Ethics of Autonomous Cars

Based on:

Holstein, T., Dodig-Crnkovic, G., & Pelliccione, P. (2021). Steps Towards Real-world Ethics for Self-driving Cars: Beyond the Trolley Problem. In Steven John Thompson (Ed.), Machine Law, Ethics, and Morality in the Age of Artificial Intelligence. IGI Global

PART 1
AUTONOMOUS CARS DEVELOPMENT

Source: Wikipedia

Source: Tesla

Source: T. Holstein (taken at IAA 2015)

Cars over the last decades

Complex driving tasks are successively replaced by advanced driving assistant systems

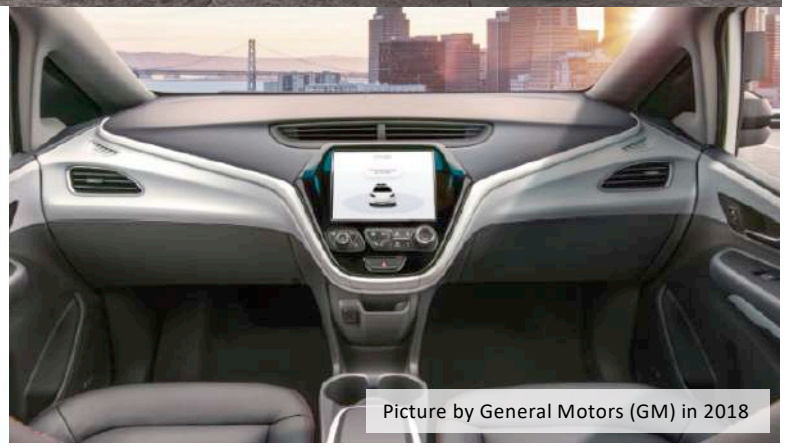
With Self-Driving Cars into the Future

Picture by Google/Waymo in 2015

Picture by General Motors (GM) in 2018

Google I/O in 2016

Picture Source: http://blog.csloxinfo.com/wp-content/uploads/2017/01/google-self-driving-car-2016-photos-15-420-90.jpg

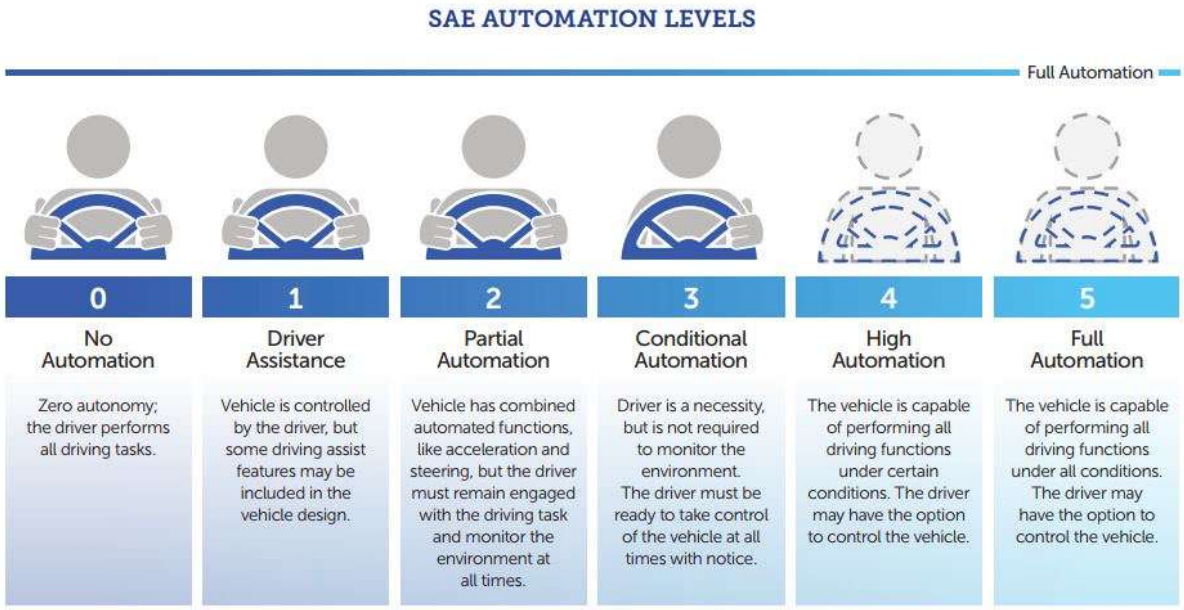


Picture by General Motors (GM) in 2018

Highest level of autonomous driving (Level 5 of 5), where a car can drive from A to B without human supervision.

