



UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Autonomes Fahren: Ein Beitrag zur Debatte aus der
Perspektive der christlichen Verantwortungsethik

Autonomous Driving: A Contribution to the Debate from
the Perspective of Christian Ethics of Responsibility

by

Dipl.-Ing. (FH) HILGER STEENBLOCK

UP-Student-Number: 22940172

Supervisor: Prof. Volker Kessler

Co-Supervisor: Prof. Daniël P. Veldsman

30. April 2025

Submitted in fulfillment of the requirements
for the degree MTh in Dogmatics and Christian Ethics in the
FACULTY OF THEOLOGY AND RELIGION
Department of Systematic and Historical Theology
at the
UNIVERSITY OF PRETORIA

Ehics statement

The author, whose name appears on the title page of this dissertation, has obtained the required research ethics approval/exemption for the research described in this work.

The author declares that he has observed the Code of ethics for scholarly activities.

Bad Zwischenahn, 30.04.2025

Ort, Datum



Unterschrift

I, Hilger Steenblock, declare that the dissertation is my own work and has not previously been submitted by me for a degree at this or any other tertiary institute.

Bad Zwischenahn, 30.04.2025.

Ort, Datum



Unterschrift

Widmung

Für meinen Neffen Malte.

Weiter widme ich diese Arbeit allen Christen in der Automobilindustrie und Politik, die sich um eine Verbesserung der Verkehrssicherheit bemühen.

Danksagung

Ein ganz herzliches Dankeschön möchte ich meinen Supervisoren Prof. Dr. habil. Dr. Volker Kessler und Prof. Dr. Daniël P. Veldsman sagen für die ermutigende Begleitung, Hilfestellung und Herausforderung während meiner Forschungsarbeit. Es war mir eine große Freude und auch eine große Bereicherung in meinem Leben, mit euch zusammengearbeitet zu haben.

Ich bedanke mich auch bei Dr. Sandy Abreu von der Elektrobit Automotive GmbH, einer Kollegin aus meinem beruflichen Umfeld. Sie war dazu bereit, den technischen Teil dieser Arbeit in Kapitel 3 auf technisch-wissenschaftliche Richtigkeit zu überprüfen und mir dabei wertvolle Rückmeldungen zu geben.

Darüber hinaus danke ich den Vielen, die mir Fragen beantworteten, Feedback gaben und Korrektur lasen. Mein besonderer Dank geht diesbezüglich an meine Frau Eita Steenblock sowie an die Lektorin Claudia Böckle für die äußerst professionelle Überprüfung meiner Forschungsarbeit. Eine weiterer Dank für eine besondere Begleitung gilt Dr. Wolfgang Egelhardt, der mir durch seine Forschungsarbeit mit dem Titel „Eine theologische Perspektive auf die ethische Debatte um den Einsatz autonomer Waffensysteme“ ein hilfreicher Gesprächspartner war.

Ein herzlicher Dank geht auch an Marian Winter und Dorothea Krämer von GBFE (Gesellschaft für Bildung und Forschung in Europa) für ihre überaus hilfreiche administrative Begleitung. Ferner bedanke ich mich auch bei Don Winter für die liebevolle Unterstützung während der Studienreise im Jahr 2023.

Auch möchte ich mich bei den Verantwortlichen von AcF (Akademie für christliche Führungskräfte) und hier insbesondere erneut bei Prof. Dr. habil. Dr. Volker Kessler bedanken. Erst auf dem Weg zum AcF-Diplom „Christian Leadership“ war es mir möglich, den Einstieg in die Theologie als Geisteswissenschaft zu finden.

Ethik-Erklärung

Der Autor, dessen Name auf der Titelseite dieser Arbeit erscheint, hat die erforderliche Forschungsethikgenehmigung/Befreiung für die in dieser Arbeit beschriebene Forschung erhalten. Der Autor erklärt, dass er den ethischen Verhaltenskodex für wissenschaftliche Tätigkeiten eingehalten hat und dass die Arbeit seine eigene Leistung ist und nicht zuvor für einen Abschluss an der Universität Pretoria oder einer anderen Institution eingereicht wurde.

Schlüsselwörter

Autonomes Fahren, selbstfahrende Fahrzeuge, theologische Ethik, künstliche Intelligenz, moralische Verantwortung, christliche Verantwortungsethik.

Zusammenfassung

Dieses Forschungsprojekt stellt theologische und ethische Fragen zum Einsatz von autonomen Fahrzeugen. Zur Beantwortung der Forschungsfrage wird eine Fallstudie durchgeführt. Diese basiert auf zwei dokumentierten Unfällen mit selbstfahrenden Fahrzeugen, die für den Personentransport freigegeben wurden. Die Auswahl der Unfälle berücksichtigt die Varianz in der technischen Umsetzung der selbstfahrenden Fahrzeuge, um dadurch ein besseres Forschungsergebnis zu erzielen. Die christliche Verantwortungsethik ist für diese Arbeit besonders relevant. Diese Perspektive auf den Begriff des Handelns wurde maßgeblich von Dietrich Bonhoeffer und Reinhold Niebuhr beeinflusst.

Keywords

Autonomous vehicles, self-driving vehicles, theological ethics, artificial intelligence, moral responsibility, Christian ethics of responsibility.

Summary

This research project analyses theological and ethical questions surrounding the use of autonomous vehicles. A case study is being conducted to answer the research question. It is based on two documented accidents involving self-driving vehicles that were approved for passenger transport. The accidents were selected to reflect the variance in the technical implementation of self-driving vehicles in order to achieve better research results. Christian ethics of responsibility are particularly relevant to this work. This perspective on the concept of action was significantly influenced by Dietrich Bonhoeffer and Reinhold Niebuhr.

Kurzübersicht

Widmung	3
Danksagung	4
Ethik-Erklärung	5
Schlüsselwörter	6
Zusammenfassung	6
Keywords	7
Summary	7
1 EINLEITUNG	13
1.1 Problemfeld	13
1.2 Definitionen	13
1.3 Forschungsgebiet.....	14
1.4 Standpunkt des Autors	14
1.5 Forschungsziele	15
1.6 Forschungsfragen	15
1.7 Methodik.....	15
1.8 Kapitelgliederung.....	17
2 FORSCHUNGSSTAND	18
3 TECHNISCHE GRUNDLEGUNG	24
3.1 Menschliches Verhalten in der Fahrzeugführung	24
3.2 Kybernetik	29
3.3 Mechatronik.....	34
3.4 Entwicklungsprozesse	37
3.5 Künstliche Intelligenz.....	40
3.6 Zwischenfazit.....	43
4 FALLSTUDIEN UND DEBATTE	45
4.1 Kollision zwischen einem Tesla und einem Sattelzug	45
4.2 Kollision zwischen einem Robotertaxi und einer Fußgängerin	50
4.3 Hauptlinien in der ethischen Debatte.....	53
5 THEOLOGISCHE GRUNDLEGUNG	72
5.1 Bonhoeffers Konzept der Verantwortung	72

5.2	Niebuhrs Konzept der Verantwortung	81
5.3	Brunners Begriff der Verantwortlichkeit	87
5.4	Schlussfolgerung.....	88
6	THEOLOGISCH-ETHISCHE ENTSCHEIDUNGSFINDUNG	89
6.1	Kontextsensibilität und Wirklichkeitsgemäßheit.....	89
6.2	Klärung der moralischen Frage	91
6.3	Argumente der Stakeholder	92
6.4	Ein Blick in die Luftfahrt.....	97
6.5	Evaluation der Argumente.....	99
7	BEANTWORTUNG DER FORSCHUNGSFRAGE.....	103
8	LITERATURVERZEICHNIS.....	108
9	ANHANG	115
9.1	Abkürzungen	115
9.2	Abbildungen	116
9.3	Tabellen	116

Ausführliches Inhaltsverzeichnis

Ehics statement	2
Widmung	3
Danksagung	4
Ethik-Erklärung	5
Schlüsselwörter	6
Zusammenfassung	6
Keywords	7
Summary	7
1 EINLEITUNG	13
1.1 Problemfeld	13
1.2 Definitionen	13
1.3 Forschungsgebiet.....	14
1.4 Standpunkt des Autors	14
1.5 Forschungsziele	15
1.6 Forschungsfragen	15
1.7 Methodik.....	15
1.8 Kapitelgliederung.....	17
2 FORSCHUNGSSTAND	18
3 TECHNISCHE GRUNDLEGUNG	24
3.1 Menschliches Verhalten in der Fahrzeugführung	24
3.1.1 Maschinenführung nach Rasmussen.....	25
3.1.2 Fahrzeugführung nach Donges	27
3.1.3 Zeithorizonte	28
3.2 Kybernetik	29
3.2.1 Abstrakte Sicht auf Systeme.....	30
3.2.2 Aspekte der Regelungstechnik	30
3.2.3 Aspekte der Informationsverarbeitung	32
3.2.4 Aspekte der Spieltheorie.....	33
3.2.5 Aspekte der Algorithmen	33
3.3 Mechatronik.....	34
3.3.1 Rationaler Agent.....	35
3.3.2 Sensorik.....	36

3.4	Entwicklungsprozesse.....	37
3.4.1	V-Modell	37
3.4.2	Kritik am V-Modell.....	39
3.5	Künstliche Intelligenz.....	40
3.5.1	Maschinelles Lernen.....	41
3.5.2	Neuronale Netzwerke	42
3.6	Zwischenfazit.....	43
4	FALLSTUDIEN UND DEBATTE	45
4.1	Kollision zwischen einem Tesla und einem Sattelzug	45
4.2	Kollision zwischen einem Robotertaxi und einer Fußgängerin	50
4.3	Hauptlinien in der ethischen Debatte.....	53
4.3.1	Ethische Theorien und Leitlinien.....	53
4.3.2	Die Debatte um Dilemma-Situationen.....	58
4.3.3	Die Debatte aus der Sicht ausgewählter Stakeholder.....	64
4.3.4	Debatte mit Fokus „Tesla Autopilot“	68
4.3.5	Debatte mit Fokus „Cruise“ Robotertaxi.....	71
5	THEOLOGISCHE GRUNDLEGUNG	72
5.1	Bonhoeffers Konzept der Verantwortung	72
5.1.1	Biografie von Dietrich Bonhoeffer	72
5.1.2	Rezeption des Bonhoefferschen Verantwortungsbegriffs	74
5.1.3	Matrix der Verantwortung	75
5.1.4	Zwischenfazit zu Bonhoeffer.....	81
5.2	Niebuhrs Konzept der Verantwortung	81
5.2.1	Biografie von H. Richard Niebuhr	81
5.2.2	Zentraler verantwortungsethischer Ansatz	82
5.2.3	Verortung der Theorie Niebuhrs	86
5.3	Brunners Begriff der Verantwortlichkeit	87
5.4	Schlussfolgerung.....	88
6	THEOLOGISCH-ETHISCHE ENTSCHEIDUNGSFINDUNG	89
6.1	Kontextsensibilität und Wirklichkeitsgemäßheit.....	89
6.2	Klärung der moralischen Frage	91
6.3	Argumente der Stakeholder	92
6.4	Ein Blick in die Luftfahrt.....	97
6.5	Evaluation der Argumente.....	99

6.5.1	Fall Tesla	99
6.5.2	Fall Cruise.....	101
7	BEANTWORTUNG DER FORSCHUNGSFRAGE.....	103
8	LITERATURVERZEICHNIS.....	108
9	ANHANG	115
9.1	Abkürzungen	115
9.2	Abbildungen	116
9.3	Tabellen	116

1 EINLEITUNG

1.1 Problemfeld

Mehrere Automobilhersteller und Technologieunternehmen entwickeln und testen derzeit autonome Fahrsysteme, und einige Städte und Länder haben damit begonnen, autonome Fahrzeuge in begrenztem Umfang für den öffentlichen Verkehr, Lieferdienste und andere Zwecke freizugeben. Autonomes Fahren bezeichnet die Fähigkeit eines Fahrzeugs, eigenständig im öffentlichen Verkehrsraum zu fahren, ohne direktes menschliches Zutun. Diese Technologie trägt wesentlich zur Verbesserung der Straßenverkehrssicherheit, zur Verringerung der Verkehrsüberlastung und zur Verbesserung der Zugänglichkeit und Mobilität von Menschen mit Behinderungen und Personen mit eingeschränkter Mobilität bei.

