



Universiteit Utrecht

The Ethics of Self-Driving Cars

Sven Nyholm, Assistant Professor of Philosophical Ethics

Before 2015 . . .

Valentine's day, 2016:



May 2016:



Similarities and differences between the Google- and Tesla-cases:

Google:	Tesla:
Assumed (partial) responsibility for what happened	Emphasized that customers need to assume “control and responsibility”
Car operated by Google employee, as part of Google’s testing of their vehicle	Car operated by a “customer”
Google promised to update the software of their car, so as to make it better able to handle similar situations	Tesla promised to update the sensors of their car, so as to make it better able to handle similar situations

March 2018:



A need for “ethics settings”?

- Always hand over control to human drivers?
- Always simply brake?
- Tech in isolation vs. tech in society

Mandatory
ethics settings?



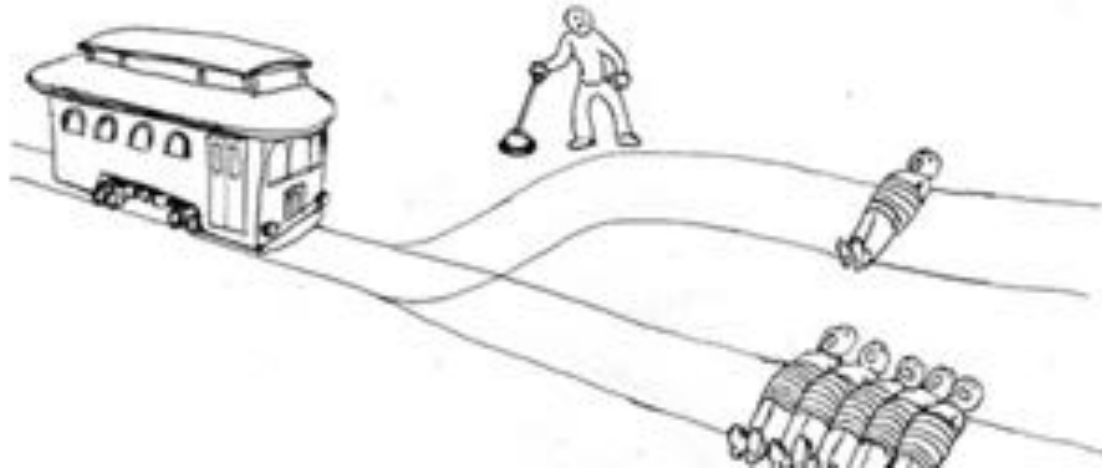
Good for
coordination

An “ethical
knob”?

