

Ethical questions and issues identified in the field of autonomous driving

© Christoph Lütge

The following represent some of the issues and (ethical) questions that need to be addressed and answered in the future to be able to advance the development and implementation of autonomous vehicles (AVs) in a responsible manner.

Technical safety of AVs:

One of the prime rationale for introducing AVs onto streets is the aim to decrease fatalities that arise from human error. To achieve that AVs reduce physical harm to individuals in traffic, their technical robustness and safety needs to be ensured. In this regard, relevant questions to be answered are:

- How can potential threats (especially cybersecurity threats) be prevented?
- What are ‘safe’ fallback plans for AVs?
- How can we experiment with AVs on the road without harming humans?

Balancing risks in AVs:

Generally, AVs should be allowed to conduct a responsible balancing of risk or estimated harm. However, this balancing should never be based on personal characteristics of individuals but rather take into consideration more objective features of traffic participants. In this regard, relevant questions to be answered are:

- What are such objective factors that AVs can rely on in their decision-making process regarding risk distributions?
- How can this be technically implemented (e.g. by adjusting the lateral position of AVs on a lane)?

Human agency in AVs:

AVs have enormous potential to influence human agency, either in a positive manner by e.g. offering mobility solutions to mobility-impaired individuals or in a negative manner by e.g. restricting self-determined and independent decisions of the driver. However, to ensure a high level of human agency in AVs, relevant questions to be answered are:

- To what extent and in which situations should humans be able to override the AV?
- Through what exact processes can we enhance human agency in AVs (e.g. training drivers and including external human-machine interfaces)?

Privacy & data governance in AVs:

AVs will need to collect a vast amount of data for proper and safe functioning. Nevertheless, personal privacy still should be respected by, for example, transparently communicating how data is governed or by explicitly requesting affirmative consent from the driver. In this regard, relevant questions to be answered are:

- What types of data inside and outside the AV need to and should be collected?
- Under which circumstances and in which format can valuable data be shared with third parties?

Responsibility, liability & accountability in AVs:

In case of an accident where an AV is involved, the vehicle itself cannot be held morally accountable for the outcomes. Responsible parties will instead be manufacturers, component suppliers, technology companies, infrastructure providers or car holders / drivers. To identify the true cause of an accident during an investigation, explicit measures of transparency need to be implemented beforehand (e.g. recording of training data, event data recorder). In this regard, relevant questions to be answered are:

- What are explicit measures of transparency that allow retrospective investigation of the true cause of an accident where an AV was involved?
- In what way do we need to change regulations on product liability for AVs?
- To what extent should AVs comply with traffic laws?

Non-discrimination & inclusiveness in AVs

In the past, studies have shown that implicit biases and discrimination may unintentionally be incorporated into algorithms. In order to ensure non-discriminatory programming and functioning, the systems of AVs need to be trained and tested for unfair bias. In addition, AVs should exhibit non-discriminatory design in that they are equally usable for and accessible to all individuals. In this regard, relevant questions to be answered are:

- How can companies ensure / test that biases are not incorporated into the systems of their AVs and that certain fairness standards are met?
- What exact features need to be included in the design of AVs to allow accessibility to all individuals?

Effects of AVs on societal & environmental wellbeing:

AVs have great potential to bring forward societal and environmental benefits such as increased mobility, better traffic flows, less congestion and decreased carbon emission. To achieve these benefits, a simultaneous adaption of infrastructure (physical and digital) is needed. In this regard, relevant questions to be answered are:

- How can AVs be combined with other technologies to increase societal and environmental benefits?
- How and where can new mobility models be rolled out and tested (e.g. mixed vehicle traffic flows)?
- What are the main barriers to updating road networks accordingly?

The need to move away from theoretical discussions towards practical recommendations:

Past literature has focused on debating theoretical dilemma situations (e.g. unavoidable accidents) with reference to the famous Trolley problem. Additionally, many policy documents have indicated potential future directions for solving ethical issues that arise from AVs. Although these theoretical considerations and high-level guidelines serve as a good starting point, more concrete, practical recommendations for industry and policymakers need to be developed. Currently, the AI4People Automotive Committee is doing so.