Neben den technischen Herausforderungen stellen sich auch neue gesellschaftliche Fragen. Basierend auf einer ersten Literaturanalyse wird das Thema von Seiten der philosophischen und theologischen Ethik diskutiert. Da bisher nur sehr wenige Beiträge aus der theologischen Ethik zu finden sind, soll diese Arbeit einen Beitrag zur Debatte aus theologisch-ethischer Sicht leisten.

1.2 Definitionen

Autonomes Fahren (AD, auch „Self-Driving“, „Roboter-Taxen“ oder „Autopilot“ genannt) bezeichnet die Fähigkeit eines Fahrzeugs, ohne direktes menschliches Zutun oder Eingreifen im öffentlichen Verkehrsraum zu fahren. Autonome Fahrsysteme verwenden eine Kombination aus verschiedenen Technologien wie dem Global Positioning System (GPS), verschiedene Sensoren für das maschinelle Sehen und fortschrittliche Algorithmen, um zu navigieren und Entscheidungen zu treffen.

Künstliche Intelligenz (KI) bezeichnet die Fähigkeit von Maschinen oder Computersystemen, Aufgaben auszuführen, die normalerweise menschliche Intelligenz erfordert. KI-Systeme lassen sich je nach ihren Fähigkeiten und Funktionen in verschiedene Typen einteilen. Drei gängige Arten von Kategorien sind folgende: Als erstes wären „Rule-based systems“ (regelbasierte Systeme) zu nennen. Diese Systeme verwenden vordefinierte Regeln und logische Schlussfolgerungen, um Entscheidungen zu treffen und auf Eingaben zu reagieren. Eine zweite Kategorie ist

die Nutzung von „Machine learning“-Methoden (maschinelles Lernen). Diese Systeme verwenden Algorithmen, um Daten zu analysieren und daraus zu lernen, ohne explizit programmiert zu werden. Das automatische Lernen kann noch weiter unterteilt werden. Beim „überwachten Lernen“ wird die Maschine mit gekennzeichneten Daten trainiert, während beim „unüberwachten Lernen“ Muster in nicht gekennzeichneten Daten erkannt werden. Und als letztes Beispiel in diesem Überblick sei die Kategorie des „Deep Learning“ (tiefes Lernen) genannt. Hierbei handelt es sich um einen Teilbereich des maschinellen Lernens, bei dem künstliche neuronale Netze zur Verarbeitung und Analyse großer Datenmengen eingesetzt werden, so dass die Maschine komplexe Muster erkennen und genauere Vorhersagen machen kann.

1.3 Forschungsgebiet

Die Arbeit ist in der theologischen Ethik angesiedelt. Zugleich ist die Arbeit von der Diskussion in der angewandten bzw. bereichsspezifischen Ethiken, etwa der Technikethik oder der Roboterethik, beeinflusst. Die angewandten Ethiken existieren dabei unabhängig von der theologischen Ethik.

Es gibt verschiedene Konzepte der Ethik wie Tugendethik, Pflichtenethik und Verantwortungsethik. Die Verantwortungsethik ist für diese Arbeit besonders relevant. Diese Perspektive auf den Begriff des Handelns wurde maßgeblich von Max Weber (1919) und Dietrich Bonhoeffer beeinflusst.

1.4 Standpunkt des Autors

In der Automobilindustrie in Deutschland arbeite ich als Projektleiter für Fahrerassistenzsysteme und automatisiertes Fahren. Als Diplom-Ingenieur habe ich eine positive Einstellung gegenüber der Technik im Allgemeinen.

Die Menschheit hat nach Budelacci ihren göttlichen Sinngaranten verloren (Budelacci 2022:18) und versucht nun, sich mittels der Technik zu „erlösen“ oder zumindest zu optimieren. Es existiert seiner Meinung nach ein grenzenloser Optimierungswille und ein zusätzlicher Druck durch sogenanntes Biohacking (:25), um sich reibungslos ins kapitalistische System einzufügen. Damit, so warnt Budelacci, gibt der Mensch auf,

was ihn als Menschen auszeichnet (:19). Für mich als Christ in der freien, evangelischen Kirche (eingebettet im Evangelischen Gnadauer Gemeinschaftsverband) existiert ein anderes Bezugssystem. Dadurch kann es mir gelingen, mich (möglicherweise nur teilweise) dem Optimierungsdruck zu entziehen.

Mich interessiert die Frage, was uns die theologische Ethik in dem neuen Bereich der Anwendung von KI lehren kann. Im Grunde genommen können wir in der Bibel keine direkten Antworten für die Entwicklung neuer Technologien finden. Aber ich beginne meine Arbeit mit der Hypothese, dass die theologische Ethik dem Bereich der Anwendung von KI eine wertvolle Stimme geben kann.

1.5 Forschungsziele

Dieses Forschungsprojekt analysiert theologische und ethische Fragen zum Einsatz von autonomen Fahrzeugen. Bisher wurde die Debatte über dieses Feld aus der Perspektive der allgemeinen Ethik geführt. Mit dieser Arbeit möchte ich einen Beitrag aus einer christlich-theologischen Perspektive leisten. Insbesondere gehe ich auf das Thema der Verantwortung ein, einem zentralen Begriff der christlichen Ethik und der christlichen Anthropologie.

1.6 Forschungsfragen

Zentrale Forschungsfrage

Ist es aus theologischer Sicht zu verantworten, autonome Fahrzeuge fahren zu lassen?

Untergeordnete Forschungsfrage

1. Welche Umstände haben zu einem Unfall mit einem autonomen Fahrzeug geführt?
2. Wie ist die Nutzung von selbstfahrenden Fahrzeugen aus der Perspektive der christlichen Verantwortungsethik zu bewerten?

1.7 Methodik

Für die Methodik bietet sich ein Modell von Bleisch, Huppenbauer & Baumberger (2021) zur ethischen Entscheidungsfindung an (siehe Abbildung 1). Im Schritt drei fließen dabei die Hintergrundtheorien ein. Dabei nennen Bleisch et al. explizit nicht die theologische Ethik als mögliche Hintergrundtheorie. Im Rahmen dieser Arbeit soll die theologische Ethik, bzw. konkreter die Verantwortungsethik nach Weber und Bonhoeffer, in die Methodik mit einbezogen werden.

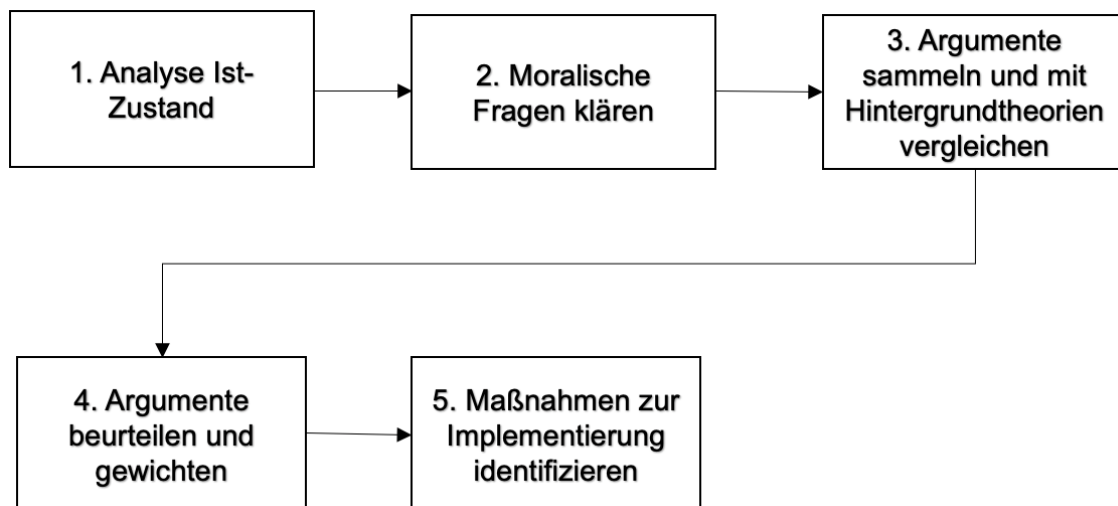


Abbildung 1 Schema ethische Entscheidungsfindung (Bleisch et.al., 2021)

Ferner spielt in dieser Arbeit die Technologie eine wichtige Rolle. Ladd (2019) nennt die verborgene Technologie, die zum Beispiel ein Flugzeug automatisch landet oder ein Kernkraftwerk steuert, „Vermittler“ („Intermediaries“ 2019:2014)¹. In der Methodik wird seinem Rat gefolgt, diese „Vermittler“ in der Analyse der Verantwortlichkeiten einzubeziehen.

Die Forschungsfragen dieser Arbeit sollen mittels zweier Fallstudien (Case Studies) beantwortet werden. Dabei fließen die Empfehlungen von Yin (2018) und Thomas (2021) in das Design der Untersuchung ein. Als Fälle dienen zwei dokumentierte Unfälle mit selbstfahrenden Fahrzeugen. Die Auswahl der Fälle berücksichtigt folgende Grundkonzepte²:

- a) Das Fahrzeug hat einen Platz für die Fahrerin. Unter bestimmten Umständen kann das automatische Fahren aktiviert werden. Die Marke Tesla bietet zum Beispiel eine entsprechende Funktion an. Es handelt sich hier also um eine Erweiterung eines am Markt verfügbaren Fahrzeuges.

¹ Im technischen Bereich wird hier oft von „eingebetteten Systemen“ gesprochen. „Vermittler“ wird aus technischer Sicht eher selten verwendet. Sobald diese eingebetteten Systeme zu einem Verbund mit einer Dateninfrastruktur (z.B. das Internet) zusammengeschlossen werden, kommt der Begriff „Cyberphysisches System“ zur Anwendung.

² Der Methodenansatz „Case Study“ bietet sich bei diesen zwei „Fällen“ besonders gut an: Ein Experiment zum Beispiel wäre für den Rahmen dieser Arbeit zu kostspielig und würde ferner unter Laborbedingungen möglicherweise zu anderen Ergebnissen kommen.

- b) Das Fahrzeug ist vom Grund her fahrerlos entworfen worden und erste Fahrversuche damit wurden unternommen (siehe BBC 2023). Leitbild für die Entwicklung ist die Automatisierung des öffentlichen Personennahverkehr.

Die Fahrzeuge bewegen sich im öffentlichen Straßenverkehr und dabei ist es zu Verkehrsunfällen gekommen. Andere Umgebungen wie das Fahren auf einem Firmengelände oder innerhalb einer Produktionsstätte wären als Fälle auch denkbar, sind jedoch nicht Fokus dieser Studie.

Da zwei Fälle unabhängig voneinander untersucht werden, bietet sich die Möglichkeit an, eine Multiple-Case Study durchzuführen und am Ende nach Yin eine Cross-Case Conclusion (2018:48ff) zu schreiben (siehe Abbildung 2).

