



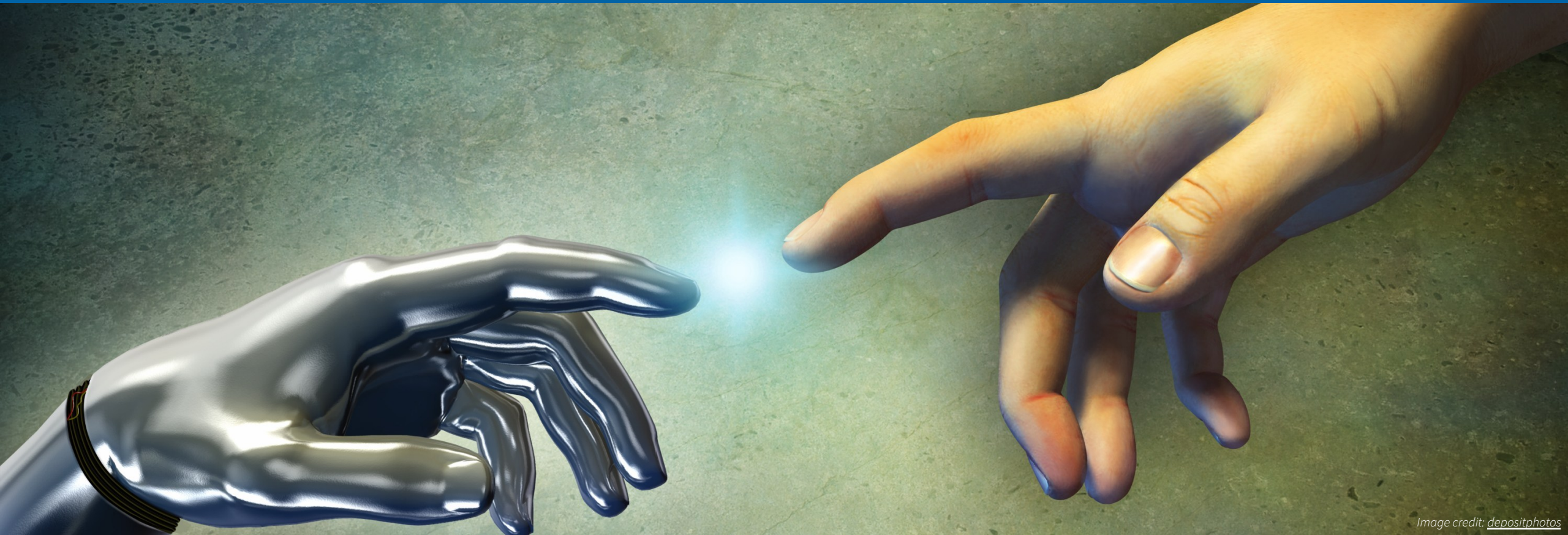
Laura Crompton

Doktorandin

Universität Wien

FoNTI

Vom Elfenbeinturm in die Praxis: was uns die Philosophie über autonomes Fahren sagen kann



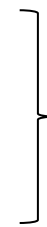
Ethik, Moral, Norm, ... – ein kleiner Überblick

- Moralische Überzeugungen als Ausgangspunkt der Ethik
→ Was ist gut? Was ist schlecht/unzulässig?

- Zwei Aufgaben der Ethik:

- 1) Allgemeine Kriterien für *gut*, *schlecht*, *richtig*, *fair*, etc. herausarbeiten
- 2) Hilfestellung bei Widersprüchen oder Unsicherheiten bieten

(Nida-Rümelin, 2005)



Ethische Kriterien
sind normativ

Zur Komplexität der Ethik

Unterschiedliche ethische Paradigmen, z.B.:

Utilitarismus (moral. Begründung durch Vermehrung von Lust, Verminderung von Leid); **Kantische Ethik** (moral. Begründung durch kategorischen Imperativ); **Tugendethik** (moral. Begründung durch Tugend), **Libertarismus** (moral. Begründung durch Individualrechte), etc. (Nida-Rümelin, 2005)

Herausforderungen:

- ➔ Welches ethische Paradigma sollen wir verfolgen? Welche Prinzipien sollen wir verfolgen?
 - ➔ Welche Werte können wir gegeneinander abwägen? Welche Werte sollen wir priorisieren?
 - ➔ Welche normative Kriterien können welche praktischen Fragestellungen/Probleme angemessen adressieren?
-

Aufteilung der Ethik in **Bereichsethiken** (z.B. Medizinethik, Umweltethik, Tierethik, Rechtsethik, Gen-Ethik, etc.) → **Technikethik als Bereichsethik**: Technikbewertung und Technikfolgenabschätzung (Ott, 2005)

Technikethik

Allgemeine ethische Herausforderungen für **i) Mensch** (Arbeitsplatzverluste; Terminator-Szenarien; zwischenmenschliche Interaktion), **ii) Umwelt** (Energieverbrauch; Öl-Bohrungen; nachhaltige Fischerei), **iii) gesellschaftliche/soziale Grundstrukturen** (Demokratie; Recht -> Verantwortungsfragen; Wirtschaft)

→ wie weit wollen vs. sollen wir gehen? (normativ! → Recht?)

„Trendfragen“ für Technik aus ethischer Sicht

Verantwortungsfragen, Bias, Undurchsichtigkeit der Systeme (*transparency* und *explainability*), Umsetzung von ethischen und moralischen Prinzipien in Code, etc.

