

Ethik am Steuer

Einführung

Wir stehen heute an einem Punkt, an dem die Mobilität, wie wir sie kennen, vor einem fundamentalen Umbruch steht. Der Übergang vom menschlichen Fahrer hin zu Systemen, die von Künstlicher Intelligenz gesteuert werden, ist längst kein Ding der Unmöglichkeit mehr, sondern Realität. (*SAE Levels of Driving Automation™ Refined for Clarity and International Audience*, 2021) Dies bringt jedoch ein schwieriges Paradoxon mit sich: Einerseits kann die KI Unfälle massiv reduzieren, da sie weder Ermüdung noch Ablenkung kennt. (*Künstliche Intelligenz: Chancen Und Risiken*, 2020) Andererseits zwingt uns genau das dazu, sie für den absoluten Ernstfall zu programmieren. Für jene Momente, in denen eine Kollision schlicht unvermeidbar ist.

Während wir Menschen in solchen Sekundenbruchteilen meist instinktiv oder reflexartig reagieren, trifft die KI eine Entscheidung, die vorab von Menschen trainiert und so bestimmt wurde. Damit verändert sich das Fahrzeug von einem reinen Werkzeug hin zu einem moralischen Akteur im Verkehr. Das wirft Fragen auf: Wer soll geschützt werden, wenn es kein eindeutiges ‚Richtig‘ mehr gibt, sondern nur noch die Abwägung verschiedener Schadensprofile? (*KI Im Straßenverkehr: Vorteile Und Herausforderungen Für Die Vision Zero — DGUV Forum*, n.d.)

An dieser Stelle treffen unterschiedliche Disziplinen aufeinander. In der Informatik liegt der Fokus darauf, wie die KI durch Deep Learning und Simulationen lernt, (*Was Ist Deep Learning?*, n.d.) Millionen von Szenarien in Echtzeit zu bewerten. Doch die Logik hinter diesen Prozessen benötigt ein ethisches Fundament.

Müssen wir utilitaristisch agieren und versuchen, die grösstmögliche Zahl an Menschenleben zu retten, selbst wenn dies bedeutet, ein Individuum gezielt zu opfern? Oder soll die Programmierung einer Deontologie folgen, die besagt, dass bestimmte Handlungen an sich falsch sind und Menschenleben niemals gegeneinander aufgewogen werden dürfen? (Sinnott, 2003) (*Automatisiertes Fahren*, 2025)

Diese Weichenstellung hat massive rechtliche Konsequenzen, denn schliesslich stellt sich die ganz praktische Frage der Haftung. Wer trägt die Verantwortung, wenn das System versagt? (*Künstliche Intelligenz Im Straßenverkehr – Chancen & Risiken*, 2025) Ist es der Hersteller, der Programmierer oder der Halter des Fahrzeugs? Wir müssen einen Konsens finden, um die Verantwortung für diese neue Technologie tragen zu können.

Deontologie

Die erste Theorie, welche es zu erläutern gibt, ist die Deontologie, sie ist auch die pflichtethische Theorie. Dabei steht die Frage: "Was soll ich tun?", im Vordergrund. (Meyer, n.d.) Also ist nicht das Ergebnis einer Handlung entscheidend, sondern die Handlung selbst. Dabei wird das Ergebnis aussen vor gelassen. Bei der Handlung geht es dabei um richtig und falsch, dabei wird nach Regeln, Pflichten und moralischen Grundsätzen bewertet. Die Moral wird immer wieder ins Zentrum gerückt, stark geprägt hat diese Ethik Immanuel Kant. Er betonte, dass das moralische Handeln aus Pflichtbewusstsein entstehen sollte.

(*Pflichtethik*, n.d.) Diese Pflicht basiert auf einem Gesetz, das je nach Zeit unterschiedlich verstanden wurde. Bei Kant kommt dieses Gesetz aus der eigenen Vernunft. Die Vernunft ist hier zentral, weil sie dem Menschen ermöglicht, moralische Gesetze selbst zu erkennen und aufzustellen und ihn unabhängig von Gefühlen oder äusseren Einflüssen macht.

(*Pflichtethik*, n.d.) Doch gibt es auch Kritik daran, dass zwischenmenschliche Beziehungen zu wenig berücksichtigt werden. Denn auch wenn eine Handlung nützliche Folgen hätte oder einer zwischenmenschliche Beziehung gut tun würde, sie aber "falsch" sind, dürfte sie trotzdem nicht getätigt werden.