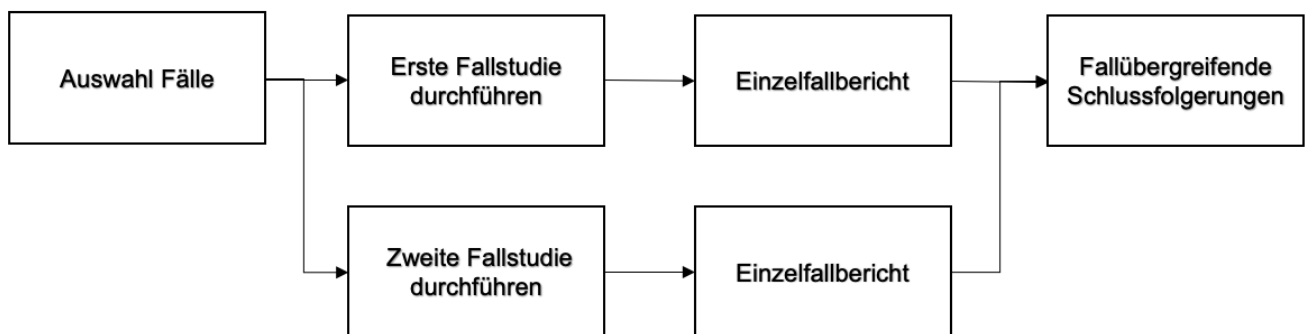


Abbildung 2: Multiple-Case Study nach Yin (2018)

1.8 Kapitelgliederung

Nach dieser Einführung wird im folgenden Kapitel der Stand der Forschung dargestellt. Darauf folgt in einem weiteren Schritt die technische Grundlegung mit ausgewählten Aspekten, die für die Entwicklung von selbstfahrenden Fahrzeugen notwendig sind. Damit ist eine Grundlage geschaffen, um die zwei Fallstudien vorzustellen sowie auf die Debatte dazu einzugehen. Anschließend wird eine theologische Grundlegung vorgenommen und auf die christliche Verantwortungsethik eingegangen. Im Anschluss daran wird eine Evaluation der Argumente durchgeführt. Abgeschlossen wird diese Arbeit mit einer Antwort auf die Forschungsfrage.

2 FORSCHUNGSSTAND

Wir stehen am Anfang der 4. Revolution in der Industrie. Nach der Entdeckung der Wasser- und Dampfenergie, der Elektrizität und der Computertechnologie werden nun die verschiedenen Technologien miteinander verbunden. Und der Übergang hat eine höhere Geschwindigkeit, als wir sie bisher erlebt haben. Henning – ein ehemaliger Professor an der RWTH Aachen – schreibt treffend, dass „die digitalen Begleiter allgegenwärtig und unaufdringlich sein werden. Sie sind äußerst nützlich und deshalb nutzen wir sie“ (Henning 2020:13). Die Frage ist, welche Auswirkungen die 4. industrielle Revolution auf die Gesellschaft haben wird. Es besteht die allgemeine Auffassung, dass dank der technologischen Entwicklung jeder ein gutes Leben haben wird. So werden beispielsweise neue Geräte wie ein „Kirchenfest“ gefeiert, wenn das „beste Smartphone, das wir hergestellt haben“ angekündigt wird. Es bleibt jedoch zu bezweifeln, dass alle besser leben werden. Wie Peckham warnt:

Wir beschäftigen uns mit sozialen Medien, dem Internet, Online-Shopping, intelligenten Städten und den neuesten Gadgets, ohne jemals darüber nachzudenken, was dies mit dem Bild Gottes in uns macht oder wie es unser Verhalten und unsere Beziehungen verändert. (Peckham 2021:27)

Diese 4. Revolution lässt sich auch als digitale Transformation definieren, als einen Schirm, unter dem zahlreiche technologische Änderungen eingeordnet werden können. Kirschschaeger spricht in diesem Aspekt zusätzlich von „super-data-based-systems“ (Kirschschaeger 2021), wie wir sie von den großen Cloud-Anbietern kennen³. Diese Transformation führt zu neuen Ungerechtigkeiten, da aus seiner Sicht nur wenige Menschen einen direkten Vorteil durch den „technology-based shift“ (Kirschschaeger 2021:107) haben. Aus ethischer Sicht müsste es jedoch gelingen, dass die meisten Menschen vom technologischen Fortschritt profitieren können.

Mehrere Automobilhersteller und Technologieunternehmen entwickeln und testen derzeit autonome Fahrsysteme, und einige Städte und Länder haben damit begonnen, autonome Fahrzeuge in begrenztem Umfang für den öffentlichen Verkehr, Lieferdienste und andere Zwecke freizugeben. Autonomes Fahren bezeichnet die Fähigkeit eines Fahrzeugs, eigenständig im öffentlichen Verkehrsraum zu fahren, ohne direktes menschliches Zutun.

³ Ein Super-data-based system könnte das „Infrastructure as a Service“ Konzept (IaaS) sein. Damit werden Rechner-Leistung und Speicher on-demand dem Kunden zur Verfügung gestellt. Amazon, Microsoft und Google gehören hier zu den bekanntesten Marktführern.

Die Kapazität des autonomen Fahrens kann in verschiedene Stufen eingeteilt werden. Nach SAE (2021) wird zwischen fünf verschiedene Stufen der Automatisierung unterschieden: Assistiert wird als Stufe 1 bezeichnet, teilautomatisiert als Stufe 2, hochautomatisiert als Stufe 3, und Stufe 4 wird als vollautomatisiert bezeichnet. Das autonome Fahren, in dem nur Passagiere ohne Fahraufgabe im Fahrzeug sich aufhalten, wird als Stufe 5 bezeichnet. Bei Stufe 3 „fahren Sie nicht, wenn diese automatisierten Fahrfunktionen aktiviert sind – auch wenn Sie auf dem Fahrersitz sitzen“ (SAE 2021). Das hat zur Folge, dass das Fahrzeug während der autonomen Fahrt die Verantwortung über das Handeln übernimmt.

Aus technischer Sicht ist eine solche Automatisierung des Fahrens erst mit der Nutzung großer Datenmengen, mit schnellen Internetverbindungen sowie mit der notwendigen Rechenkapazität denkbar. Darüber hinaus hilft die künstliche Intelligenz, um beispielsweise ein maschinelles Sehen (Stiller et.al. 2015) mit geeigneter Objekterkennung zu realisieren. Für eine Berechnung der Sicherheitsabstände eines Fahrzeugs ist eine 3D-Szenengeometrie unverzichtbar (Stiller et.al., 2015:370). Da jedoch eine Kamera die dreidimensionale reale Welt um eine ganze Dimension reduziert, müssen geeignete Verfahren eingesetzt werden, um diesen Informationsverlust auszugleichen.

In der Literatur werden unterschiedliche Ausgangspunkte für eine Einführung in die künstliche Intelligenz gewählt. Crawford (2021) beginnt mit einem Beispiel vom Ende des neunzehnten Jahrhunderts. Damals konnte das „klügste Pferd der Welt“ (Crawford 2021) Dinge tun, die für Menschen möglich sind, aber nicht für Tiere. Später wurde bei der Analyse der Gründe für die tierische Intelligenz herausgefunden, dass das Pferd „Kluger Hans“ gelernt hatte, die Körperhaltung, die Atmung und andere Dinge des Fragenden zu interpretieren, um die richtige Antwort zu geben. Laut Crawford (:4) wird das Beispiel „jetzt beim maschinellen Lernen als warnende Erinnerung daran verwendet, dass man nicht immer sicher sein kann, was ein Modell aus den ihm gegebenen Daten gelernt hat“. Selbst wenn das System für die aktuellen Daten korrekte Antworten liefert, kann es zu „schrecklichen Vorhersagen“ kommen, wenn neue Eingabedaten verwendet werden.

Technisch gesehen lernt die KI durch die Auswahl eines großen Datensatzes. Wie Crawford hervorgehoben hat, muss eine solche Auswahl der Trainingsdaten sorgfältig getroffen werden, da eine falsche Auswahl zu einer Fehlinterpretation führen und eine Ungleichheit zur Folge haben kann. So haben Forscher beispielsweise herausgefunden, dass eine falsche Vorbedingung, nämlich „Zugang zu einer Lernplattform zu haben“, am Ende zu einem Algorithmus führen kann, der eine Ungerechtigkeit enthält (Dieterle, Holland & Dede 2021).

Andere Autoren (z.B. Bartneck et.al., 2019, Nida-Rümelin, J, Weidenfeld, N., 2020) nutzen Science-Fiction Filme, um über deren Inhalt ethische Betrachtungen zum Zeitalter der künstlichen Intelligenz anzuregen. Zum Beispiel der Film „Ich, der Roboter“ basiert auf dem gleichnamige Buch von Isaac Asimovs (Asimov 2015), das erstmalig im Jahr 1950 erschienen ist. Die Geschichte spielt im Jahr 2035, in dem menschenähnliche (d.h. humanoide) Roboter Realität geworden sind. Die Roboter werden unterwürfig dargestellt, „deren Existenzberechtigung nur darin besteht, von Menschen benutzt und eingesetzt zu werden“ (Nida-Rümelin, J, Weidenfeld, N., 2020:24). In dem Film wird ein für diese Arbeit interessantes Dilemma aufgegriffen: Bei einem Unfall rettet ein Roboter den Detective Del Spooner vor dem Ertrinkungstod, dagegen wurde ein Mädchen nicht gerettet. Das Entscheidungskriterium des Roboters basierte dabei auf einem Abwägen der Überlebenschancen und dem Detective wurden von dem Algorithmus größere Chancen eingeräumt. Ferner haben sich die Asimov’schen Gesetze als fiktionale Moralprinzipien für Roboter überliefert. Das erste Gesetz sei das Schützen von Menschen vor den Robotern. Und so lange dies gegeben ist, sollen die Maschinen ihre eigene Existenz sichern (siehe dazu auch Misselhorn 2019:165).

In dem Film „Ex Machina“ aus dem Jahre 2015 wird eine Androidin Ava geschaffen. In der Handlung soll der junge Programmierer Caleb anhand eines sogenannten Turing-Tests⁴ feststellen, ob die Androidin Ava ein Bewusstsein besitzt. Der Film stellt dem Zuschauer eine interessante Frage: „Hat Ava nur gelernt, bestimmte

⁴ Der Turing-Test geht auf Alan Turing zurück. Er formulierte 1950 die Idee, dass eine Maschine „intelligent und denkend sei, wenn sie Menschen in umfassender Hinsicht täuschen könnte“. D.h. wenn ein Mensch sein künstliches Gegenüber auch als Mensch wahrgenommen hat. Siehe dazu auch Misselhorn 2019:272).

Verhaltens-weisen zu imitieren, um damit den falschen Eindruck zu erwecken, sie hätte Gefühle?“ (Nida-Rümelin, J, Weidenfeld, N., 2020:37).

Nach Budelacci (2022) ist der Begriff der künstlichen Intelligenz (KI) irreführend und lässt einen gewissen Spielraum für Interpretationen. Hilfreicher ist die Unterscheidung zwischen einer „schwachen KI“ und einer „starken KI“ (:62), die erstmals von John R. Searle (1980) verwendet wurde. Maschinen, die mit einer „schwachen KI“ ausgestattet sind, verhalten sich zwar intelligent, sind aber immer noch dumm im Vergleich zur Intelligenz eines Menschen. Eine „starke KI“ hingegen ist der Intelligenz des Menschen ebenbürtig. Budelacci (2022) warnt vor dem Einsatz einer „starken KI“: "Sollte dies gelingen, würde der Mensch seiner Einzigartigkeit beraubt werden" (:62). Das Hauptziel der KI-Forschung sei „eine künstliche Intelligenz, die über umfassende Fähigkeiten verfügt (Artificial General Intelligence, AGI) und nicht nur einzelne Fähigkeiten des Menschen übertrifft (Artificial Narrow Intelligence, ANI)“ (:58). Russell & Norvig definieren die künstliche Intelligenz als einen „intelligenten Agenten“ (2010:61), der abhängig von den Eingaben über weitere Aktionen entscheiden kann. So gesehen kann auch das selbstfahrende Fahrzeug als intelligenter Agent gesehen werden.

In der Ethik gibt es verschiedene Denkschulen. Zum Beispiel die teleologische Ethik oder die deontologische Ethik. Mühling sieht diese nicht als ausschließende Positionen, sondern als komplementäre Perspektiven und spricht von einer „Theorie des guten Handelns“ (Mühling 2012:37). Damit spricht er die Notwendigkeit an, „das vorzugswürdige Handeln unter der Bedingung der möglicherweise bewussten Übernahme von Schuld verantwortungsvoll zu gestalten“ (:37).

Im Rahmen der digitalen Ethik nutzen Grimm, Keber und Zöllner die ethischen Denkschulen, um eine Einführung in die angewandte Ethik zu geben. Zum Beispiel passt die Folgeabschätzung einer digitalen Innovation zur teleologischen Perspektive (Grimm et.al. 2022:14).