Ethik des autonomen Fahrens

Verantwortung: welche Akteure sind involviert? (Fahrer, Automobilhersteller, Programmierer, und evtl. Staatliche Instanz (StVO)); Moralische vs. Rechtliche Verantwortung; *Human in the loop?*

Bias: Benachteiligung und Ausschließung von best. gesellschaftlichen Gruppen? Sozio-ökonomischer Bias (in Entwicklung von Auto; Zugang zu autonom. Fahrzeugen?)/Geschlechter-Bias (in Entwicklung von Auto, Werbung)/Rassistischer-Bias (in Entwicklung von Auto; aus technischer Sicht; Zugang zu autonom. Fahrzeugen)

Vertrauen: Können und sollen wir dem Fahrzeug vertrauen? (Hacking, Fehlfunktion, ungewollte autom. Updates), gerade in Anbetracht von *transparency* und *explainability*!

Umsetzbarkeit von ‚moralisch‘-autonomen Auto: welche Paradigmen?; Ungelöste moralisch, ethische (und rechtl.) Dilemmata; Rationalität und moralische Handlungsgründe überhaupt in Code übersetzbar?

- **Verantwortungsfrage** direkt mit moral agency verbunden: können und wollen wir technologische moral agents? (Siehe Wallach & Allen, 2008; Asaro, 2006)
- **Bias** als ‚Spiegel‘ unserer Gesellschaft: (wie) können wir Bias aus den Systemen bringen? (Coeckelbergh, 2020)
- *Transparency* und *Explainability* als Herausforderungen für *Trustworthy AI* (HLEG, 2019)
- Kann Ethik jemals Kriterien finden, die so gelten können, dass wir autonomes Auto allgemein als ethisch, moralisch OK ansehen können? Landesweit? EU weit? Weltweit? Unterschiedliche StVOs verkörpern eventuell unterschiedliche ethische Prinzipien?



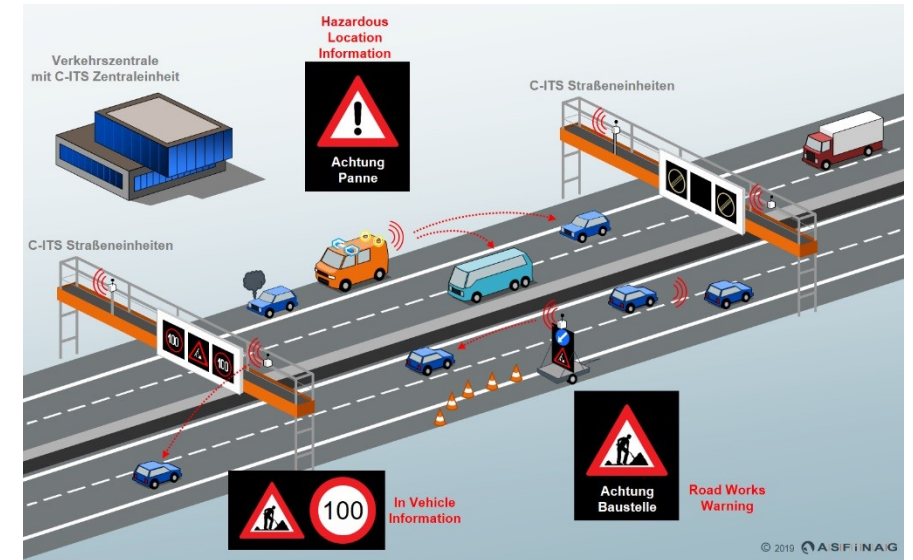
Vielfalt ethischer Paradigmen führt zu allgemeinen Herausforderungen in philosophischer Betrachtung (ungelöste Dilemmata, Wertkonflikte,...)

Use Case

Implementierung & automatisiertes updaten von StVO (rechtl. bindend) in autonomem Fahrzeug bei normaler Fahroutine prinzipiell als weniger (moralisch) herausfordernd (Loh & Loh, 2017)

⚡ ABER: adressiert digitale StVO mögliche moralische/ethische Dilemmata (siehe z.B. Trolley-Probleme)?

- Was deckt digitale StVO ab?
- Was deckt Hard- & Software des Autos ab?
- Deckt Fahrer etwas ab, und wenn ja, was?



Literatur

Julian Nida-Rümelin, *Angewandte Ethik – Die Bereichsethiken und ihre theoretische Fundierung*. Kröner Verlag, Stuttgart, 2005.

Konrad Ott. Technikethik. In Julian Nida-Rümelin, editor, *Angewandte Ethik – Die Bereichsethiken und ihre theoretische Fundierung*, Kröner Verlag, Stuttgart, 2005.

Mark Coeckelbergh. *AI Ethics*. The MIT Press, Cambridge, MA, 2020. ISBN 0262538199.

AI HLEG. *ETHICS GUIDELINES FOR TRUSTWORTHY AI* [ebook]. European Commission, Brussels, 2019. Available at:
<<https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai>>

Peter Asaro. What Should We Want From a Robot Ethic?, *International Review of Information Ethics*, 6, December 2006

Wendell Wallach and Colin Allen. *Moral Machines*. Oxford University Press, New York, 2009. 10.1093/acprof:oso/9780195374049.001.0001

Janina Loh and Wulf Loh. Autonomy and responsibility in hybrid systems –the example of autonomous cars. In Patrick Lin, Keith Abney, and Ryan Jenkins, editors, *Robot Ethics 2.0: From Autonomous Cars to Artificial Intelligence*. Oxford University Press, New York, 2017.