Genau dies ist nun Interessant anzuschauen in Bezug auf KI-gesteuerte Fahrzeuge. Denn wir Menschen programmieren sie und geben dabei Regeln und Gesetze weiter. Vor allem in einer Ausnahmesituation, bei der es zu einem Unfall kommen könnte, eine Entscheidung getroffen werden muss oder sogar eine Abwägung in Betracht gezogen wird, handelt das Programm nach seinem Inhalt und Gelernten und dabei spielt die Art des Gelernten eine Rolle. Wenn das Fahrzeug nach deontologischer Art programmiert wäre, dann würde es sich in kritischen Situationen vor allem an festen Regeln orientieren und nicht daran, welche Lösung am meisten Menschenleben rettet. In diesem Fall wäre zum Beispiel wichtig, dass das Auto bestimmte Handlungen grundsätzlich vermeidet, etwa Menschen absichtlich zu gefährden oder Leben gegeneinander abzuwägen. (*Ethische Fragen Zum Vernetzten Und Automatisierten Fahren*, 2023) Laut ADAC ist es klar, dass ein Fahrzeug in einer Dilemma-Situation nicht abwägen darf und soll. Es sollte lieber defensiv und vorausschauend fahren. Ein deontologisch programmiertes Fahrzeug würde also eher nach dem Grundsatz handeln, dass bestimmte Dinge verboten sind, auch wenn dadurch vielleicht ein grösserer Gesamtnutzen entstehen könnte.

Doch zu erkennen ist, dass ein KI-gesteuertes Fahrzeug regelgetreuer handeln kann und wahrscheinlich auch wird. Wir können also dem Fahrzeug die Pflichtethik im Sinne von Regeln und Verhaltensmechanismen mitgeben. Doch eine solche starre Regelbefolgung kann natürlich eine gewisse Unflexibilität mit sich bringen, die sich negativ auf den Verkehr auswirken kann, welcher von Spontanität, Verantwortung und menschlichem Verhalten geprägt ist. ("Das Auto, Das Entscheiden Muss, Ob Es Alte Oder Kinder überfährt," 2018) Ein deontologisch programmiertes Fahrzeug im Vergleich zu einem menschlichen Fahrer ist natürlich viel eingeschränkter, was die Komplexität der menschlichen Moral angeht. Daher kann pflichtethisch gehandelt werden, aber nicht immer im besten Sinne, mit den besten Folgen. Dafür gilt es sich nun auch den Utilitarismus anzuschauen.

Utilitarismus

Der Utilitarismus ist eine der einflussreichsten Strömungen der modernen Ethik. Das Grundprinzip ist so simpel wie radikal. Er besagt, dass eine Handlung dann moralisch richtig ist, wenn sie das Glück der grösstmöglichen Zahl maximiert und somit das Leid minimiert. Anstatt sich auf starre Verbote oder religiöse Dogmen zu stützen, bewertet der Utilitarismus Taten ausschliesslich nach ihren Folgen. Der Wert einer Handlung bemisst sich also an ihrem Nutzen (utilitas). Wobei das Wohlergehen aller von der Handlung betroffenen Personen gleichermassen berücksichtigt werden muss. Die Theorie wurde von zwei Philosophen geprägt, die unterschiedliche Schwerpunkte setzen. Jeremy Bentham gilt als der Begründer des klassischen Utilitarismus. Bentham verfolgte einen rein quantitativen Ansatz. Für ihn war Glück gleichbedeutend mit Lust und der Abwesenheit von Schmerz. Um den moralischen Wert einer Handlung zu berechnen, entwarf er das hedonistische Kalkül, eine Art mathematische Formel zur Messung der Intensität und Dauer von Glück. John Stuart entwickelte die Theorie von Bentham weiter. Er führte eine qualitative Entscheidung ein. Er argumentierte, dass geistige und intellektuelle Freuden wertvoller seien als rein körperliche Genüsse. (Assmann & Bergmann, 2004)