Für Pieper (2017) ist Grundvoraussetzung für eine ethische Diskussion der „gute Wille“:

Guter Wille meint hier die grundsätzliche Bereitschaft, sich nicht nur auf Argumente einzulassen, sondern das als gut Erkannte auch tatsächlich zum

Prinzip des eigenen Handelns zu machen und in jeder einzelnen Handlung umzusetzen. (Pieper 2017)

Emil Brunner, der 1937 seine eigene Anthropologie entwickelt hat, definiert die menschliche Verantwortung als ein Kriterium für das Menschsein selbst. Der Begriff der Verantwortungsethik wird mit Weber und Bonhoeffer in Verbindung gebracht. Weber (1919) verwendete den Begriff erstmals in einem Vortrag im Jahr 1919. Bonhoeffer, der zu seiner Zeit im Widerstand gegen den Nationalsozialismus aktiv war, spricht von der „Bereitschaft der Schuldübernahme“ (Bonhoeffer 2006:275). Angesichts der modernen Entwicklungen kann die Legitimität von Verhalten nicht mehr rein teleologisch-ethisch begründet werden, denn nach Lin kann „ohne Berücksichtigung der Fernwirkung der Handlungsfolgen die Legitimität der Handlungsweise nicht mehr begründet werden“ (Lin 2003:10).

Die Arbeitsteilung ist ein Zeichen unserer Zeit. Dies bedeutet, dass auch die Verantwortlichkeiten geteilt sind und es im Falle eines Fehlers oder Unfalls schwierig ist, eine Person zur Verantwortung zu ziehen. In Anlehnung an die Überlegungen von Jonas (1984) ist es für de Villiers daher wichtig, dass (Technologie-) Unternehmen „die notwendigen organisatorischen Vorkehrungen treffen, um die Ausübung der voraussichtlichen Verantwortung in den eigenen Reihen zu gewährleisten“ (de Villiers 2002:20).

In seiner Masterarbeit hat sich Verhaegen (2021) auf die Verantwortung eines autonomen Fahrzeugs konzentriert. Nach Verhaegen „entsteht die Verantwortungslücke, wenn eine autonome Maschine außerhalb der Kontrolle oder Aufsicht eines menschlichen Agenten handelt“ (2021:18). Verhaegen konzentrierte sich in seiner Arbeit auf die Verantwortungsethik für KI. Wie später in dieser Arbeit gezeigt wird, kommen differenzierte Aspekte für autonome Fahrzeuge zum Tragen.

Ein in der Literatur typischer Ansatz zum Aufzeigen des Moraldilemmas im autonomen Fahren ist die Diskussion über das Trolleyproblem (Lin 2015:78). Darunter versteht man das Gedankenexperiment, bei dem ein Weichensteller einer Straßenbahn einen Menschen opfern könnte, um einen Zusammenstoß mit mehreren anderen Menschen zu vermeiden. Das gleiche Beispiel lässt sich mit einem Fahrzeug und Personen unterschiedlicher Altersklassen durchdenken. Die

moralische Frage dahinter ist die, ob man den Wert eines Lebens höher bewerten darf als den eines anderen Lebens. Dies darf nach dem deutschen Grundgesetz nicht sein (:71). Da dieses Dilemma nicht einfach zu lösen ist, braucht es nach Lin die Diskussion über die Ethik in der Entwicklung der autonomen Fahrzeuge.

Ein Skandal, der im Jahr 2015 in der Welt des Automobils bekannt geworden ist, war die Manipulation der Abgasreinigung bei Dieselmotoren (auch Abgasskandal oder Dieselgate genannt), um die Abgasnorm EU 6 zu erfüllen. Insbesondere der Volkswagenkonzern wird mit diesem Skandal in Verbindung gebracht (di Rattalma 2017). Aus technischer Sicht hat die Software im Fahrzeug erkannt, dass es sich aktuell in einem Prüfmodus befindet. Im Alltagsbetrieb dagegen wurde die Abgasreinigung abgeschaltet. Dadurch war es dem Volkswagenkonzern möglich, die Fahrzeuge ohne zusätzlichen Harnstoff-Zugabe im Fahralltag zu betreiben (auch als „Add-Blue“ bekannt). Infolgedessen mussten führende Manager des Konzerns wie Martin Winterkorn und der Audi-Manager, Rupert Stadler, von ihren Führungspositionen zurücktreten. Strafrechtlich konnte Winterkorn bisher keine Verantwortung nachgewiesen werden (Posocco 2017:13). Dagegen hat Stadler den Betrug gestanden, um einer Bestrafung durch das Gericht zu entgehen. Der Skandal hat jedoch das Vertrauen in die „Vorzeige-Industrie“ erschüttert. Ferner wurden dadurch die Umwälzungen in der Branche noch weiter beschleunigt: VW hat als Konsequenz zum Beispiel die Elektrifizierung der Fahrzeugflotte als neue Strategie definiert (:26). In Bezug auf diese Arbeit kann festgestellt werden, dass unbemerkt von den staatlichen Prüfstellen bewusste Fehlfunktionen in der Software implementiert wurden, um sich einen wirtschaftlichen Vorteil zu verschaffen. Und es zeigt auch, wie schwierig es ist, hochrangige Manager zur Verantwortung zu ziehen.

3 TECHNISCHE GRUNDLEGUNG

Die Entwicklung von selbstfahrenden Fahrzeugen kann als eine weitere Revolution in der Mobilität angesehen werden. Dabei greifen eine Vielzahl von technischen Innovationen so ineinander, dass sie zu einer Erweiterung der Lösungsmöglichkeiten geführt haben bzw. in Zukunft noch führen werden. Um die zwei Fallstudien zu den Kollisionen (siehe Kapitel 4.1 und Kapitel 4.2) besser bewerten zu können, wird darum in diesem Kapitel eine technische Grundlegung vorgenommen. Dabei geht es nicht um eine vollständige Darstellung des Themengebietes, sondern um eine stichpunktartige Abgrenzung, so dass eine Unterscheidung der Argumente möglich wird.

3.1 Menschliches Verhalten in der Fahrzeugführung

Die aktive Teilnahme am Straßenverkehr stellt für die Verkehrsteilnehmer und Verkehrsteilnehmerinnen eine komplexe Aufgabe dar, die eine Vielzahl von Fähigkeiten und Fertigkeiten erfordert, „für deren Gelingen der Fahrer bei heutiger Rechtslage und heutigem Stand der Technik voll verantwortlich ist“ (Donges 2015:18). Wenn die Fahrzeugführung durch Assistenzsysteme unterstützt oder gar übernommen werden soll, ist es zunächst hilfreich zu verstehen, wie der Mensch die Fahrzeugführung umsetzt. Dieser Aspekt wird zu Beginn dieses Kapitels erläutert. Darauf aufbauend folgt die abstrakte Systemsicht und die Vorstellung einiger ingenieurwissenschaftlicher Grundlagen, die für den Entwurf selbstfahrender Fahrzeuge relevant sind.

Die Fahrzeugführung durch einen Menschen lässt sich in eine schematische Darstellung unterteilen (siehe Abbildung 3: Leistungsstufen von qualifizierten menschlichen Bedienern (Rasmussen 1983:258)). Dieser Versuch wird im Folgenden dargestellt. Ferner kommt die Diskussion zur zeitlichen Abhängigkeit von Entscheidungen zur Sprache.

Die vorgestellten Modelle von Donges und Rasmussen wurden aufgrund zweier Punkte ausgewählt. Als erstes helfen sie zu verstehen, wie der menschliche Fahrer oder die Fahrerin die Fahrzeugführung durchführt. Ein zweiter wichtiger Aspekt ist, dass sie ein Verständnis für die Aufgaben in einem selbstfahrenden Fahrzeug vorbereiten. Im Untersuchungsbericht zum Unfall mit dem Robotertaxi von Cruise

von Exponent (2023) wird das komplexe System auf einem abstrakten Niveau erläutert. Viele Teile in dem Bericht wurden für die Öffentlichkeit leider geschwärzt. Jedoch lässt sich der Systemaufbau aufgrund der Sub-Kapitel-Überschriften zum Kapitel „Cruise ADS System Description“ (Exponent 2023:131-140) gut erahnen und diese Unterteilung ist weitgehend identisch mit dem nachfolgenden Modell.

3.1.1 Maschinenführung nach Rasmussen

Professor Jens Rasmussen (1926–2018⁵) forschte an der Technischen Universität in Dänemark im Bereich Cognitive Systems Engineering („CSE“). Der Fokus dieses Bereichs liegt an der Schnittstelle zwischen Menschen und Technologie mit Schwerpunkt auf sicherheitskritische Systeme. Auch in einem Fahrzeug sind Elemente von sicherheitskritischer Relevanz enthalten, wie z.B. die Bremsen oder die elektrische Lenkung.

Zur Erörterung der menschlichen Leistungsfähigkeit in Bezug auf das Bedienen von Maschinen hat Rasmussen (1983:258) ein abstrahierendes Modell entwickelt. Dieses Modell ist auch eine Hilfe zum Verständnis der Fahrzeugführung, unabhängig davon, ob es sich um ein Modell mit Automatisierung nach höheren SAE-Levels oder um ein Fahrzeug ohne Fahrerassistenzsysteme handelt.

Rasmussen schlägt eine Unterscheidung in drei Kategorien vor: Als erste Kategorie definiert er das „Knowledge-Based Behaviour“ (wissensbasiertes Verhalten), als zweite Kategorie das „Rule-Based Behaviour“ (regelbasiertes Verhalten) und als dritte und letzte Kategorie das „Skill-Based Behaviour“ (fähigkeitsbasiertes Verhalten) (:258). Die schematische Darstellung in Abbildung 1 veranschaulicht seinen Ansatz, der die Leistungsebenen eines qualifizierten Maschinenbedieners zeigt.

Die Ebene des „Skill-Based-Behaviour“ nutzt die reflexartige Reaktion des Menschen. Die Schnelligkeit in dieser Ebene beruht nach Rasmussen auf einem „feed-forward“-Konzept (:259), d.h. Störungen werden berücksichtigt, bevor sie sich auf das System auswirken können. Gute Beispiele hierfür finden sich z.B. im Sport, wenn ein Torwart reflexartig die Hand ausstreckt, um einen Ball zu fangen.

⁵ [https://en.wikipedia.org/wiki/Jens_Rasmussen_\(human_factors_expert\)](https://en.wikipedia.org/wiki/Jens_Rasmussen_(human_factors_expert)), angesehen am 04.04.2024.