No Steering-Wheel or other primary driving controls => the former driver becomes solely passenger.

LEVELS OF AUTOMATION



PART 2 CURRENT ETHICS OF AUTONOMOUS CARS – TROLLEY PROBLEM

Current discussions about self-driving cars repeatedly take form of decision-making problem borrowed from philosophy

THE TROLLEY PROBLEM:
whom WILL THE SELF-DRIVING CAR kill when it has to decide?

Source: The New York Times; Illustration by Frank O'Connell

The Trolley Problem Ethical Dilemma

Ethical thought experiment defined by philosopher Philippa Foot in "The Problem of Abortion and the Doctrine of the Double Effect," pp. 5-15, *Oxford Review,* 5, (1967). Focus on the difference between responsibility of acting vs. non-acting.

Many different variants, such as the use of personas to include an emotional perspective. But there is always a single decision: Whom to kill?

Typical Approaches to the Trolley Problem...

…are based on the following ethical theories:

- Utilitarianism
- Other forms of consequentialism
- Deontological ethics

For example, **utilitarianism would aim to minimize casualties**,

even if it means to kill the passenger in the car, by following the principle: **the moral action is the one that maximizes utility**

(or in this case minimizes the damage).

Depending on the ethics framework,

different arguments can be used to justify the decision.

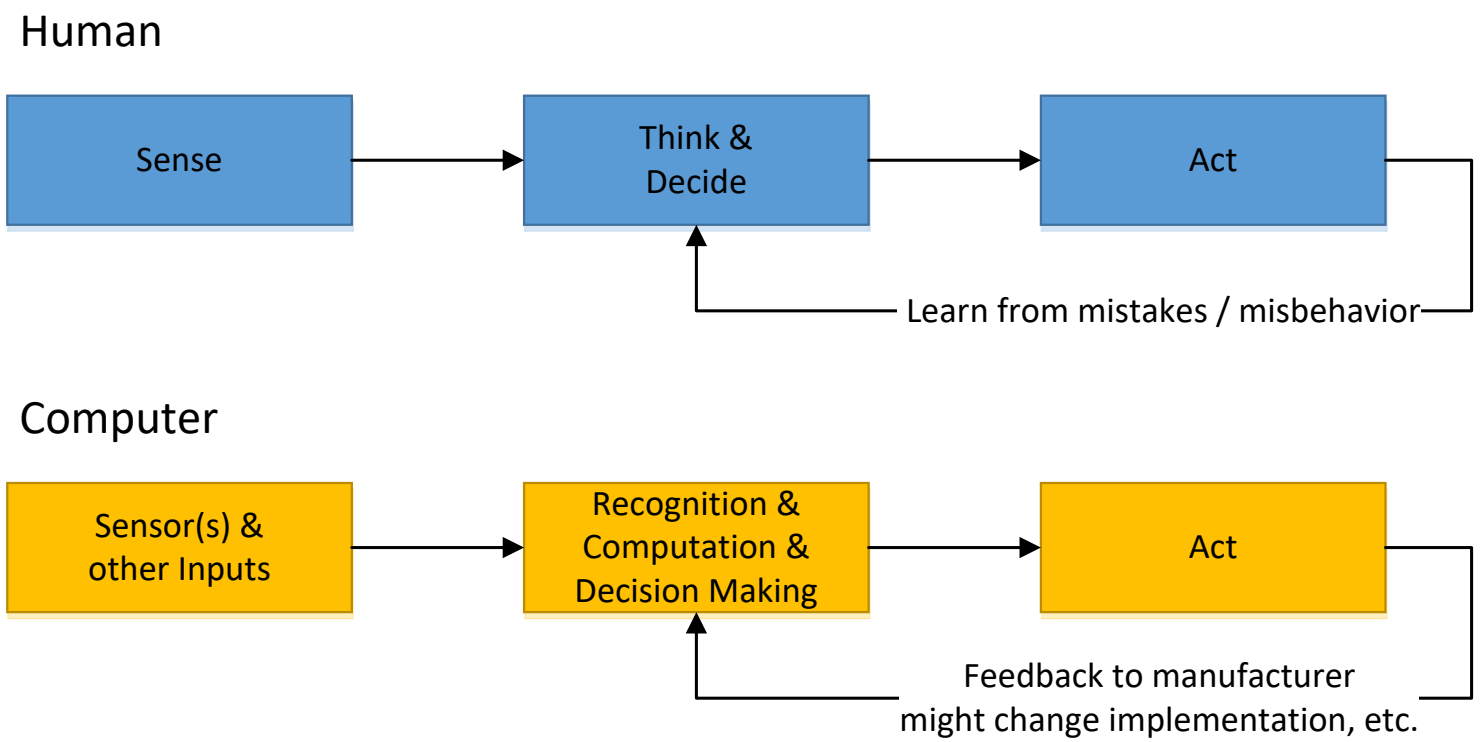
The Trolley Problem is Unsolvable by Construction!

