

We use cookies to improve your experience on our website. By using our website you consent to all cookies in accordance with our updated [Cookie Notice](#).

I accept



Why we have the ethics of self-driving cars all wrong



Autonomous vehicles are never drunk, drowsy, or distracted.

Image: REUTERS/Noah Berger

21 Jan 2018

Karl Iagnemma

President, nuTonomy

This article is part of the **World Economic Forum Annual Meeting**

A trolley barrels down a track at high speed. Suddenly, the driver sees five people crossing the track up ahead. There's not enough time to brake. If the driver does nothing, all five will die. But there is enough time to switch onto a side track, killing one person instead. Should the driver pull the switch?

Philosophers have debated trolley problems like this for decades. It's a useful thought experiment for testing our intuitions about the moral difference between doing and allowing harm. The artificial scenario allows us to ignore empirical questions that might cloud the ethical issue, such as could the trolley stop in time? Could the collision be avoided in another way?

Recently the trolley problem has been invoked within the real-world policy debate about regulating autonomous vehicles (AVs). The issue at hand is how AVs will choose between harms to one set of people or another.

In September 2016, the National Highway Traffic Safety Administration (NHTSA) asked companies developing AVs to certify that they have taken [ethical considerations](#) into account in assessing the safety of their vehicles.

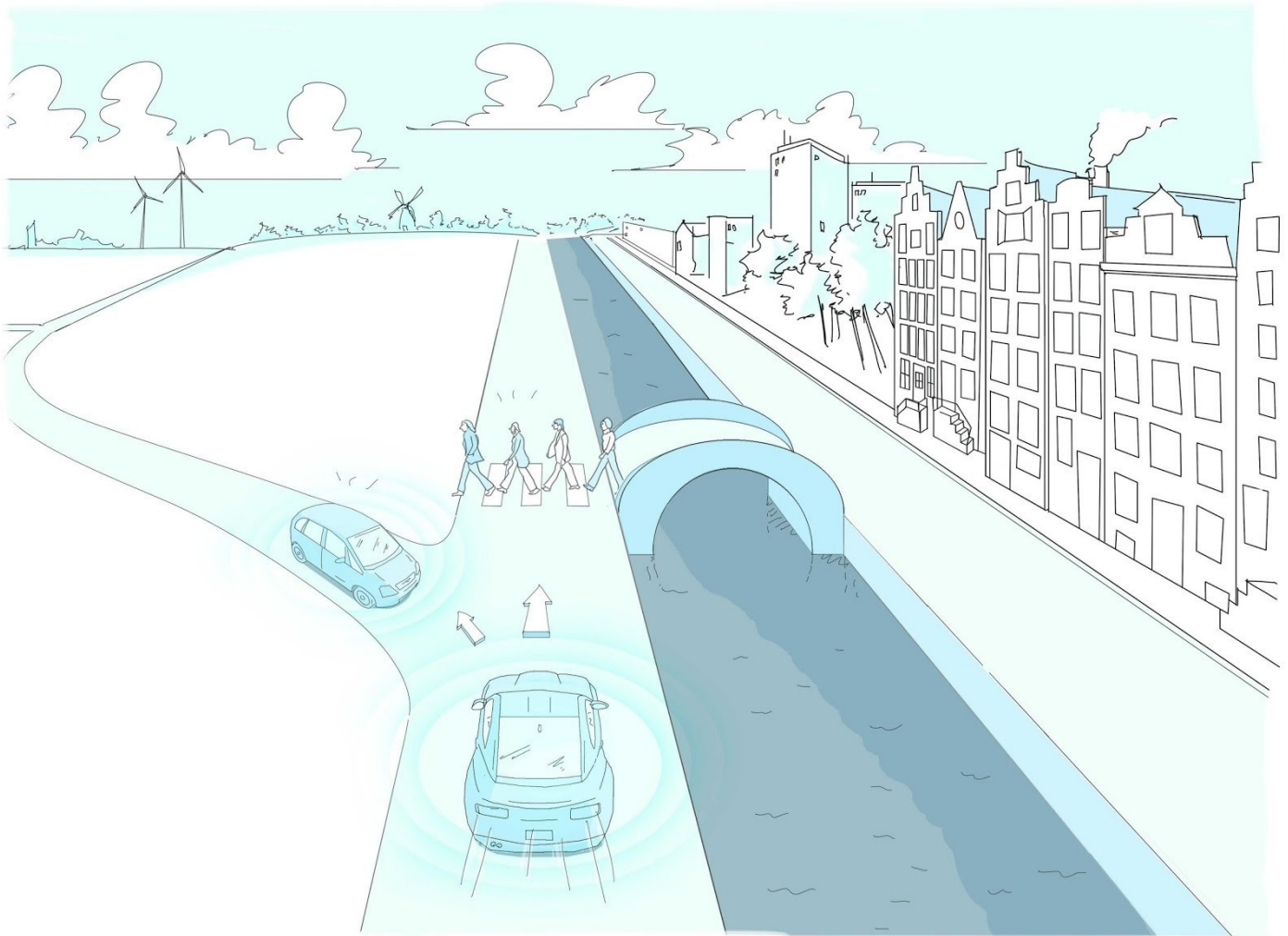


Image: Delft Technology University

Engineers and lawyers actually working on AV technology, however, largely agree that the trolley problem is at best a distraction and at worst a dangerously misleading model for public policy.

The trolley problem is the wrong guide for regulating AVs for three reasons:

1. Trolley problem scenarios are extremely rare

Even in a world of human-driven vehicles, for a driver to encounter a real-world trolley problem, he or she must 1) perceive an imminent collision in time to consider alternative paths; 2) have only one viable alternative path, which just happens to involve another fatal collision; and yet 3) be able to react in time to steer the car into the alternative collision. The combination of these three circumstances is vanishingly unlikely. It's not surprising, then, that we never see trolley problem-like collisions in the news, let alone in judicial decisions.