The total performance is smooth and integrated, and sense input is not selected or observed: the senses are only directed towards the aspects of the environment needed subconsciously to update and orient the internal map. The man looks rather than sees. (Rasmussen 1983:259)

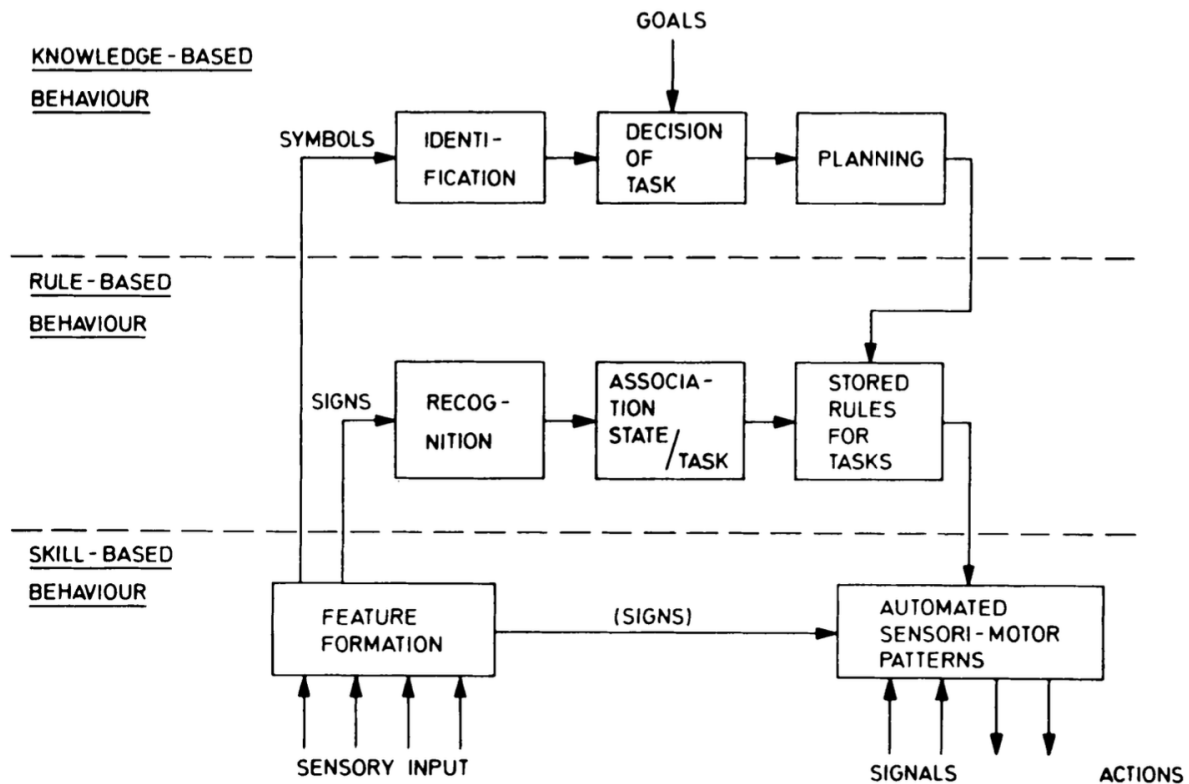


Abbildung 3: Leistungsstufen von qualifizierten menschlichen Bedienern (Rasmussen 1983:258)

Die nächste Ebene „Rule-based Behavior“ wird von gespeicherten Regeln beeinflusst, „which may have been derived empirically during previous occasions, communicated from other persons' know-how as instruction or a cookbook recipe“ (:259).

Das Ziel kann möglicherweise erst nach einer Abfolge von Handlungen erreicht werden. Darum ist die Betrachtung des Systems durch eine „direkte Rückkopplung“ nicht immer sofort möglich. Diese Analyse der Reaktion der Umwelt kann laut Rasmussen als eine unabhängige, gleichzeitige Aktivität der nächsthöheren Ebene („Knowledge-based“ Behaviour“) betrachtet werden (:259). Auf dieser Ebene ist die Entwicklung eines Plans das Hauptmerkmal. Auch das Abwägen von Alternativen oder der Versuch, die Auswirkungen vorherzusagen, sind hier zu finden.

3.1.2 Fahrzeugführung nach Donges

Die Arbeiten von Rasmussen wurden von Donges in dem Artikel „Fahrerverhaltensmodelle“ (Donges 2015) für das *Handbuch Fahrerassistenzsysteme* erweitert. Edmund Donges ist promovierter Ingenieur und hat in seiner Dissertation das Lenkverhalten des Fahrers in einem regelungstechnischen Modell abgebildet (Donges 1977). Beruflich war Donges bis zu seiner Pensionierung als Ingenieur beim Automobilhersteller BMW AG in München tätig.

Die oberste Ebene der Fahraufgabe ist die Navigationsaufgabe. Der Fahrer muss sich Gedanken über das Fahrtziel und die zu wählende Route machen. Während der Fahrt können Störungen im Straßennetz wie Unfälle oder Staus auftreten, die eine Änderung der Route erforderlich machen.

Befindet sich der Fahrer oder die Fahrerin in einem unbekanntem Gebiet, erfordert die Umplanung bei einer Verkehrsstörung einen bewussten Prozess und wird daher von Donges der Ebene des wissensbasierten Verhaltens zugeordnet (Donges 2015:19). Erfolgt die Umplanung in einem bekanntem Gebiet, ist sie wesentlich einfacher und wird daher dem regel-basierten Verhalten zugeordnet.

Die eigentliche Fahraufgabe der Fahrzeugführung wird in den Ebenen „Führen“ und „Stabilisieren“ abgebildet. Auf der Ebene „Führen“ muss der Fahrer bzw. die Fahrerin eine „ständig wechselnde Konstellation sensorischer, insbesondere optischer Eingangsinformationen“ (:19) berücksichtigen. Technisch gesprochen besteht die Fahraufgabe darin, aus den Eingangsinformationen eine sinnvolle Sollspur und Sollgeschwindigkeit abzuleiten, um eine möglichst geringe Abweichung zwischen Ist- und Sollzustand zu erreichen.

Auf der Ebene der „Stabilisierung“ muss der Fahrer oder die Fahrerin bei Bedarf „korrigierende Stelleingriffe“ (:19) vornehmen, um das Fahrzeug in eine sinnvolle Fahrtrichtung zu bewegen. Als Beispiel kann ein Überholmanöver eines LKW auf der Autobahn dienen, das aufgrund von starkem Seitenwind zu einer plötzlichen Richtungsänderung beim Ein- oder Ausscheren aus dem Windschatten führen kann. Die beiden letzten Teilaufgaben „Führen“ und „Stabilisieren“ werden entscheidend durch den Erfahrungsschatz des Fahrers bzw. der Fahrerin beeinflusst. Es ist leicht

nachvollziehbar, dass sich ein Fahranfänger zu Beginn seiner Fahrkarriere zunächst auf das theoretische Wissen aus der Fahrausbildung verlassen muss und daher das Fahrzeug vor allem über die Ebene des „wissensbasierten Verhaltens“ steuert. Erst mit zunehmender Fahrpraxis wird ein Fundus an Verhaltensregeln aufgebaut und gespeichert, der in bestimmten Situationen abgerufen werden kann. Um die Dauer des Lernvorganges der Fahrzeugführung zu verdeutlichen, macht Donges die Unfallbeteiligung von Fahranfängern deutlich. Aus seiner Sicht vergehen „7 Jahre bzw. 100.000 km Fahrleistung, bis ein Fahrer den ausgelernten Zustand erreicht hat“ (:20).

3.1.3 Zeithorizonte

Für die oben dargestellten Ebenen der Fahraufgabe „Navigation“, „Führung“ und „Stabilisierung“ lassen sich bestimmte Zeithorizonte für den Menschen ableiten, innerhalb derer er auf eine potenziell kritische Situation angemessen reagieren kann. Im Bereich der Navigation kann der Zeithorizont für bestimmte Routenwahlen im Stundenbereich liegen, bei einer unerwarteten Streckensperrung (um ein Beispiel zu nennen) zumindest im Minutenbereich.

In der Führungsebene befindet sich der Zeithorizont im Sekundenbereich. Laut Donges sollte die Antizipationszeit (d.h. das Vorhersehen oder das Vorausschauen von Ereignissen), die eine „kognitive Verarbeitung“ erfordert, mindestens einen Wert von zwei bis drei Sekunden haben (:23). Die Wichtigkeit der rechtzeitigen Fahrerreaktion macht Donges mit folgender Aussage deutlich: „Die Hälfte aller Kollisionsunfälle könnte durch Vorverlegung der Fahrerreaktion um eine halbe Sekunde vermieden werden“ (:23).

Das erweiterte Modell der Fahraufgabe ist in Abbildung 4 ersichtlich. Nach Donges gehört das Führen eines Fahrzeuges im wahrsten Sinne des Wortes „zu den zielgerichteten sensumotorischen Tätigkeiten des Menschen“ (:18). Die erste Ebene ist durch „reflexartige Reiz-Reaktions-Mechanismen charakterisiert“ (:18). Diese Ebene wird durch einen kürzeren oder zeitlich längeren Lernprozess geformt, bis in dieser Ebene und außerhalb der bewussten Kontrolle Entscheidungen ablaufen können. Für Donges sind derart eintrainierte Fähigkeiten die zeitlich effektivste Form menschlichen Verhaltens (:18). Zum Beispiel kann das Schalten bei einem Wagen mit Schaltgetriebe bei längerer Erfahrung wie „von selbst“ bzw. als Routineaufgabe

vom Gehirn erledigt werden. Durch diese Routine-Verarbeitung spart das Gehirn Energie für Informationen, die eine bewusste Entscheidung notwendig machen.

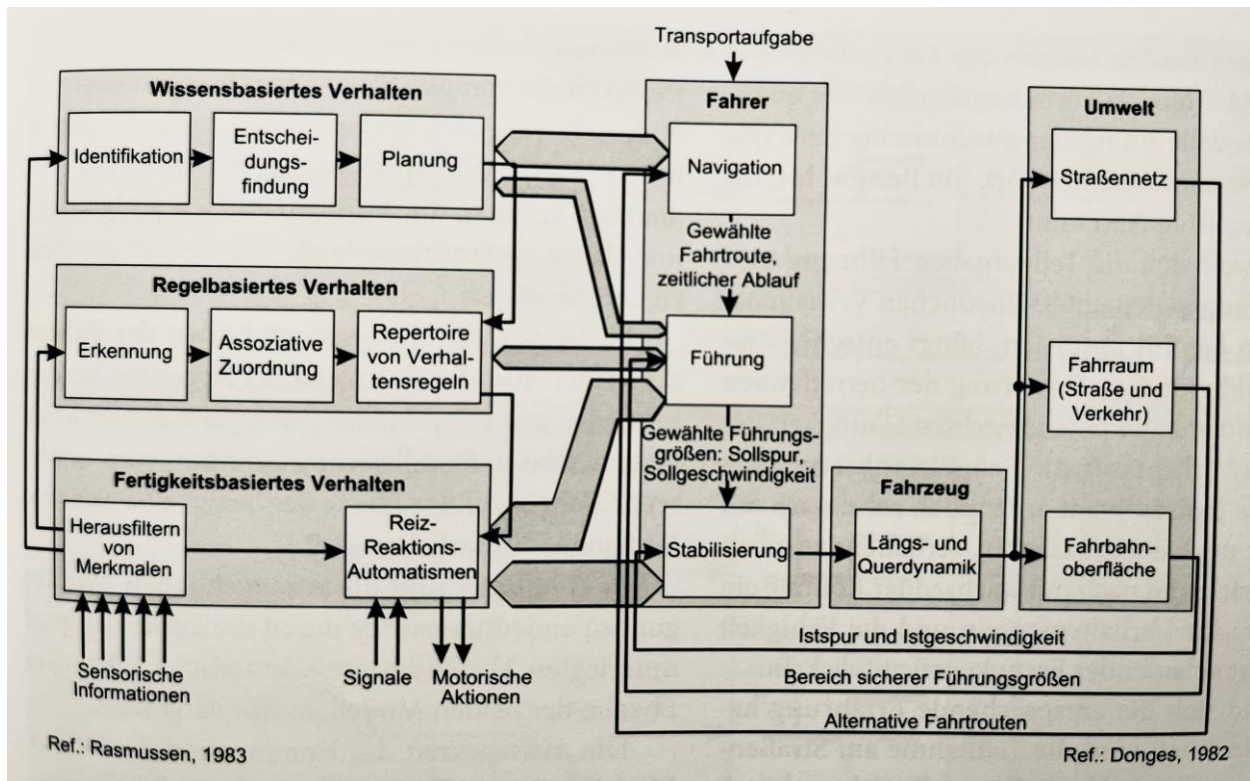


Abbildung 4: Erweitertes Modell der Fahraufgabe (Donges 2015:19)

Kommt der Mensch in eine unvorbereitete Situation und verlangt diese „untrainierte Handlungsweisen“ (:18), führt dies den Menschen zu der Ebene des „wissensbasierten Verhaltens“. In dieser Ebene werden auf Basis von vorhandenem und neu erworbenem Wissen in einem mentalen Prozess Handlungsalternativen durchgespielt. Gute Lösungen kann der Mensch dann in sein Repertoire an bereits getätigten Entscheidungen ablegen und diese dann bei zukünftigen Fällen abrufen.

3.2 Kybernetik

Eine weitere Wissenschaft, die für diese Arbeit relevant ist, ist die Kybernetik (engl. Cybernetics). Wie im Folgenden gezeigt wird (siehe Kapitel 3.2.1), ist die Sichtweise der Kybernetik für die Fahrzeugentwicklung und damit auch für die Entwicklung selbstfahrender Fahrzeuge nützlich.