The problem is that for the question “whom to kill?” **all answers are** ethically questionable and perceived as **bad or wrong**.

There is no correct answer to the Trolley Problem and therefore it is not the right kind of problem representative of a real-life situation and even less possible to approach by engineering.

PART 3 THE REAL-WORLD ENGINEERING PROBLEM IS NOT WHOM TO KILL BUT HOW NOT TO KILL!

Human Decision-making Process versus Self-Driving Car (Computer)



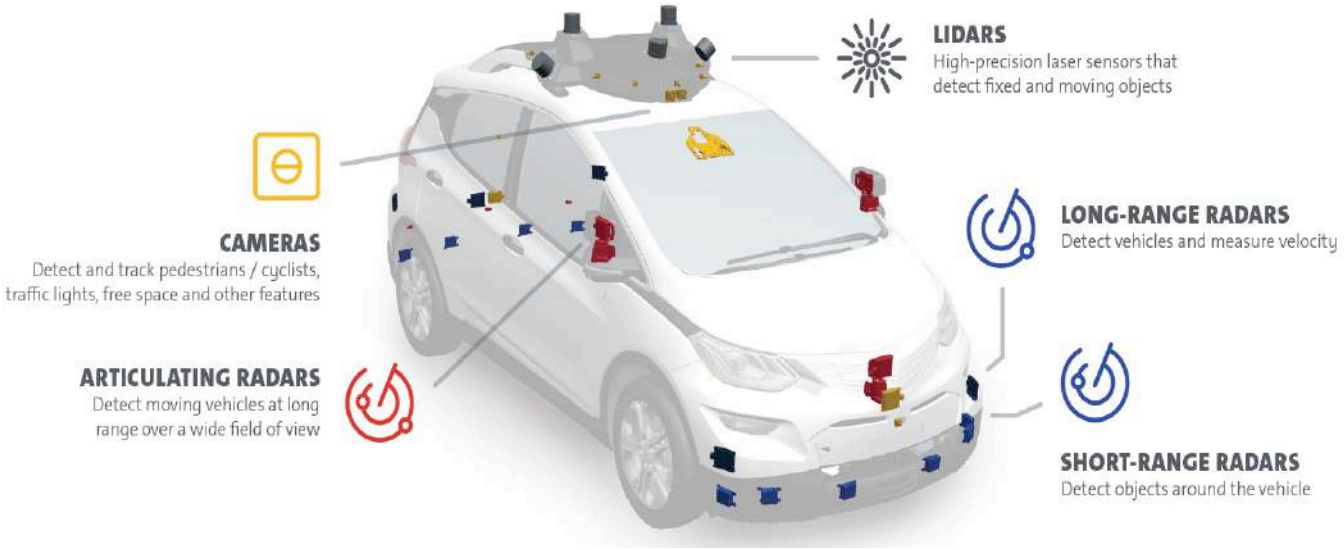
Decision Making in Self-Driving Cars

Decision making process involves sensors, external sources of information, networks, hardware, software, etc.

Environmental influences, such as weather conditions (rain, bright sun, storm, …)

Complex input has to be filtered and only represents an abstraction of the real world.

