Legitimacy

Exercise 1:

Here's a thought experiment. Imagine a state that wants to make the roads safer by suspending the licenses of drivers determined to pose an unacceptably high accident risk. We'll propose a series of ways in which the state could go about this. The first is widely adopted today and generally accepted by the public (although it is not without its concerns); the successive scenarios gradually introduce more controversial elements.

1. Accurate & Aligned, Effective, Reliable Well executed, Well justified

1. Drivers get points on their license for violating the law. The number of points is tailored to the severity of the offense. The Traffic Safety Bureau suspends a driver's license if they accumulate a certain number of points. The specific number of points assigned to different offenses and the threshold for suspension are worked out by hand by members of the Traffic Safety Bureau; these are both communicated to drivers.
2. Statisticians at the Traffic Safety Bureau analyze data on accidents and conclude that they can derive a more accurate predictor of accidents using a more sophisticated formula. The formula is posted on the Traffic Safety Bureau's website, but it is no longer something that drivers can remember or mentally compute. 2. reliable, accurate & aligned
3. Since it is laborious and error-prone to compute the formula with pen and paper, the Bureau computerizes its calculation. The Bureau also automates the suspension of the licenses when the relevant threshold is met. 3. Accurate, effective, not well justified -
4. Next, the agency turns to additional data sources to improve the classifier. It partners with the police department to train officers to question occupants and collect as much data as possible during stops. It ingests data from insurance companies who put tracking devices on vehicles to compute accident risk. 4. not well aligned, reliable, effective, not well executed not well justified

4. Using a variety of data sources to take the decision

What legitimacy concerns do you have with each of these scenarios? What is it specifically about the scenario that triggers the concern?

Exercise 2:

When someone cuts a long line during the morning rush at our regular coffee shop, our objection to such behavior might rest on the belief that the store should operate on a principle of **first come, first served**—that no one should have special priority beyond their time of arrival. This particular principle of queuing is so deeply ingrained in certain societies that it can feel self-evident.

1. In the context of a coffee shop, what are some alternatives to first come, first served? How might each of these be justified? Why don't coffee shops tend to use these, despite those justifications?
2. What if we changed the scenario from queuing at the coffee shop to boarding an airplane? How would the answers to these questions change?
3. And now consider the different ways that we might prioritize the provision of vaccines. How would these answers change once again in this scenario?

Exercise 3:

How would you construct training data for each of the following applications under the automating judgment approach (emulating the informal decision-making processes of humans) and the predictive optimization approach (Learning decision-making rules from labeled data)? In each case, what are the pros and cons of each approach from an ethical perspective?

1. Ranking candidates who have applied to a PhD program.
2. Lending: estimating the likelihood that a loan applicant will default.
3. Pretrial release: deciding whether a defendant poses too much of a risk.

(Credit: https://fairmlbook.org/)