Als einer der Urväter dieser Wissenschaft gilt Nobert Wiener, der 1963 das Buch *Kybernetik, Regelung und Nachrichtenübertragung im Lebewesen und in der Maschine* in deutscher Übersetzung herausbrachte. Klaus und Liebscher brachten 1974

das Buch *Was ist – Was soll Kybernetik?* heraus. In ihrer Arbeit haben sie fünf wesentliche Merkmale definiert, die aus einem abstrakten System ein „kybernetisches System“ (Klaus & Liebscher 1974:26) machen: Es sind die abstrakte Sicht auf das System, Aspekte der Regelungstechnik, der Informationsverarbeitung, Themen der Algorithmen-Entwicklung sowie die Spieltheorie. Diese fünf Merkmale werden im Folgenden näher erläutert.

3.2.1 Abstrakte Sicht auf Systeme

Ein selbstfahrendes Fahrzeug kann aus verschiedenen Blickwinkeln betrachtet werden. Ein Fahrer oder eine Fahrerin kann sich fragen, wie viel es kostet, welche Funktionen es hat und vielleicht auch, wie gut die selbstfahrenden Eigenschaften genutzt werden können, und natürlich, wie hoch die Kosten für die Anschaffung oder den Betrieb sind. Ein Techniker oder eine Technikerin aus dem Automobilbereich stellt sich dagegen eher Fragen zum Steuergerät, zur maschinellen Umgebungswahrnehmung oder wie der Einsatz von künstlicher Intelligenz realisiert wurde (und vieles mehr).

Als ein erstes Prinzip legt der Kybernetiker nach Küppers seinen Fokus auf die „Systemaspekte“ (2023:27). Der Kybernetiker betrachtet das Auto „as an abstract dynamic system, as a moving system, as a function-oriented system“ (:27). Dieses System kann bei Bedarf weiter aufgebrochen werden in kleinere Teilsysteme. Dieses Aufteilen bietet sich so lange an, bis eine weitere Verfeinerung keinen Sinn mehr ergibt. Solche kleinen Teilsysteme nennen Klaus und Liebscher dann „Elemente“, „deren Zusammenwirken eben die Funktion des betrachteten dynamischen Systems zustandebringt“ (Klaus & Liebscher 1974:13). Auf der anderen Seite steht das gesamte System in einer Wechselbeziehung zu seiner „Umgebung“ (:13): System und die Umgebung bzw. Umwelt bilden eine „neue, höhere Einheit“ (:14).

3.2.2 Aspekte der Regelungstechnik

Eine weitere wichtige Wissenschaft im Fahrzeugbau ist die Regelungstechnik. Innovationen im Bereich der Assistenzsysteme, wie z.B. das seit einiger Zeit bekannte Antiblockiersystem (ABS), wären ohne die Regelungstechnik kaum denkbar. Ein einfaches Beispiel für den Einsatz der Regelungstechnik ist das Zusammenspiel zwischen dem Gaspedal und der daraus resultierenden Beschleunigung des Fahrzeugs.

Als ein Standardwerk für das Wissenschaftsgebiet „Regelungstechnik“ gilt in Deutschland das gleichnamige Werk von Otto Föllinger (Föllinger, Konigorski, Lohmann & et al. 2022). Das Lehrbuch ist erstmals 1972 erschienen und wird von seinen vier ehemaligen Mitarbeitern fortgeführt, die ebenfalls einen Hochschullehrauftrag haben. Das Buch führt in die Verwendung mathematischer Methoden sowie in die Verfahren der Regelungstechnik ein. Nachfolgend wird auf die Terminologie der Regelungstechnik eingegangen. Für tiefergehende Studien sei an dieser Stelle auf das o.g. Buch verwiesen.

Ist ein dynamisches System in Bewegung, so können äußere Einflüsse die Richtung des Systems beeinflussen. Regelungstechniker nennen diese äußeren Einflüsse auch Störgrößen. Ein gutes Beispiel für den Einfluss von Störgrößen ist die Einfahrt oder Ausfahrt in einen Windschatten von einem LKW. Zeigt ein System eine gewisse „Beständigkeit“ oder eine „zweckgerichtete Reaktion“ (Klaus & Liebscher 1974:15) gegenüber äußeren Einflüssen, dann nennt der Kybernetiker dieses Verhalten „geregelt“. Weitere Bezeichnungen nach Klaus und Liebscher sind „geregelt Systeme“ oder auch „selbstregulierende Systeme“ (:15). Das wesentliche Merkmal einer Regelung ist die „Rückkopplung“. Damit ist die Beobachtung des zu steuernden Teilsystems gemeint, wie z.B. die Fahrtrichtung des Fahrzeuges. Der „Regler“ wird versuchen, durch geeignete Entscheidungen und Methoden die „Regelabweichung“ auszugleichen, d.h. die Differenz zwischen Istwert und Sollwert sollte im Idealfall Null sein. Um bei dem o.g. Beispiel des Windschattens zu bleiben, wird der erfahrene Fahrer oder die erfahrene Fahrerin durch einen angepassten Lenkwinkel versuchen, die Fahrtrichtung auch bei Seitenwind beizubehalten. Möglicherweise erfolgt diese Anpassung so geschickt, dass sie von einem anderen Insassen nicht bemerkt wird. Bei einem Fahranfänger oder einer Fahranfängerin ist es hingegen denkbar, dass er oder sie eine Überreaktion zeigt und es zu einem „Übersteuern“ kommt. Durch die Überreaktion in Kombination mit anderen Umständen könnte es zu einer gefährlichen Verkehrssituation kommen. Wird die Regelgröße wie z.B. die Zielrichtung des Fahrzeuges nicht erreicht, so wird dies als „Regelabweichung“ bezeichnet. In der Praxis hat sich die Darstellung mittels Blockschaltbildern etabliert. Abbildung 5 zeigt die Anwendung dieser Blockschaltbilder und stellt ferner die Unterschiede zwischen Regelung und Steuerung dar.

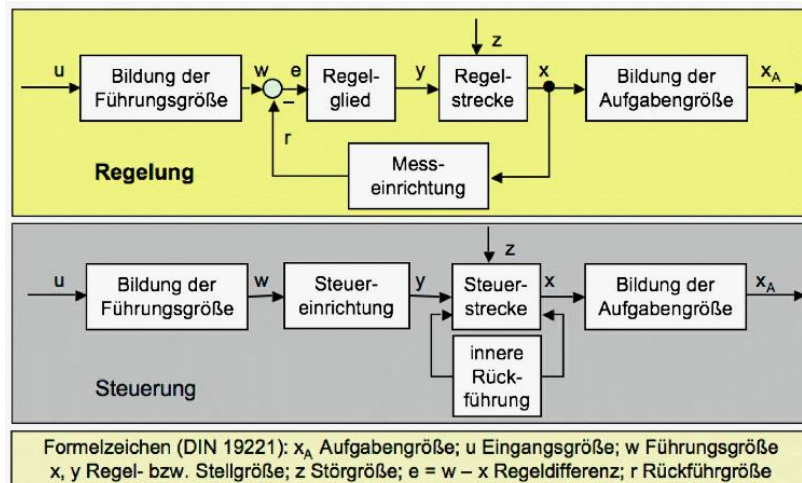


Abbildung 5: Blockschaltbilder einer Regelung und einer Steuerung (Czichos 2019:110)

Als Zusammenfassung der obigen Darstellung lässt sich feststellen, dass das „Prinzip der Rückkopplung“ (principle of feedback) allen kybernetischen Steuerungssystemen innewohnt.

3.2.3 Aspekte der Informationsverarbeitung

Die Fahrzeugsteuerung lässt sich abstrakt durch einen Prozess von Erzeugung, Speicherung und Übermittlung von Informationen verstehen. Das Fahrzeug zeigt dem Fahrer oder der Fahrerin wichtige Informationen an, die für die Fahrzeugführung relevant sind. In einem mentalen Prozess muss der Mensch diese Daten aufnehmen, bewerten und entscheiden. Am Ende führen Signale zu einer Muskelbewegung, die wiederum zu einer Änderung der Lenkeinstellung führen. Schematisch ist dieses Vorgehen in Abbildung 6 ersichtlich. Die Entscheidung fällt der Mensch in seinem Gehirn, in dem „biological subsystem, the driver’s neural network“ (Küppers 2023:29).

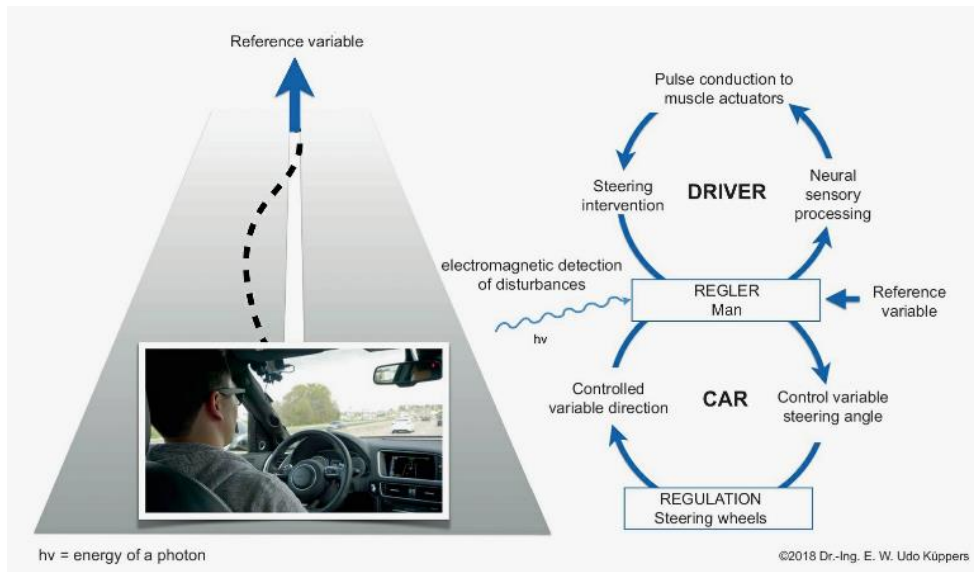


Abbildung 6: Kybernetisches System mit doppelter Rückkopplung (Küppers 2023: 27)

Und damit kann eine dritte Eigenschaft eines kybernetischen Systems formuliert werden: „Kybernetische Systeme sind auch Systeme, die Informationen empfangen, speichern, verarbeiten und in eine Wirkung auf die Umwelt umsetzen“ (:28).

3.2.4 Aspekte der Spieltheorie

Ein viertes Prinzip soll anhand von der Erfahrung der Fahrzeugführer erläutert werden. Weiter oben wurde das Fahrerverhaltensmodell nach Donges vorgestellt. Eine erfahrene Fahrerin oder ein erfahrener Fahrer haben sich durch die Fahrpraxis ein Repertoire an Lösungen abgespeichert, mit denen sie gefährliche Situationen gemeistert haben und diese können sie dann unbewusst abrufen. Dies können extreme Situationen sein wie Glatteis oder starker Regen. Nach Küppers hat der erfahrene Fahrer, die erfahrene Fahrerin eine „optimization strategy at hand“ (:30), die sich durch das „Spielen“ mit verschiedenen vorherigen Verkehrssituationen entwickelt hat. Dieses Spielen bzw. Optimieren von alten Situationen führt Küppers zu dem vierten Prinzip: “Finding optimal strategies and processes is a game-theoretical feature of cybernetic systems” (:30).

3.2.5 Aspekte der Algorithmen

Und zuletzt noch ein fünftes Prinzip, mit dem sich kybernetische Systeme beschreiben lassen: Algorithmen sind ein weiteres Charakteristikum von kybernetischen Systemen. Passend zu den bisherigen Ausführungen folgt das Fahrverhalten im Straßenverkehr bestimmten logischen Schritten wie Lenken, Bremsen und Weiterfahren. Diese lassen sich nach Küppers als Algorithmen in verschiedener

Ausprägung beschreiben. Auch die Entwicklung der künstlichen Intelligenz gehört nach Küppers zu diesen Ausprägungen.