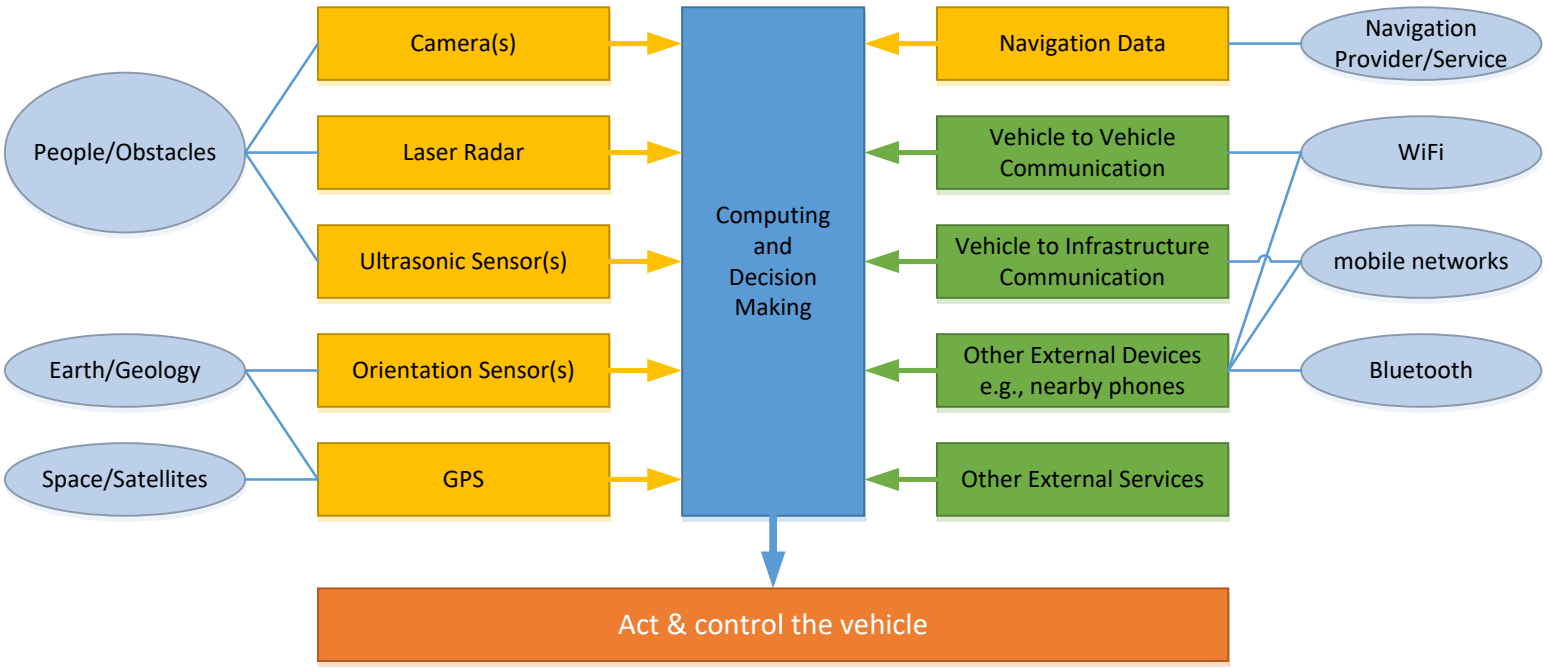
Technical Components



Picture Source: General Motors Safety Report 2018

What a Self-Driving Car “sees”…

Abstract Decision Making Process



This is an outline of what a decision making process might include.
It is based on a literature review and official press releases (Tesla, Google, GM).

PART 4 TECHNICAL CHALLENGES WITH ETHICAL CONSEQUENCES IN AUTONOMOUS CARS

Safety

How can we test self-driving cars? and when is testing sufficient?

