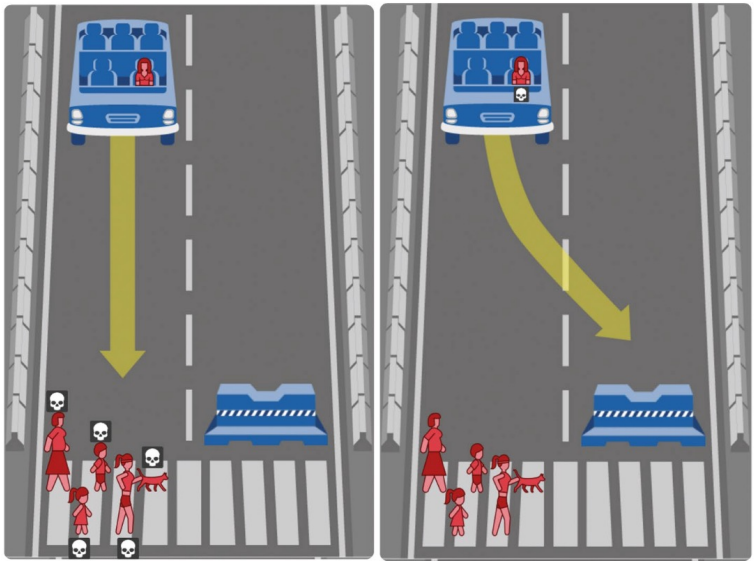
Today:

- Fairness, Accountability & Transparency (FAT*) of AI and Algorithmic Systems in general



Autonomous vehicles:
What are the chances for such an ethical dilemma to occur?

Today:

Across multiple other task domains, *every single decision* made by an algorithm involves an ethical dimensions, e.g.:

- Predictive policing and jurisdiction (recidivism decisions)
- Predicting financial worthiness (credit scoring)
- Predicting employees' success (hiring decisions)

Line of argumentation:

In all of these scenarios, one must necessarily ask:

- Is the decision *fair?*
- *Who* made the decision?
  - Who is responsible?
  - Where rests accountability?
- *How* was the decision made?
  - Can we understand the decision-process?
  - How transparent is the process?

Why bother?

Line of argumentation: Why bother?

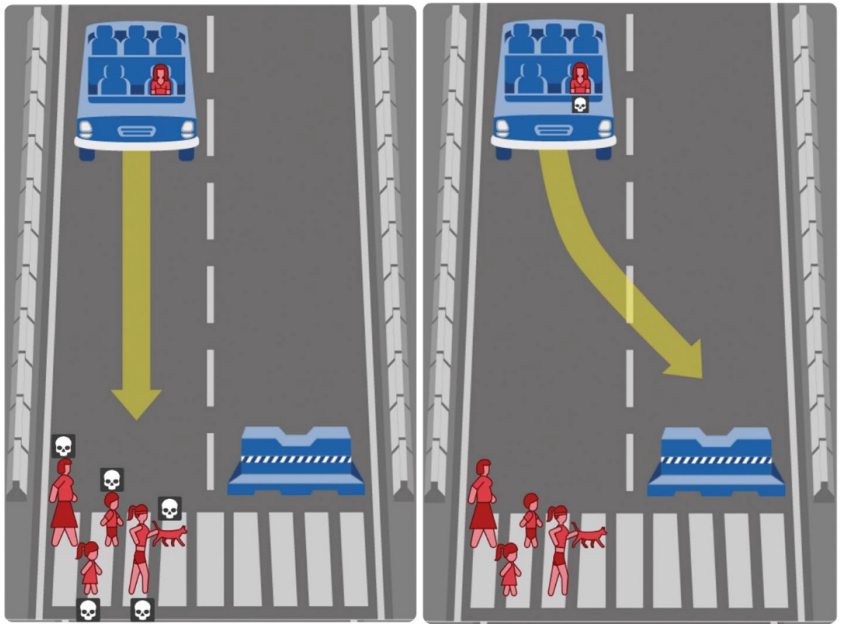
Fairness, accountability and transparency *can* serve as ethical measurements.

- Products are subjected to legal requirements (based on ethical considerations)...
- Fairness, accountability and transparency are *trust-enhancing* factors.

Ethics → trust-enhancing factors (FAT) → product adoption

78% of Americans Do Not Trust AVs

Ethics → trust-enhancing factors (FAT*) → product adoption



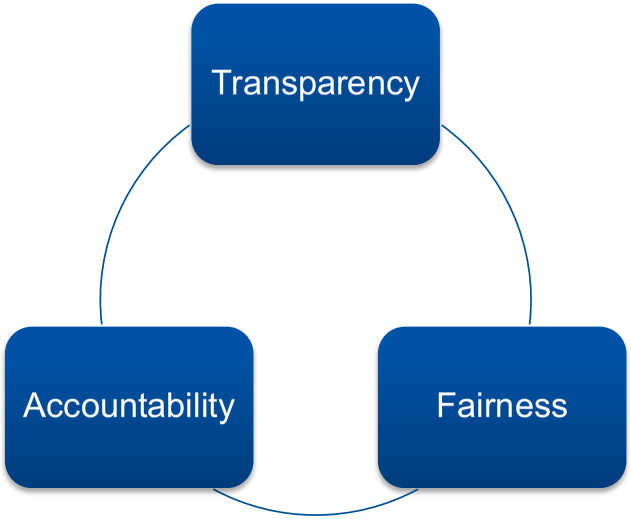
Americans Feel Unsafe Sharing the Road with Fully Self-Driving Vehicles.
American Automobile Association, see in *Nature*, 2017

We are frequent subjects of algorithms…

- Predicting employees' success (Highhouse 2008)
- Predicting academic performance (Dawes, 1971)
- Predictive policing and jurisdiction (Wormith et al., 1984)
- Predicting driving outcomes (Koo et al., 2015)
- Predicting sport judgments…

How fair, accountable and transparent is AI-based algorithmic decision making?

FAT: Trust-enhancing factors for AI adoption



- Without transparency, can we know whether the decision was fair or who is responsible for it?
- Is transparency a necessary (and sufficient?) condition to determine accountability and fairness in an algorithmic system?

Case-study: Assessment tools to predict recidivism risk

- How likely is a defendant to commit a felony or misdemeanor once released from prison?

What's the appeal to use risk assessment tools?

- The United States **locks up far more people than any other country**, a disproportionate number of them black.
- Key decisions in the legal process have been in the **hands of human beings guided by their instincts and personal biases**.
- If computers could accurately predict which defendants were likely to commit new crimes, the **criminal justice system could be fairer**.

Eric Loomis

The New York Times

Sent to Prison by a Software Program's Secret Algorithms

NY Times, May, 2017

Classified by *COMPAS* software tool as
"*individual who is at high risk to the community*"

→Judge sentences Eric Loomis to 6 years in prison.

Is “COMPAS” fair?

- How does the algorithm calculate the score?
  - Developed by company Northpoint (now “equivant”).
  - *COMPAS* in use since the year of 2000 (predictions for > 1 million offenders).
  - Scores from 1 – 10 (10 = highest risk score).
  - Algorithm is proprietary and thus a trade secret:
    - Little transparency over decision-making process.