Algorithms appear in many variations and specific applications, such as control algorithms, game algorithms, optimization algorithms, and last but not least, algorithms in the context of artificial intelligence development. (:31)

Für die Entwicklung von regelungstechnischen Anwendungen hat sich eine rechnergestützte Vorgehensweise (engl. computer-aided engineering oder CAE) in der Industrie etabliert. Als ein Beispiel seien hier die Produkte von The MathWorks genannt⁶. Das US-amerikanische Unternehmen hatten sich am Anfang der Firmengeschichte auf Lösungen für numerische Berechnungen spezialisiert. Regelungstechniker erkannten die Vorteile dieser Lösung und so hat sich das Unternehmen mit seinen Produkten schnell in der Branche etabliert. Für die Simulation dynamischer Systeme hat das gleiche Unternehmen das Produkt „Simulink“ entwickelt, das sich ebenfalls als Standardwerkzeug in der Automobilindustrie etablieren konnte.

3.3 Mechatronik

Die Mechatronik ist ein weiteres interdisziplinäres Gebiet, in dem klassische Ingenieur-Disziplinen wie Maschinenbau und Elektrotechnik eingebettet sind. Ziel des Ansatzes ist es, eine ganzheitliche Sichtweise einzunehmen, „to improve the functionality of technical systems and the creation of new concepts of machinery and equipment with built-in „artificial intelligence“ (Pannaga, Ganesh, & Gupta 2013:129). Dabei ist der Begriff „Mechatronics“ ein Kunstwort aus den Wörtern Mechanics und Electronics und wurde erstmalig von der japanischen Firma Yaskawa im Jahre 1969 verwendet. Nach Pannaga, Ganesh und Gupta erhielt die Firma Namensschutzrechte auf die Bezeichnung und gab die Verwendung aufgrund der weltweiten Akzeptanz im Jahre 1982 wieder frei (:129). Das Institute of Electrical and Electronics Engineers (IEEE) nutzt für Mechatronics folgende Definition: *“Mechatronics is the synergistic integration of mechanical engineering with the electronic control and the intelligent, PC-based control, in the design and manufacturing of goods and processes”* (:129).

⁶ siehe <https://ch.mathworks.com/help/overview/control-systems.html>.

Auch im Fahrzeugbau hat die Mechatronik ihre Wirkung entfaltet. Seit jeher kamen die klassischen Ingenieur-Disziplinen wie Maschinenbau und Elektrotechnik zum Einsatz. Durch den Trend zur Digitalisierung zog in das Gebiet zusätzlich die Disziplin „Informatik“ ein und ist heutzutage ein wichtiges Mittel, um eine Differenzierung zwischen den Herstellern zu erzielen. Die kleinen Mini-Computer in Fahrzeugen oder Haushaltsgeräten werden üblicherweise als eingebettete Systeme (embedded systems) bezeichnet. In der Automobilbranche hat sich zudem für eingebettete Systeme die Bezeichnung „Steuergerät“ etabliert (oder engl. Electronic Control Unit mit der Abkürzung ECU).

Eine Übersicht über das Prinzip mechatronischer Systeme zeigt die Abbildung 7. In dem Bild sind zwei Themen gegenübergestellt: Oben sieht man die mechanische Grundstruktur und darunter das Zusammenwirken der Sensorik, der Prozessorik und der Aktorik. Zudem sind auch einige übliche Formelzeichen in der Grafik ersichtlich.

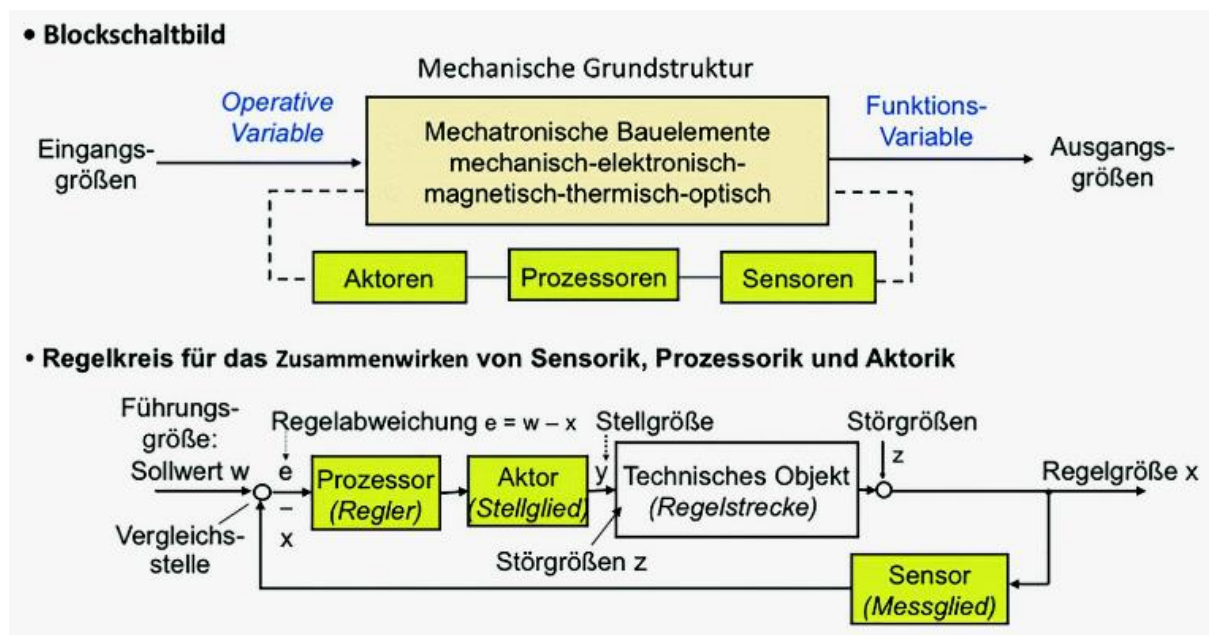


Abbildung 7: Grundprinzip mechatronischer Systeme (Czichos 2019:23)

3.3.1 Rationaler Agent

Nach Russel und Norwig kommuniziert ein Agent über Sensoren und Aktuatoren mit ihrer Umgebung (Russel & Norwig 2012:61). Bei einem Menschen dienen als Sensoren zum Beispiel die Augen, die Ohren und andere Organe. Die Hände, die Füße oder der Mund lassen sich als Aktuatoren verstehen. Durch die „Wahrneh-

mung“ (:61) aus der Umgebung lassen sich die Eingaben in den Agenten beschreiben. Wie schon Donges (2015) veranschaulichte, plant bzw. vollzieht der Mensch in einem mentalen Prozess die nächsten Schritte seines Handelns. Auf dieser Grundlage lassen sich die grundsätzlichen Operationen eines Roboters ebenfalls unterteilen in ein Fühlen oder Wahrnehmen der Umgebung, als Zweites in einen Planvorgang und als Letztes in einen Prozess des Handelns.

3.3.2 Sensorik

Es existieren unterschiedliche Technologien für die Wahrnehmung der Umgebung. Je genauer und verlässlicher ein Fahrzeug die Umgebung erkennen soll, desto wichtiger werden die Sensoren. Wie in Abbildung 7: Grundprinzip mechatronischer Systeme (Czichos 2019:23) ersichtlich, können Sensoren nur in der Kombination mit Aktoren ihre Wirkung entfalten. Für die Beantwortung der Forschungsfrage sind die Sensoren in dieser Arbeit relevanter und darum werden die Aktoren nicht weiter vorgestellt.

Die Ingenieure müssen die verschiedenen Sensortechnologien auswählen und die Rohdaten via angemessener Datenverarbeitung zu einem künstlichen Abbild der Realität zusammenfassen. Doch welche Sensorarten lassen sich nennen? Allgemein bekannt sind Ultraschallsensoren, die für die nahe Distanz eingesetzt werden, wie zum Beispiel bei den seit einiger Zeit verfügbaren Parkassistenten. Als weitere Technologie lassen sich die Kamerasysteme nennen, die auch für die Spurerkennung eingesetzt werden. Eine zentrale Rolle für die Einführung der selbstfahrenden Fahrzeuge ist die Verwendung der LiDAR-Technologie (engl. für Light Detection and Ranging). Durch den Laserstrahl wird eine hochauflösende 3D-Information der Umgebung in Echtzeit erzeugt. Auch in anderen Bereichen des Lebens wird für die Distanzmessung die Laser-Technologie eingesetzt. Was auf der einen Seite charmant und als Lösung für das gesamte Problem der Wahrnehmung zu sein scheint, hat leider auch seine Schattenseiten. Denn die LiDAR-Sensoren sind wie die Kamerasysteme anfällig für schlechte Sichtverhältnisse, wie sie bei Nebel oder Dämmerung herrschen. Ein weiterer Nachteil von LiDAR ist die fehlende Erkennung von Schriftzeichen auf Verkehrsschildern.

In jüngster Zeit spielt die Wärmebildtechnik eine immer wichtigere Rolle. Nach Eltahan und Elsayed (2024:147) kann der Einsatz einer Wärmebildkamera in der

Fahrzeugsensorik die Zuverlässigkeit erhöhen. Sie übertrifft andere Sensoren bei schlechten Lichtverhältnissen und auch bei Tageslicht kann mit dem Ansatz eine gute Leistung erzielt werden. Ferner weisen Eltahan und Elsayed in ihrem Artikel darauf hin, dass in letzter Zeit verschiedene Forschungen zur Sensorfusion, d.h. eine Zusammenschaltung unterschiedlicher Sensoren, gestartet wurden. Darin werden zum Beispiel Varianten aus Wärmebild- und LiDAR-Technologie untersucht, um die Vorteile der hochpräzisen LiDAR-Bilder mit den Wärmebildern für schlechte Lichtverhältnisse zu kombinieren (:147).

3.4 Entwicklungsprozesse

Für die Steuerung eines selbstfahrenden Fahrzeuges ist die Qualität und der Funktionsumfang der installierten Steuerungssoftware von großer Bedeutung. Ein Fehler kann zum Beispiel zu weitreichenden Konsequenzen für die Insassen führen. Auch aus rein finanzieller Sicht haben die Hersteller ihre eigene Motivation, kostspielige Rückrufaktionen zur Fehlerbeseitigung zu vermeiden. Aus diesem Grund nutzt die Automobilindustrie verschiedene Prozesse und Normen, die die Qualität der Software sicherstellen sollen. Im Folgenden soll auf eine in der Automobilindustrie weit verbreitete Methode, das sogenannte V-Modell, eingegangen werden.

3.4.1 V-Modell

Das Vorgehensmodell mit dem kurzen Namen „V-Modell“ ist ein Hilfsmittel, um die „Erfolgswahrscheinlichkeiten von Projekten“ zu verbessern (Rausch & Broy 2008:2). Es basiert auf einer Konzeption des Bundes für die Entwicklung von IT-Systemen, das sich aufgrund der Offenheit auch in anderen Branchen durchgesetzt hat. Das Modell als „Projektdurchführungsstandard“ (:4) hat verschiedene Ziele. Zum einen soll das Vorgehen eine bessere Planbarkeit und Projekttransparenz erzielen. Ein weiterer Punkt ist die Gewährleistung der notwendigen Qualität.