But sadly, unlike trolley problems, fatal collisions are not rare. The National Safety Council estimates that about 40,200 Americans died on the highway in 2016, a 6% increase over the previous year. By comparison, about 40,610 women in the US will die from breast cancer this year, as estimated by the American Cancer Society. A NHTSA study concluded that driver error is the critical reason for 94% of crashes. Policymakers need to keep the real causes of preventable highway deaths, like alcohol and texting, in mind to save lives.

2. Autonomous vehicles will make them even rarer

To the extent that trolley problem scenarios exist in the real world, AVs will make them rarer, not more frequent. One might think that, since AVs will have superior perception and decision-making capacities and faster reaction times, an AV might be able to make trolley problem-like choices in situations where a human driver wouldn't. But those same advantages will also enable an AV to avoid a collision entirely – or reduce the speed and severity of impact – when a human driver wouldn't.

Unlike the track-bound trolley, an AV will almost never be restricted to two discrete paths, both of which involve a collision. AVs are equipped with sensors that provide a continuously updated, three-dimensional, 360-degree representation of the world around the vehicle, enabling it to know, and be able to act on, many alternative paths. More importantly, since AVs are never drunk, drowsy, or distracted, they are less likely to be in near-collision situations in the first place.

3. There is not much regulators can do about them

Even if trolley problems were a realistic concern for AVs, it is not clear what, if anything, regulators or companies developing AVs should do about them. The trolley problem is an intensely debated thought experiment precisely because there isn't a consensus on what should be done.

Generally, if commentators applying the trolley problem to AVs give any conclusions at all, they propose that AVs should not distinguish among different types of people, based on age, sex, or other characteristics. But it doesn't take a trolley problem to reach that common sense conclusion.

Focusing on the trolley problem could distract regulators from the important task of ensuring a safe transition to the deployment of AVs, or mislead the public into thinking either that AVs are programmed to target certain types of people or simply that AVs are dangerous.