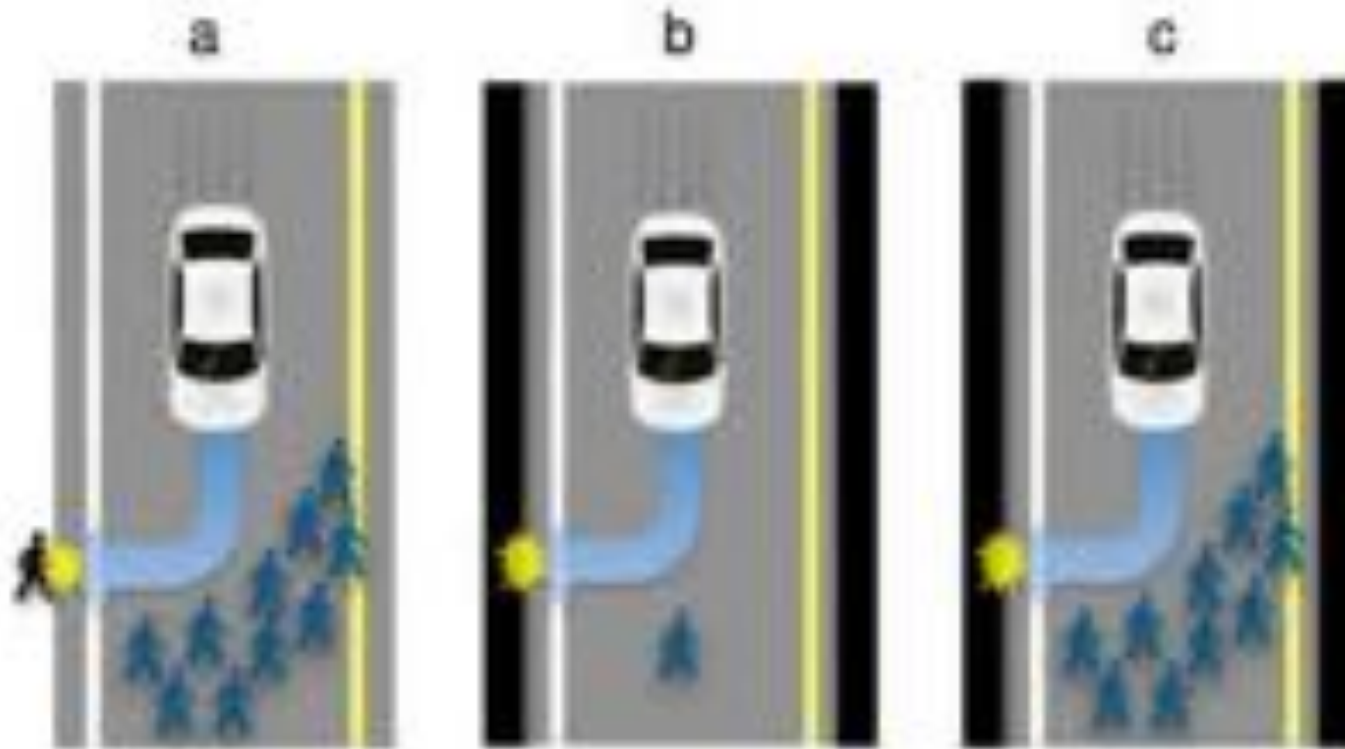


Clearer
responsibility-
allocation

The trolley problem . . .