Real world vs Abstract World Training of Neural Networks

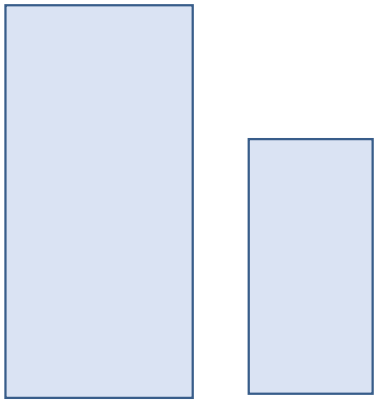
Security

- Attacks against car systems and sensors
- System & security updates
- Do we need a Black Box in self-driving cars like in aircrafts?

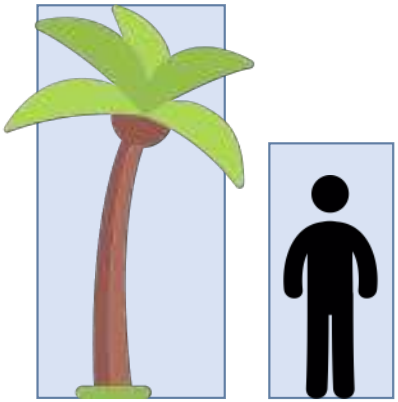
Privacy

- What data should the car have access to?
  - Who will have access to that data?
  - How will the data be used?
- What data is collected?

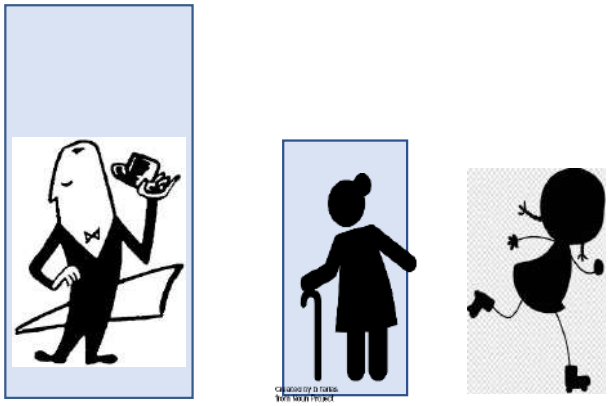
Privacy
What does the car “recognize”?



Objects, different size,
Position, moving or stationary



Objects vs Person(s)



„Everything“ including human identity
– connected to data-bases

Trust

How trustworthy are data sources?

- E.g., GPS, map data, external services
- Trust between self-driving car and services

How trustworthy is the self-driving car?

- E.g., Trust between user and car

Transparency

Multi-disciplinary challenge to ensure transparency, while respecting intellectual property rights, corporate secrets, security concerns, etc.

How much should be disclosed,
and disclosed to whom?

Reliability

- What do we have to rely on?
  - What if sensor(s) fail?
  - What if networks fail?

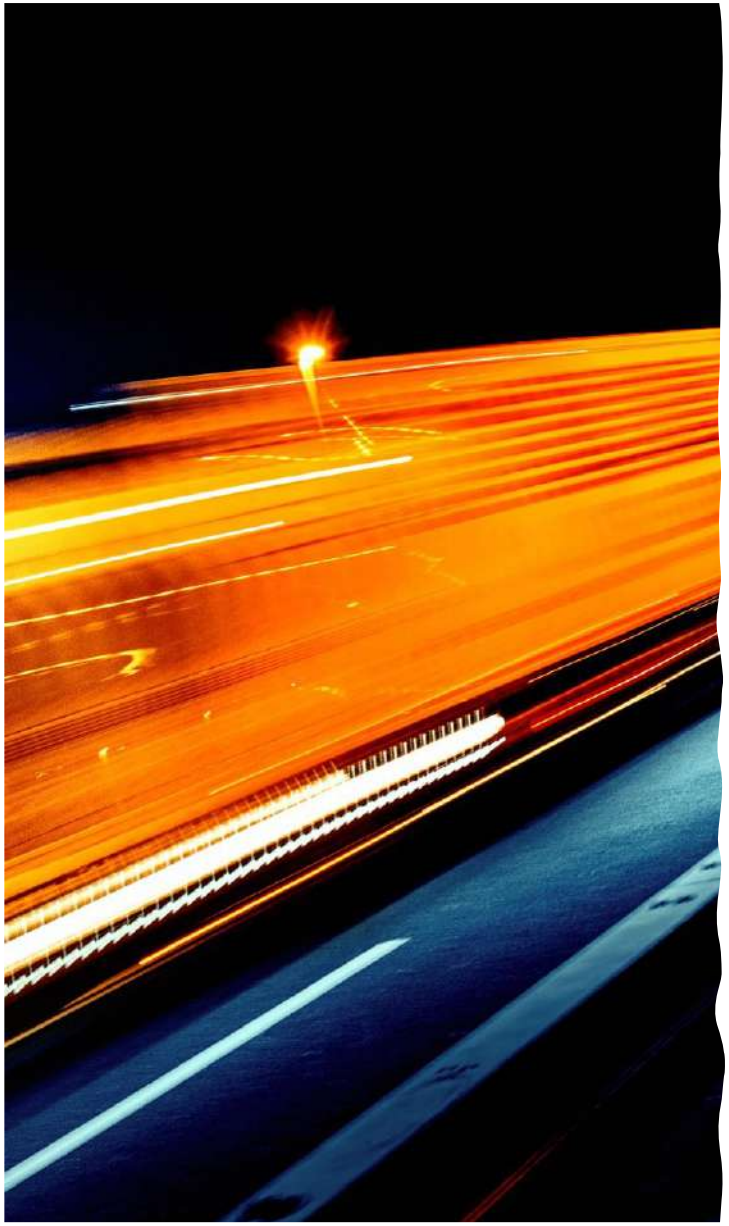
- Redundancy for everything?