We are all vulnerable to the tendency to overestimate the likelihood of vivid, cognitively available risks rather than statistically likelier, but less salient, risks. We often neglect the base rate of conventional traffic accidents, even though the statistical risk is high. Associating AVs with deadly trolley collisions could only exacerbate this problem.

Conflating thought experiments with reality could slow the deployment of AVs that are reliably safer than human drivers. Let's not go down that wrong track when it comes to regulating self-driving cars.

Written by

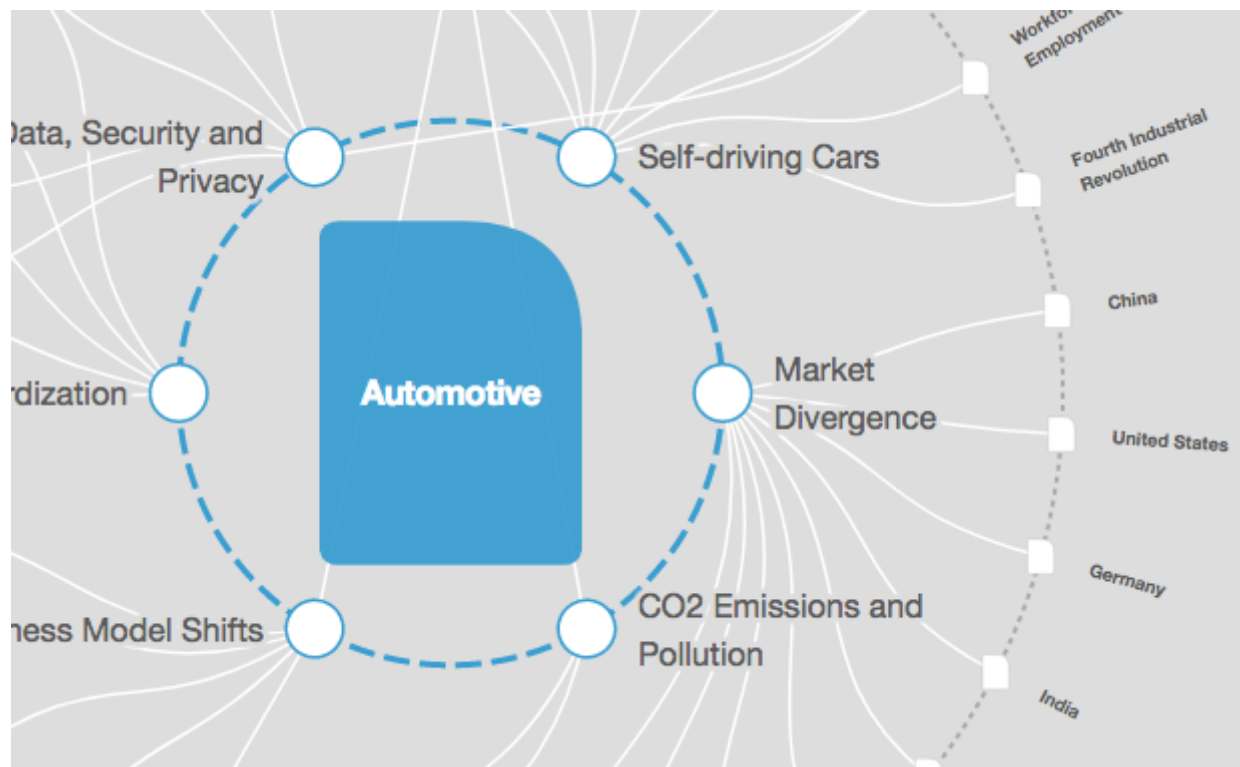
[Karl Iagnemma](#), President, nuTonomy

The views expressed in this article are those of the author alone and not the World Economic Forum.

Explore context



Automotive



Explore the latest strategic trends, research and analysis



Subscribe for updates

A weekly update of what's on the Global Agenda

© 2018 World Economic Forum Privacy Policy & Terms of Service