Wenn wir nun ein KI gesteuertes Fahrzeug auf der Grundlage utilitaristischer Werte programmieren, erschaffen wir einen unbestechlichen, moralischen Rechenmeister. Das Ziel dieses Fahrzeuges ist es, in jeder Sekunde die Summe des menschlichen Leids zu minimieren und das allgemeine Wohl zu maximieren. Im alltäglichen Fahrbetrieb würde ein menschlicher Fahrer oft dazu neigen, den eigenen Vorteil über die Bedürfnisse anderer zu stellen. Zum Beispiel die schnellere Ankunft oder den bequemsten Weg. Die KI würde ihre Geschwindigkeit und Routenwahl so anpassen, dass der Verkehrsfluss für alle Teilnehmer optimiert wird. Zum Beispiel in dem sie einen Umweg in Kauf nimmt, wenn dann ein Stau für andere verhindert werden kann. (*Ethische Entscheidungen Autonomer Fahrzeuge* | Bidt, 2024) Die wahre moralische Tragweite offenbart sich jedoch erst im Moment einer unvermeidbaren Katastrophe. In einem Szenario, in dem ein Unfall nicht mehr verhindert werden kann. Ein solches Szenario kann mit dem klassischen „Trolley-Problem“ verglichen werden. Ein utilitaristisch programmiertes Fahrzeug kennt keine Loyalitäten gegenüber seinem Fahrer. Müsste die KI entscheiden in eine Menschenmenge zu steuern oder ein Ausweichmanöver zu machen, was den sicheren Tod des Fahrers heissen würde, würde die KI ohne Zögern den Fahrer opfern. (Knollmann, 2020) In der Logik von Jeremy Bentham wiegt die Rettung von fünf Passanten schwerer als der Schutz einer einzelnen Person. Ausserdem würde die KI Faktoren wie die Schwere von Verletzungen gegeneinander aufrechnen. Wenn die KI die Wahl hat, einen Radfahrer schwer zu verletzen oder drei Insassen leichte Prellungen zuzuführen, entscheidet es sich für die Variante, die in der Summe das wenigste Leid verursacht. Dabei bleibt die Maschine unbeeindruckt von sozialen Bedingungen oder Selbsterhaltungstrieben. (Bentham & Dalege, 2025)

Ein Fahrzeug, das so programmiert ist, zwar die moralisch rationalste Lösung zur Schadensminimierung bietet. Uns aber vor ein tiefes psychologisches Dilemma stellt. Während das System durch seine objektive Kalkulation die Gesamtzahl der Opfer senken würde, opfert es gleichzeitig das individuelle Sicherheitsversprechen. Ein solches Fahrzeug wäre somit der perfekte altruistische Verkehrsteilnehmer. Doch es bleibt fraglich, ob eine Gesellschaft bereit ist, persönliche Sicherheit gegen mathematische Gemeinwohlmaximierung einzutauschen. („Moralisches Dilemma Auch Bei Autonomen Autos,” 2018)

Technische Umsetzung und gesellschaftlicher Prozess

Die technische Grundlage des automatisierten Fahrens basiert auf der Verbindung von Sensoren, einem eindeutigen Rechtsrahmen, intelligenter Datenverarbeitung und lernenden Algorithmen. Die verschiedenen Arten von Sensoren wie Radar, Kameras, LIDAR oder auch GPS erfassen die Umgebung des Fahrzeugs fortlaufend. Die von den Sensoren gesammelten Daten liefern die Grundlage für fahrdynamische Entscheidungen.

Sicherheitskritische Systeme sind mehrfach ausgelegt, um die Betriebssicherheit zu gewährleisten, sodass sie auch bei Ausfällen funktionsfähig bleiben. (Automatisiertes Fahren, 2025)

Algorithmen der künstlichen Intelligenz verarbeiten die erfassten Daten, identifizieren Objekte sowie Verkehrssituationen. Dafür greifen die Algorithmen auch auf hochpräzise Echtzeitkarten zurück. In Zukunft soll diese Datenbasis durch die Kommunikation mit Infrastruktur und anderen Verkehrsteilnehmenden erweitert werden. Etwa zur Übermittlung von Ampelphasen, Verkehrszeichen oder Störmeldungen. Dafür ist jedoch noch eine entsprechende Anpassung der Infrastruktur erforderlich. (Automatisiertes Fahren, 2025)

Künstliche Intelligenz lernt das autonome Fahren, indem sie diese Daten durch Methoden wie Deep Learning verarbeitet. ("Deep Learning nutzt mehrere verarbeitende Ebenen (Schichten), wodurch komplexe Muster erkannt werden können." (Mildner et al., 2024, 71))

Ein zentraler Bestandteil des Lernprozesses sind Simulationen. In diesen Simulationen werden etliche Verkehrssituationen und -szenarien erzeugt. Die KI lernt dadurch unterschiedlichste Situationen zu bewerten. Neueste Ansätze der Simulation erzeugen sogar gezielt riskante und selten gefährliche Situationen, welche selbst in echten Daten kaum vorkommen. (Wang et al., 2026,)

Durch die Simulationen wird auch das Verhalten in Dilemma Situationen trainiert. Das Problem dabei ist, dass die KI keinen eigenen moralischen Kompass besitzt. Sie übernimmt die moralischen Werte, mit denen sie trainiert wurde. (Müller, 2020) Wenn ein System also darauf trainiert wird, möglichst viele Menschenleben zu retten, folgt es dem Utilitarismus.