Responsibility and Accountability

- Who is responsible and for what?
- Who is accountable and for what?
- How is responsibility distributed among:
  - Developers
  - Car Manufacturers
  - Safety Inspectorates
  - Governmental Institutions
  - Involved participants in the traffic
  - Other stakeholders

Quality Assurance Process

- Lifetime of components
- Maintenance
- Ethics-aware decision making in all processes will help to make ethically justified decisions.



PART 5 SOCIAL CHALLENGES OF AUTONOMOUS CARS WITH ETHICAL CONSEQUENCES

Stakeholders Interests

- Loss of jobs (for cabs/taxi/truck/heavy industrial vehicles drivers)
- Humans in the loop
- Impact on Society

Stakeholders Interests

Freedom of movement

Will the car go, where I want it to go?

Implementation of restrictions

Route to Destination

Can the passenger define the route, or is it determined by the system?

Road trips?

PART 6 Ethical Guidelines for Self-driving Cars

The First Ethical Guidelines For Automated Driving

German Ethics Commission's report comprises 20 propositions. The key elements are:

Automated and connected driving is an ethical imperative if the systems cause fewer accidents than human drivers (positive balance of risk).

Damage to property must take precedence over personal injury. In hazardous situations, the protection of human life must always have top priority.

https://www.bmvi.de/SharedDocs/EN/publications/report-ethics-commission.pdf?__blob=publicationFile

The First Ethical Guidelines For Automated Driving

In the event of unavoidable accident situations, any distinction between individuals based on personal features (age, gender, physical or mental constitution) is impermissible.

In every driving situation, it must be clearly regulated and apparent who is responsible for the driving task: the human or the computer.

It must be documented and stored who is driving (to resolve possible issues of liability, among other things).

Drivers must always be able to decide themselves whether their vehicle data are to be forwarded and used (data sovereignty)

https://www.bmvi.de/SharedDocs/EN/publications/report-ethics-commission.pdf?__blob=publicationFile

“Learning By Experience” And “Proven In Use” Concepts

“Learning by experience” (recording data from autonomous cars) presupposes a functioning socio-technological system that provides strong coupling among legislation, guidelines, standards and use, and promptly adapts to lessons learned.

H. Schäbe and J. Braband. Basic requirements for proven-in-use arguments. CoRR, abs/1511.01839, 2015.

Challenges

- Keeping legislation up-to-date with current level of automated driving, and emergence of self-driving cars
- Creating and defining global legislation frameworks for the implementation of interoperable and development of increasingly automated vehicles
- Defining the guidelines that will be adopted by society for building self-driving cars
- Including ethical guidelines in design and development processes

Holstein, Dodig-Crnkovic, Pellizzione: Ethical and Social Aspects of Self-Driving Cars, ArXives

Recommendations

Car producers supporting and collaborating with legislators in their task to keep up-to-date with the current level of automated driving

Legislative support and contribution to global frameworks to ensure a smooth enrollment of the emerging technology

Include ethics in the overall process of design, development and implementation of self-driving cars. Ensure Ethics training for involved engineers

Establish and maintain a functioning socio-technological system in addition to functional safety standards.

