome (customers are frustrated)
- Canadian financial services firm uses a biometrics to identify customers by voice, eliminating authentication questions and improving customer service routing by 50%
- European Bank uses biometrics to identify high-profile clients as their conversation progresses. System has reduced call handling time by 15s, with 93% of clients rating the system 9/10
- Australian organization receives 9m calls per year, 75% require authentication. Voice biometrics eliminates authentication questions, cuts avg call by 40s.

(Alter & Shukla, 2016)

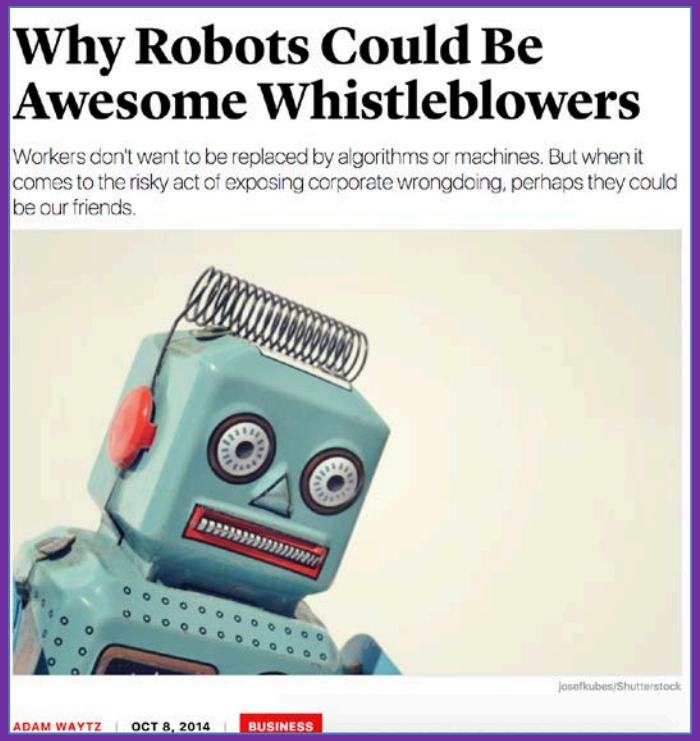
Optimizing human-machine partnerships builds trust in technology

Final thoughts: Implications Beyond Trust

- How does Automation affect attitudes toward humans. We show automation concerns → anti-immigrant attitudes (Gamez-Djokic & Waytz, 2018)

- Understanding when technology use hampers or helps empathy (Waytz & Gray, 2018)

-- Optimizing division of labor by identifying tasks that robots would perform better than humans (Waytz, 2018)



THANK YOU

a-waytz@kellogg.northwestern.edu