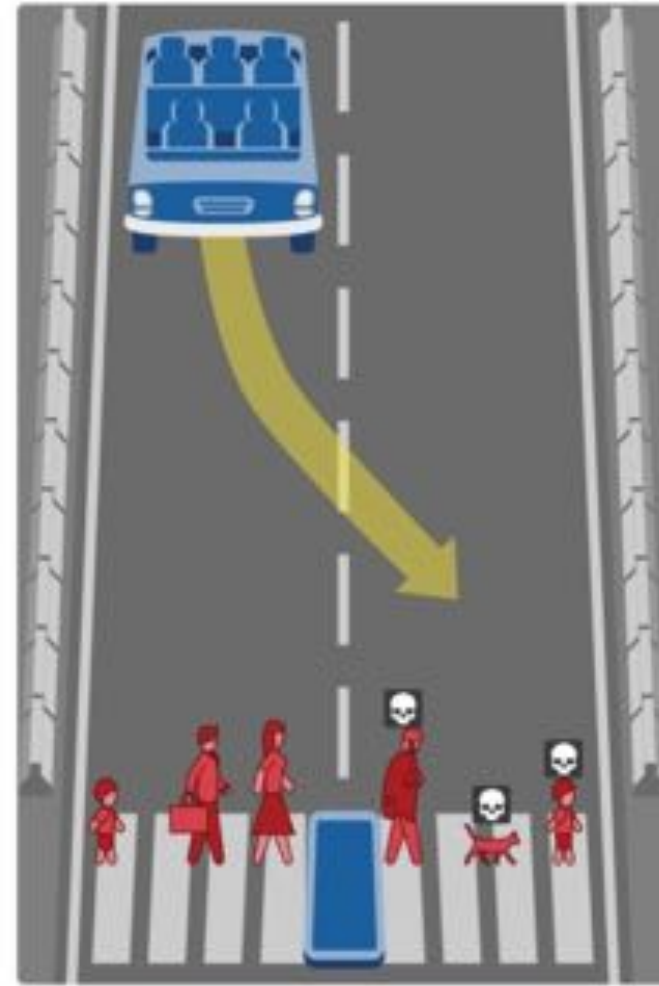
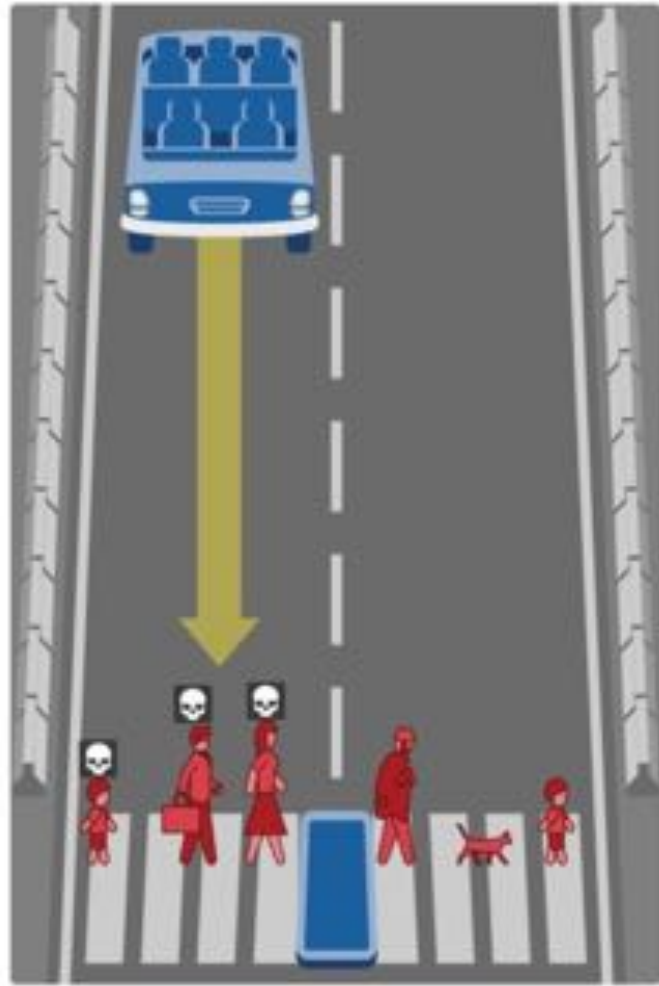
Crashing self-driving cars: an applied trolley problem?