Conclusions

Stop discussing unsolvable ethical dilemmas that obfuscate much bigger actual ethical challenges.

Discuss the real-world ethical challenges surrounding autonomous and driverless vehicles.

Define what is technically possible and ethically justifiable.

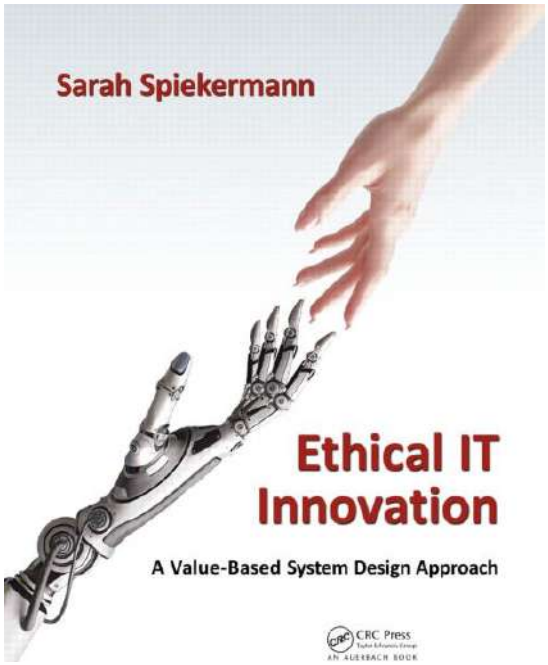
Create transparency to support evaluations by independent organisations/experts.

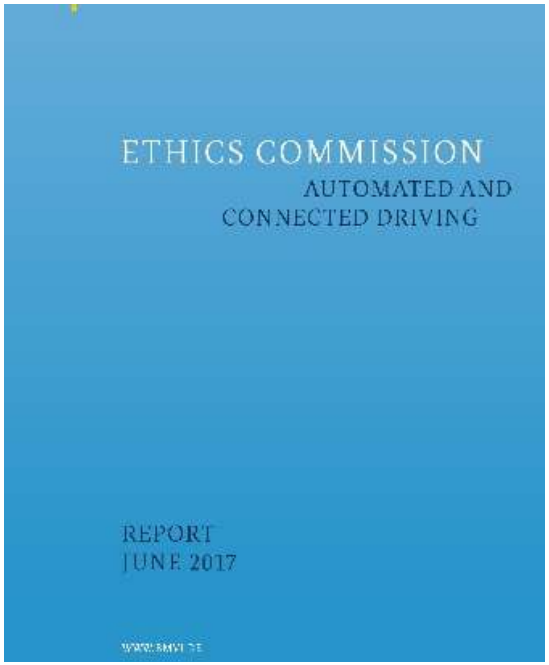
Ethicality as non-functional property?

- Ethicality: the state, quality, or manner of being ethical.

There is already a body of normative documents that can support ethicality of design and implementation.

Conclusions





Existing normative documents

Ethical IT Innovation:
A Value-Based System Design Approach
(by Sarah Spiekermann)

Ethics Commission: Automated and connected driving (Report by Federal Ministry of Transport and Digital Infrastructure of Germany [BMVI])

BMVI = Bundesministerium für Verkehr und digitale Infrastruktur

PART 7 FOR FURTHER READING

Declaration of Amsterdam

- 14 April 2016 EU member states endorsed the Declaration of Amsterdam[1] that addresses legislation frameworks, use of data, liability, exchange of knowledge and cross-border testing for the emerging technology.
- It prepares a European framework[2] (based on report by **Ethics Commission on Automated and Connected Driving**) for the implementation of interoperable connected and automated vehicles by 2019.

- [1]https://www.government.nl/documents/leaflets/2017/05/18/on-our-way-towards-connected-and-automated-driving-in-europe
- [2] https://www.bmvi.de/SharedDocs/EN/publications/report-ethics-commission.pdf?__blob=publicationFile

Legislation and Standards

- Legislation implemented with rigorous monitoring the behavior of cars
- Implementation is within the responsibility of producers. That means that design and implementation of software should follow ethical guidelines.