Wird jedoch festgelegt, dass bestimmte Handlungen grundsätzlich ausgeschlossen werden, folgt es der Deontologie.

Ein konkretes Szenario könnte beispielsweise so aussehen:

Ein autonomes Fahrzeug fährt einer Strasse entlang und aus dem Nichts springen 5 Kinder auf die Strasse, um einem Schmetterling hinterherzujagen. Das Ausweichen ist nur möglich, wenn das Fahrzeug in eine Mauer neben der Strasse lenkt. Dabei würde die Person im Auto jedoch verunglücken.

Wäre das Fahrzeug utilitaristisch trainiert, so würde es ausweichen und in die Mauer fahren, da so das grösstmögliche Wohl für die meisten Personen in Kraft treten würde.

Wäre das Fahrzeug deontologisch trainiert und die Handlung, Menschen gezielt zu töten (in die Mauer fahren), ausgeschlossen wäre, würde das Fahrzeug nicht in die Mauer fahren und folglich die fünf Kinder überfahren.

An diesem Beispiel wird deutlich, dass autonome Fahrzeuge nicht nur eine technische, sondern genauso eine gesellschaftliche und ethische Herausforderung darstellen. Die Frage, nach welchem ethischen Modell die KI handeln soll, kann und darf nicht allein in der Verantwortung der Ingenieure liegen. Stattdessen muss diese Frage mit einem gesellschaftlichen Diskurs geklärt werden. Die Ethik-Kommission automatisiertes und vernetztes Fahren hat dafür ethische Leitfaden für autonome Fahrzeuge entwickelt.

(Könneker, 2017) Darin wird unter anderem festgehalten, dass die Systeme so entwickelt werden sollen, dass kritische Situationen möglichst gar nicht erst entstehen oder dass echte

Dilemma-Situationen moralisch nicht eindeutig programmiert werden können und Menschen weder nach Geschlecht noch Alter oder sonstigen Merkmalen bewertet werden dürfen.
(Könneker, 2017)

Fazit und Erwartung

Autonome Fahrzeuge verpflichten uns dazu, eine Verantwortung zu übernehmen, welche bislang dem menschlichen Instinkt oder dem Zufall überlassen wurde. Durch die Programmierung der KI wird verdeutlicht, dass moralische Entscheidungen klar getroffen werden müssen, wodurch die gesellschaftliche Auseinandersetzung mit diesen Fragen zunehmend in den Vordergrund rückt.

Vermutlich wird sich ein gesellschaftlicher Konsens nicht auf eine einzige, allgemeingültige moralische und ethische Regel stützen lassen. Vielmehr ist ein vielschichtiger Ansatz wahrscheinlicher, in welchem verschiedene ethische Prinzipien miteinander verbunden werden.

Es braucht klare rechtliche Gegebenheiten, welche international abgestimmt werden müssen und beispielsweise die Verantwortung definieren und gewisse Standards festlegen. Auch sollte Transparenz bei den Entscheidungsprozessen autonomer Fahrzeuge gegeben sein, damit diese nachvollziehbar sind.

Ebenso sollte die Gesellschaft aktiv in die Diskussion über ethische Fragen bei autonomen Fahrzeugen einbezogen werden. Diese würden dazu führen, dass gemeinsam definierte Werte in die Programmierung einfließen.

Künftig wird sich wahrscheinlich ein gesellschaftlich akzeptierter Mittelweg durchsetzen, bei welchem die Systeme der autonomen Fahrzeuge darauf ausgerichtet sind, die Schäden so klein wie möglich zu halten und dabei keine ethischen Prinzipien zu verletzen.

Quellenangabe

Assmann, L., & Bergmann, R. (2004). *Zugänge zur Philosophie* (R. W. Henke, Ed.).

Cornelsen.

Automatisiertes Fahren. (2025). Bundesamt für Strassen (ASTRA). Retrieved May 7, 2026,

from

<https://www.astra.admin.ch/astra/de/home/themen/intelligente-mobilitaet/automatisiertes-fahren.html>

Bentham, J., & Dalege, M. (2025, February 20). *Philosoph des Monats: Jeremy Bentham*.

Reli-Ethik-Blog. Retrieved June 5, 2026, from

<https://reli-ethik-blog.de/philosoph-des-monats-jeremy-bentham/>

Das Auto, das entscheiden muss, ob es Alte oder Kinder überfährt. (2018). *Die Zeit*.