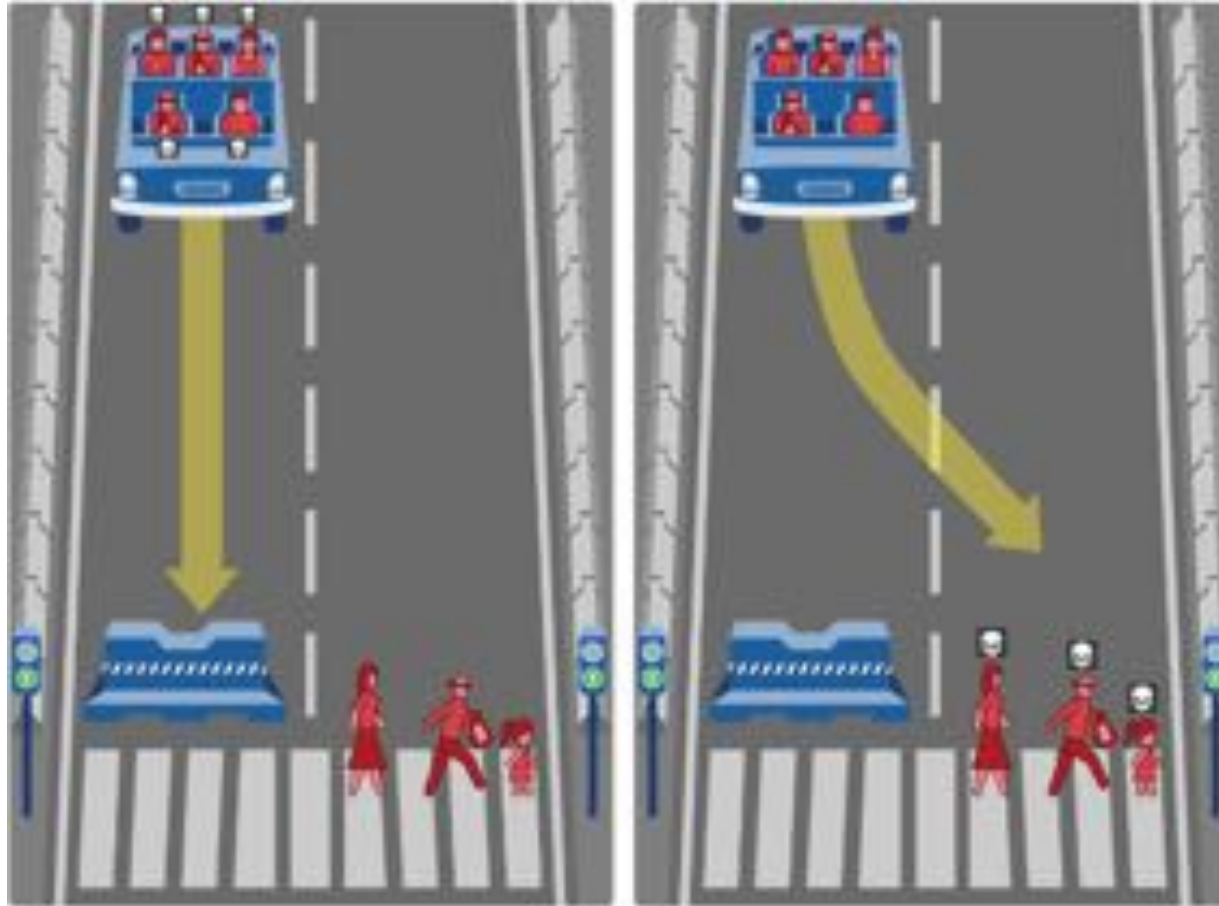
Three worries about the trolley problem comparison:

- 1: Very different decision making scenarios
- 2: Moral & legal responsibility
- **3: Risk & uncertainty**