Ethics & Law Aspects

Autonomous Vehicles Ethics & Law: Towards an Overlapping Consensus

https://www.academia.edu/29332066/Autonomous_Vehicles_Ethics_and_Law_Towards_an_Overlapping_Consensus

Patrick Lin: Why Ethics Matters for Autonomous Cars.

In: Autonomes Fahren Technische, rechtlische und gesellschaftliche Aspekte

https://www.springerprofessional.de/en/why-ethics-matters-for-autonomous-cars/4397684

Ethically Aligned Design

A Vision for Prioritizing Human Well-being With Autonomous and Intelligent Systems

https://ethicsinaction.ieee.org/

Embedding Values into Autonomous Intelligent Systems - The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems

https://standards.ieee.org/develop/indconn/ec/ead_embedding_values.pdf

An example of ethical guidelines thinking one step further is described in the book:

Sarah Spiekermann. Ethical IT Innovation: A Value-Based System Design Approach. Taylor & Francis, 2015.

Policy Concerning Automated Vehicles (US DOT)

"DOT/NHTSA Policy statement concerning Automated Vehicles" 2016 update to "Preliminary statement of policy concerning automated vehicles".

Technical report, National Highway Traffic Safety Administration (NHTSA).

http://www.nhtsa. gov/staticfiles/rulemaking/pdf/Autonomous-Vehicles-Policy-Update-2016.pdf

References

- Holstein, T., Dodig-Crnkovic, G., & Pelliccione, P. (2021). Steps Towards Real-world Ethics for Self-driving Cars: Beyond the Trolley Problem. In Steven John Thompson (Ed.), Machine Law, Ethics, and Morality in the Age of Artificial Intelligence. IGI Global
- Holstein, T., Dodig-Crnkovic, G., & Pelliccione, P. (2020). Real-world Ethics for Self-Driving Cars. In Proceedings of the 42nd International Conference on Software Engineering (ICSE '20) Poster Track. https://ethics.se
- Holstein, T. Dodig-Crnkovic G. Avoiding the Intrinsic Unfairness of the Trolley Problem. Accepted for the Proceedings of FairWare workshop at ICSE2018, to be published by ACM.
- Holstein, T. Dodig-Crnkovic G. and Pelliccione P. Ethical and Social Aspects of Self-Driving Cars, http://arxiv.org/abs/1802.04103
- Dodig Crnkovic, G. and B. Çürüklü. Robots: ethical by design. Ethics and Information Technology, 14(1):61–71, Mar 2012.
- Dodig Crnkovic, G. and B. Çürüklü. Robots: ethical by design. Ethics and Information Technology, 14(1):61–71, Mar 2012.
- Dodig-Crnkovic, G. and D. Persson. Sharing moral responsibility with robots: A pragmatic approach. In Proceedings of the 2008 Conference on Tenth Scandinavian Conference on Artificial Intelligence: SCAI 2008, pages 165–168, Amsterdam, The Netherlands, IOS Press. 2008.
- Dodig-Crnkovic, G. and D. Persson. Sharing moral responsibility with robots: A pragmatic approach. In Proceedings of the 2008 Conference on Tenth Scandinavian Conference on Artificial Intelligence: SCAI 2008, pages 165–168, Amsterdam, The Netherlands, IOS Press. 2008.
- Johnsen A., G. Dodig- Crnkovic, K. Lundqvist, K. Hänninen, and P. Pettersson. Risk- based decision-making fallacies: Why present functional safety standards are not enough. In 2017 IEEE International Conference on Software Architecture Workshops (ICSAW), pages 153–160, April 2017.
- Sapienza, G., Dodig-Crnkovic, G. and I. Crnkovic. Inclusion of ethical aspects in multi-criteria decision analysis. In 2016 1st International Workshop on Decision Making in Software ARCHitecture (MARCH), pages 1–8, April 2016.
- Thekkilakattil A. and G. Dodig-Crnkovic. Ethics aspects of embedded and cyber-physical systems. In 2015 IEEE 39th Annual Computer Software and Applications Conference, volume 2, pages 39–44, July 2015.