Ethische Entscheidungen autonomer Fahrzeuge | bidt. (2024, August 12). Bayerisches

Forschungsinstitut für Digitale Transformation. Retrieved June 5, 2026, from

<https://www.bidt.digital/phaenomene/ethische-entscheidungen-autonomer-fahrzeuge/>

Ethische Fragen zum vernetzten und automatisierten Fahren. (2023, April 17). ADAC.

Retrieved June 5, 2026, from

<https://www.adac.de/rund-ums-fahrzeug/ausstattung-technik-zubehoer/autonomes-fahren/recht/ethische-fragen/>

Hoffmann-Kolss, V., & Rathgeb, N. (Eds.). (2024). *Handbuch Philosophie des Geistes*. J.B.

Metzler.

KI im Straßenverkehr: Vorteile und Herausforderungen für die Vision Zero — DGUV forum.

(n.d.). DGUV forum. Retrieved June 5, 2026, from

<https://forum.dguv.de/ausgabe-3-2025/ki-im-strassenverkehr-vorteile-und-herausforderungen-fuer-die-vision-zero/>

Knollmann, D. (2020, August 31). *Autonoblog*. Autonoblog. Retrieved June 5, 2026, from

https://www.hs-osnabrueck.de/fileadmin/HSOS/Homepages/Deframe/20190322_Autonoblog_Ethik___autonomes_Fahren_II___Trolley-Probleme.pdf

- Könneker, C. (2017). *WWW.BMVI.DE*. BMV. Retrieved May 7, 2026, from https://www.bmv.de/SharedDocs/DE/Publikationen/StV/bericht-der-ethik-kommission.pdf?__blob=publicationFile
- Künstliche Intelligenz: Chancen und Risiken*. (2020, September 29). Themen Europäisches Parlament. <https://www.europarl.europa.eu/topics/de/article/20200918STO87404/kunstliche-intelligenz-chancen-und-risiken>
- Künstliche Intelligenz im Straßenverkehr – Chancen & Risiken*. (2025, April 24). Björn Steiger Stiftung. Retrieved June 5, 2026, from <https://www.steiger-stiftung.de/blog/kuenstliche-intelligenz-im-strassenverkehr/>
- Meyer, L. (n.d.). *Pflichtenethik*. Ethik-Lexikon. Retrieved June 5, 2026, from <https://www.ethik-lexikon.de/lexikon/pflichtenethik>
- Mildner, R., Ziller, T., & Baiocchi, F. (2024). *Car IT kompakt Reloaded: Das Auto der Zukunft – Vernetzt und autonom fahren*. Springer Vieweg. Retrieved May 7, 2026, from https://link.springer.com/chapter/10.1007/978-3-658-44319-1_2#citeas
- Moralisches Dilemma auch bei autonomen Autos. (2018, October 25). *Forschung & Lehre*, 3. <https://www.forschung-und-lehre.de/forschung/moralisches-dilemma-auch-bei-autonomen-autos-1135>
- Müller, V. C. (2020, April 30). *Ethics of Artificial Intelligence and Robotics (Stanford Encyclopedia of Philosophy)*. Stanford Encyclopedia of Philosophy. Retrieved May 7, 2026, from <https://plato.stanford.edu/entries/ethics-ai/>
- Pflichtethik*. (n.d.). HfH Ethik-Plattform. <https://ethik.hfh.ch/begriffe/pflichtethik/>
- SAE Levels of Driving Automation™ Refined for Clarity and International Audience*. (2021, May 2). SAE International. Retrieved June 5, 2026, from <https://www.sae.org/news/blog/sae-levels-driving-automation-clarity-refinements>

Sinnott, W. (2003, May 20). *Consequentialism* (*Stanford Encyclopedia of Philosophy*).

Stanford Encyclopedia of Philosophy. Retrieved June 5, 2026, from

<https://plato.stanford.edu/entries/consequentialism/>

Toivonen, H., & Kramer, S. (2025). *Was ist künstliche Intelligenz? 100 Fragen und*

Antworten. TEOS VERLAG. Retrieved May 7, 2026, from

https://link.springer.com/chapter/10.1007/978-3-658-49787-3_13

Wang, C., Liu, Q., Fang, W., & Xiong, C. (2026, Mai). HSPG: An open-loop testing

framework for autonomous driving based on proactive generation of hazardous scenario. *Science Direct*.

<https://www.sciencedirect.com/science/article/pii/S0001457526000588?>

Was ist Deep Learning? (n.d.). IBM. Retrieved June 5, 2026, from

<https://www.ibm.com/de-de/think/topics/deep-learning?>