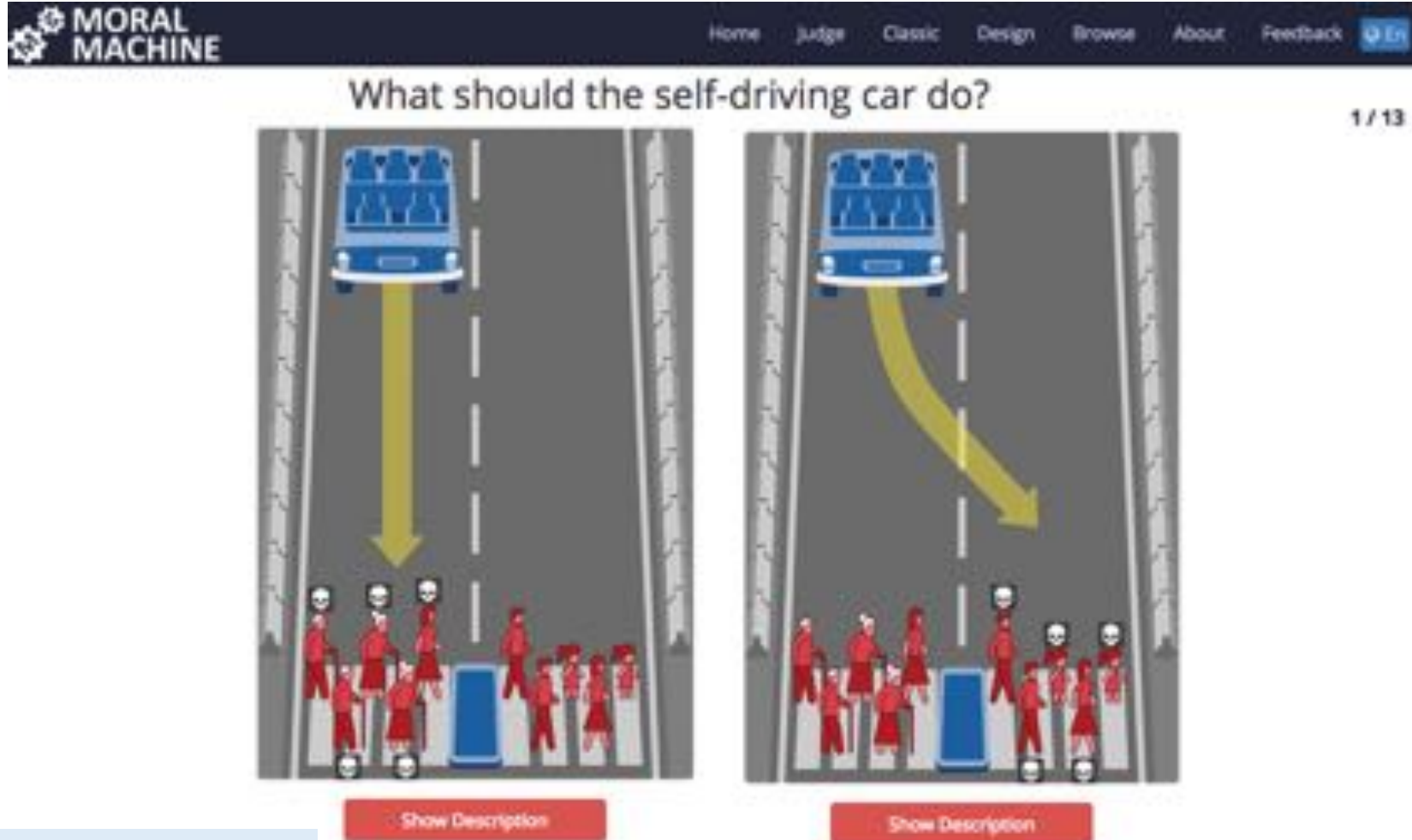
The "moral machine" experiment:



The "moral machine" experiment:



The "moral machine" experiment:



<http://moralmachine.mit.edu>

The Moral Machine: Two interesting findings:

- 1: self vs. others asymmetry
- 2: cultural variations

Three worries about such "empirical ethics":

- 1: changing attitudes with more real-world experience
- 2: gut reactions vs. moral arguments
- 3: conflicting attitudes/clash with "golden rule"



Interesting real-world
case: Christoph von
Hugo from Mercedes

What about the traditional ethical theories?

- Utilitarianism
- Kantian ethics
- Virtue ethics
- Social contract theories



What about the traditional ethical theories?

- Utilitarianism



Example: What would a utilitarian recommend?

- Kantian ethics

- Virtue ethics

- Social contract theories



Social contract theory support for harm-minimization?



After a crash. . .

- Who is morally and legally responsible?

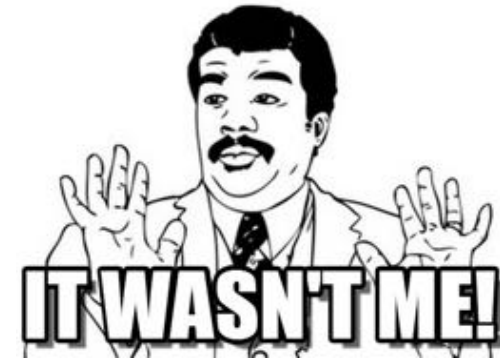
- *Car manufacturer?*

- *The driver?*

This is where
“existential”
questions arise. . .

- *Legislative body that permits automated vehicles on public roads?*

- *All users of automated vehicles collectively?*



Sartre: “existence precedes essence. . .”



