## The corrupting influence of AI as a boss or Counterparty

Hal Ashton, University of Cambridge Matija Franklin, UCL

The corrupting influence of AI as a boss or Counterparty

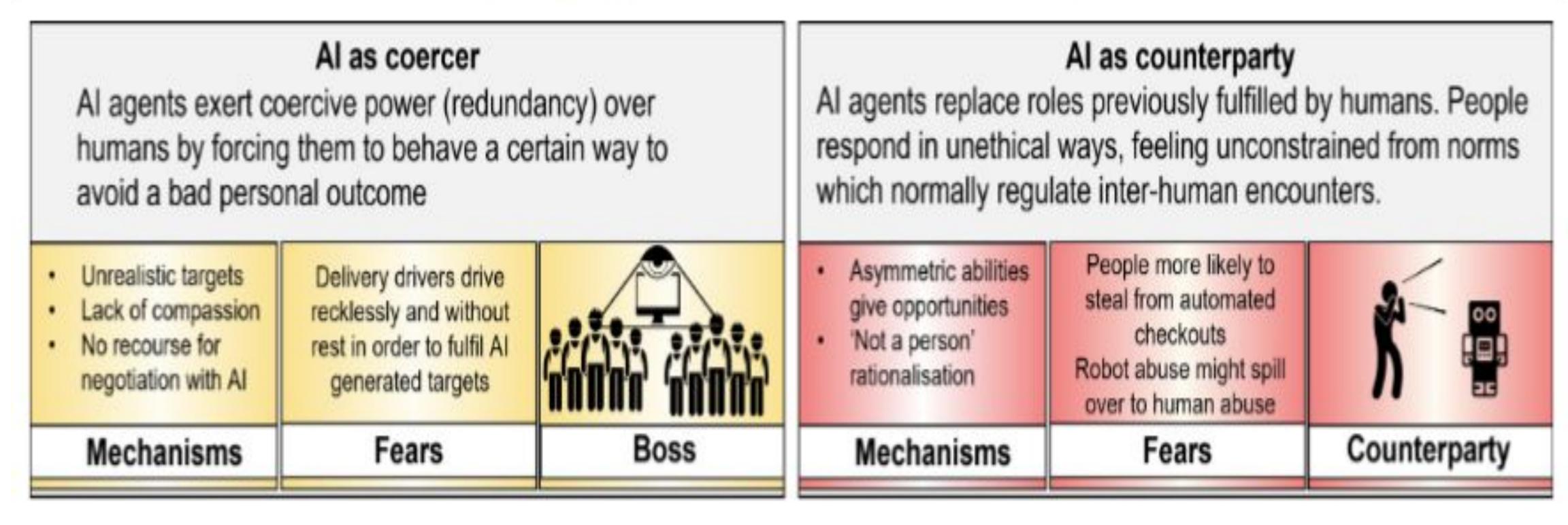


Figure 1. Two additional roles in which AI agents and humans influence ethical behavior. The main mechanisms and some fears (aka examples) are displayed in each role.

## Al Boss

- Workers in the 21st century are increasingly being told what to do by AI overseers
- Workers in the gig economy are subject to algorithmic domination facilitated in part by the asymmetric access to information that the AI boss has.
- Task scheduling algorithms operate in an idealised world, where there are no difficulties in finding items or locations and humans do not tire and require breaks.

## **Al Counterpart**

- Humans will respond to the asymmetry between their abilities and that of the Al with adversarial behaviour
- Because Al counterparties are not human, people will feel freer to abuse them.
- Virtue ethics would suggest that such behaviour is potentially damaging to one's moral character
- Such a perspective argues that virtue is performed, creating habits.
- Some argue for the design of robots which promotes better virtue from peo- ple, such as incorporating an ability for consent.