The “existential” question of automated driving:

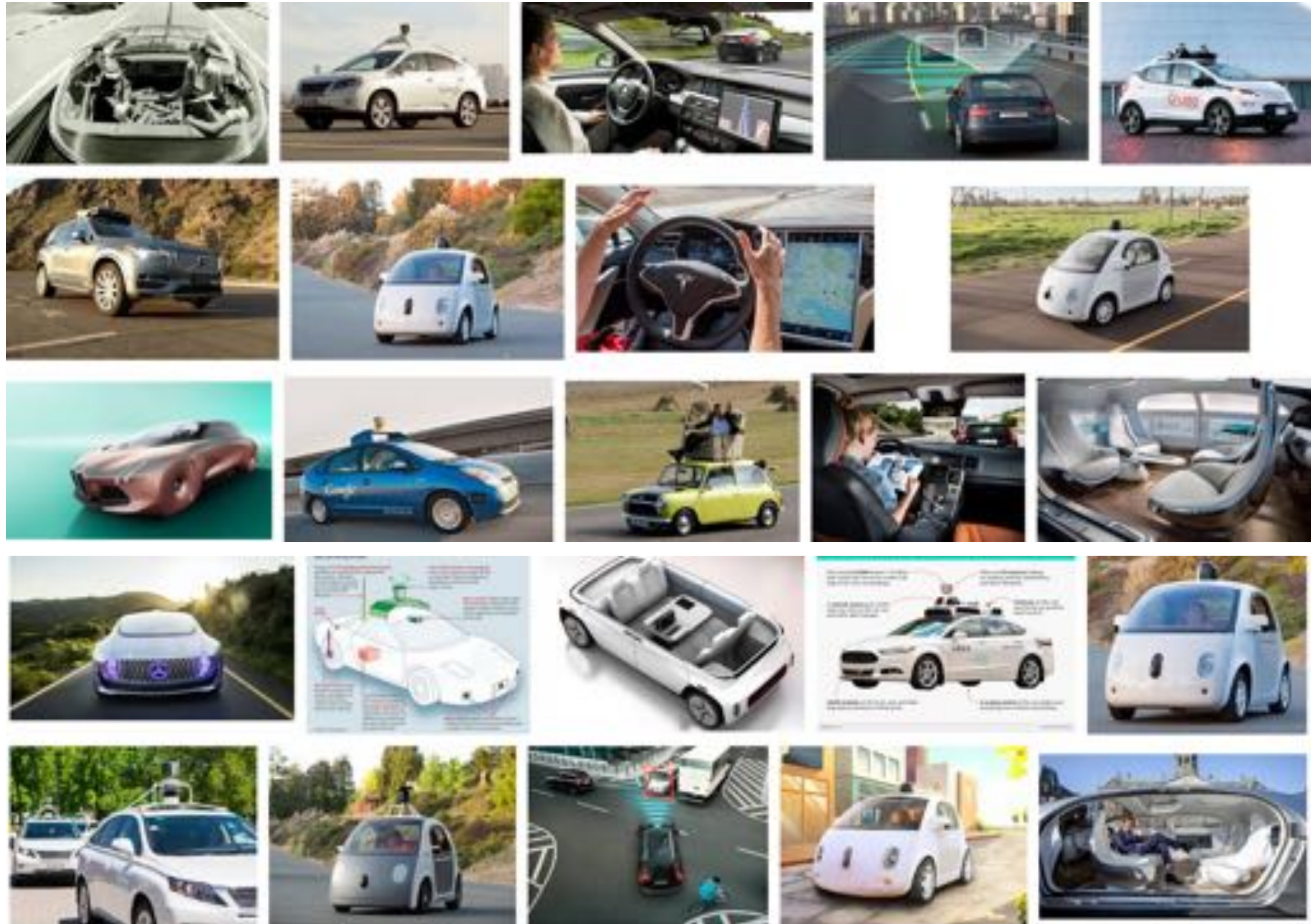
- Automated vehicles are emerging before we have a moral and legal framework for them. . .

“Driver”?
Manager?
Supervisor?
Owner?
User?
Passenger?

- **What is the new role of the driver?**

- **What is the relation between an automated vehicle and the person operating it?**

Employer – employee?
Master – slave?
Principal – agent?
Owner – property?
Pilot – co-pilot?
Or what???





Universiteit Utrecht

The Ethics of Self-Driving Cars

Sven Nyholm, Assistant Professor of Philosophical Ethics