Die Grundstruktur folgt grafisch dem „V“ (siehe dazu die Abbildung 8). Das Modell lässt sich anwenden auf die Software-, die Hardware- sowie die Systementwicklung. Auf der linken Seite befindet sich die Spezifikations- und Zerlegungsphase. In einer ersten Phase werden die Anforderungen (eng. Requirements) mit dem Auftraggeber geklärt. Danach können Software-Architekten das Systemdesign ausarbeiten. Dieses

Systemdesign kann verglichen werden mit einer Bauzeichnung für ein Haus. Die einzelnen Räume sind definiert und wesentliche Ausstattungsmerkmale wurden ausgewählt, die die Anforderungen des Kunden erfüllen. In einer weiteren Phase kann es auf der linken Seite des „V“s zu einer Detailplanung und vor allem zur Umsetzung des vom Architekten definierten Plans kommen. Damit wurde die untere Spitze des „V“s erreicht. In weiteren Phasen hat das Vorgehensmodell auf der rechten Seite Überprüfungen der vorherigen Schritte vorgesehen. Die Mitarbeiter und Mitarbeiterinnen beginnen auf der kleinsten Ebene, der Unit-Ebene (oder auch Komponenten genannt), um entsprechende Überprüfungen durchzuführen. Im Wesentlichen geht es immer darum, die entsprechende Phase von der linken V-Seite durch eine passende Qualitätssicherungsmaßnahme auf der rechten Seite des Modells zu überprüfen. Um das Design des Architekten zu überprüfen, werden verschiedene Komponenten zu einer größeren Einheit integriert und das Verhalten getestet. In einem weiteren Schritt kann dann das System getestet werden. Dazu dienen die Anforderungen von der linken Seite als Eingabe. Die Tester überprüfen somit, ob das erstellte System die Anforderungen des Kunden erfüllt. Aufgefächert werden kann die Überprüfung in eine reine Software-Überprüfung oder auch in eine einzige System-Überprüfung mit Hardware- und Software-Anteilen.

Insgesamt lässt sich dieses Vorgehensmodell auch als „Wasserfall“-Modell bezeichnen. Damit wird deutlich gemacht, dass eine neue Phase erst begonnen werden kann, wenn die vorherige vollständig umgesetzt wurde. Vorteil dieser strukturierten Vorgehensweise ist das Überprüfen der einzelnen Entwicklungsschritte durch jeweilige Testaktivitäten. Dazu haben sich verschiedene Messmethoden etabliert, die den Qualitätsstand aktuell und oft auch automatisch ermitteln können.

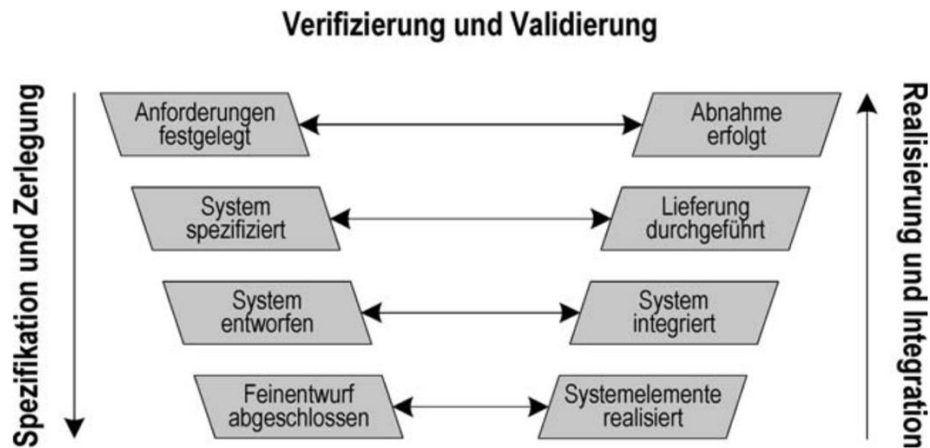


Abbildung 8: Vorgehensmodell V-Modell (Rausch & Broy 2008:17)

3.4.2 Kritik am V-Modell

Auf der anderen Seite hat dieses Modell seine Grenzen, wenn ein Prozessschritt nicht vollständig umgesetzt werden kann. Können die Anforderungen beispielsweise zu Beginn nicht genau oder vollständig definiert werden, so hat dies Einfluss auf die nachfolgenden Prozesse. Im Bereich der Software-Entwicklung hat sich die Anwendung einer „Wasserfall“-Methode als weniger erfolgsversprechend gezeigt und dagegen verbreitet sich die agile Prozessmethodik, um bei der Entwicklung in neue Gebiete erfolgreich voranzuschreiten.

Ein Grundgedanke des agilen Ansatzes ist es, die Zukunft in kleinere Einheiten zu unterteilen und im laufenden Fortschritt die Stakeholder einzubinden und so gemeinsam das Ziel zu erreichen. Dem Menschen fällt es im Allgemeinen schwer, eine sichere Zeitabschätzung für eine unbekannte Entwicklungsaufgabe zu machen. Auf der anderen Seite ist es einem Fachexperten möglich, aufgrund seiner Expertise eine Vorhersage für eine nahe Zukunft, wie zum Beispiel für die nächsten zwei Wochen, zu treffen. Diesen Umstand der Vorausschau auf kleineren Einheiten nutzt die agile Entwicklung aus. Die agile Methodik hilft somit, mit Unzulänglichkeiten und Unbekannten umzugehen.

Nach Anjum und Wolff versucht die Automobilindustrie, die Integration agiler Methoden mit dem V-Modell zu kombinieren (Anjum & Wolff 2021:141), obwohl dies im

Agilen Manifest⁷ nicht vorgesehen ist. Eine Lösung könnte daher sein, das V-Modell mit der agilen Sichtweise zu kombinieren, indem das V-Modell in kleineren Zeitscheiben wiederholt durchgeführt wird.

3.5 Künstliche Intelligenz

Weiter oben wurden die interdisziplinären Felder der Kybernetik und Mechatronik vorgestellt, die für die Automobilentwicklung von Bedeutung sind. Für die Entwicklung von selbstfahrenden Fahrzeugen kommt als weiteres interdisziplinäres Feld die künstliche Intelligenz (KI oder engl. AI für Artificial Intelligence) als Problemlösungs-Strategie in Frage. Beispiele für den Einsatz können das maschinelle Wahrnehmen der Umgebung (engl. Perception) oder auch die Vorhersage (engl. Prediction) sein. In diesem Kapitel werden dazu einige wesentliche Konzepte der KI vorgestellt.

Künstliche Intelligenz (KI) ist von Natur aus datenbasiert und stützt sich auf riesige Datensätze, um Modelle zu trainieren, die es Maschinen ermöglichen, Aufgaben auszuführen, die Intelligenz erfordern. Dieser Oberbegriff umfasst Technologien, die es Maschinen ermöglichen, sich in unterschiedlichem Maße intelligent zu verhalten. Diese datengesteuerten Modelle bilden die Grundlage für die Erstellung von Vorhersagen, Antworten oder Empfehlungen.

Die KI ist zu einem Hype geworden und hier fehlt ein halber Satz! Aktienkurse von Firmen, die notwendige Hardware für die Berechnungen von KI liefern, haben eine unglaubliche Wertsteigerung erzielt⁸. Als anderes Beispiel lässt sich die generative KI nennen, die für Suchanfragen ein Ergebnis aus einem gelernten Modell generiert und so Textinhalte, Bilder oder beispielsweise die Erstellung eines Podcast möglich macht.

Nun ist die KI nicht so neu, wie es gerade erscheint, da erste Arbeiten schon in den fünfziger Jahren des 20. Jahrhunderts gemacht wurden. Als einer der Urväter gilt

⁷ Das Agile Manifest basiert auf einer Erklärung von Werten und Prinzipien, die als Grundlage für die moderne Softwareentwicklung dienen.

⁸ Der Aktienkursverlauf des Chipherstellers NVIDIA Corporation (NVDA) ist zum Beispiel rasant angestiegen, da seine Produkte sich besonders gut für das Maschinelle Lernen eignen.

John McCarthy, der 1956 einen zweimonatigen Workshop zur künstlichen Intelligenz durchführte (Russell & Norvig, 2012:40). Russell und Norvig prägten den Begriff des „intelligenten Agenten“ (:55) in ihrem Werk *Artificial Intelligence: A Modern Approach*, das in der ersten Auflage 1995 erschien. Der Agent soll die bestmögliche Aktion in einer Situation ausführen. Als weiteres Beispiel ist ein Prozess des „Reinforcement Learning“ bekannt, bei dem es sich um eine Art maschinelles Lernen handelt, bei dem ein Agent lernt, Entscheidungen in einer Umgebung zu treffen, um möglichst viele Belohnungssignale zu erhalten. Er ahmt nach, wie Menschen durch Versuch und Irrtum lernen, wobei der Agent Handlungen ausführt, die Reaktion der Umgebung beobachtet und für seine Handlungen Belohnungen oder Strafen erhält.

Zusammenfassend gesagt wurde bei regelbasierten Ansätzen die Intelligenz der Maschine direkt vom Menschen beigebracht. Diese Regeln bestimmen das Verhalten und den Entscheidungsprozess des Systems. Die Logik wird direkt vom Programmierer „eingepägt“. Im Gegensatz dazu stehen KI-Systeme, insbesondere solche, die maschinelles Lernen einsetzen, denn sie lernen die Entscheidungslogik aus Daten. Während ein menschlicher Programmierer die Lernalgorithmen entwirft und die Daten auswählt, werden die spezifischen Entscheidungsregeln vom System entdeckt und nicht explizit programmiert.

Die Automatisierung dieses Prozesses des Lernens brachte die Innovationsschübe, die die Aufmerksamkeit der Öffentlichkeit auf sich gezogen haben. Im Folgenden wird eine Auswahl von wesentlichen Konzepten und Besonderheiten erläutert.

3.5.1 Maschinelles Lernen

Maschinelles Lernen (engl. machine learning oder kurz ML) ist ein angewandtes Verfahren innerhalb der Künstlichen Intelligenz. In diesem Prozess wird das Wissen automatisch von einer Maschine auf der Grundlage der während der Trainingsphase verwendeten Daten erworben.

Der Ansatz lässt sich in zwei Phasen unterteilen. In einer ersten Phase wird ein Modell mit problemspezifischen Daten (sogenannten Trainingsdaten) trainiert. In diesem Modell steckt letztendlich das Know-How über einen bestimmten Kontext. Dieser Schritt wird auch als KI-Modellbildung (AI model development) bezeichnet. In einer zweiten Phase kann das Modell genutzt werden, um Vorhersagen über die gegebenen Eingangsbedingungen zu machen. Diese notwendige Aufteilung der

Phasen bietet auf der anderen Seite auch neue Geschäftschancen: Ein Anwender oder eine Anwenderin könnte für seine oder ihre Zwecke ein trainiertes Modell von einem spezialisierten Anbieter mieten oder kaufen und spart sich damit die Kosten für die Modellentwicklung.

Durch das maschinelle Lernen wird der Mensch dabei entlastet, einer Maschine ein Wissen aufzubereiten und es für die Maschine in einer zugreifbaren Weise bereitzustellen (Janiesch, Zschech, Heinrich 2021:1). Um das maschinelle Lernen als Lösungsansatz erfolgreich einzusetzen, sind zwei Aspekte wesentlich: Zum einen braucht es genügend problemspezifische Daten und zum anderen genügend Rechenleistung.

3.5.2 Neuronale Netzwerke

Die Neurowissenschaften sammeln Erkenntnisse über die Funktionsweise des menschlichen Gehirns. Diese Erkenntnisse inspirierten ein weiteres Teilgebiet des maschinellen Lernens, nämlich die Nutzung von künstlichen neuronalen Netzwerken (auch Deep Learning⁹ genannt), die das Konzept des menschlichen Gehirns nachahmen. Das Gehirn besteht aus Nervenzellen, den Neuronen. Und diese einzelnen Neuronen verbinden sich mit anderen Neuronen über Verbindungspunkte, die als Synapsen bezeichnet werden. Über die Verbindungen werden Signale versendet und man nimmt an, dass das menschliche Gehirn über diese Mechanismen lernt (Russell & Norvig, 2012:33).

Die auf Deep Learning basierenden Modelle erzielen vielversprechende Ergebnisse. Hier bleibt es dem Fachmann bzw. der Fachfrau überlassen, ein solches System entsprechend für das Lernen zu konfigurieren und die Trainingsdaten auszuwählen. Janiesch, Zschech und Heinrich sprechen von „Hyperparametern eines Modells“, mit dem die Anzahl der Schichten und Neuronen sowie andere Eigenschaften festgelegt werden (Janiesch, Zschech & Heinrich 2021:688).

⁹ Künstliche neuronale Netze sind eine viel einfachere Form von Deep-Learning-Modellen, die aufgrund der hohen Anzahl von Neuronen und Schichten viel komplexer sind.

Die Zusammenhänge zwischen dem Maschinenlernen und dem Deep-Learning sind in Abbildung 9 dargestellt.

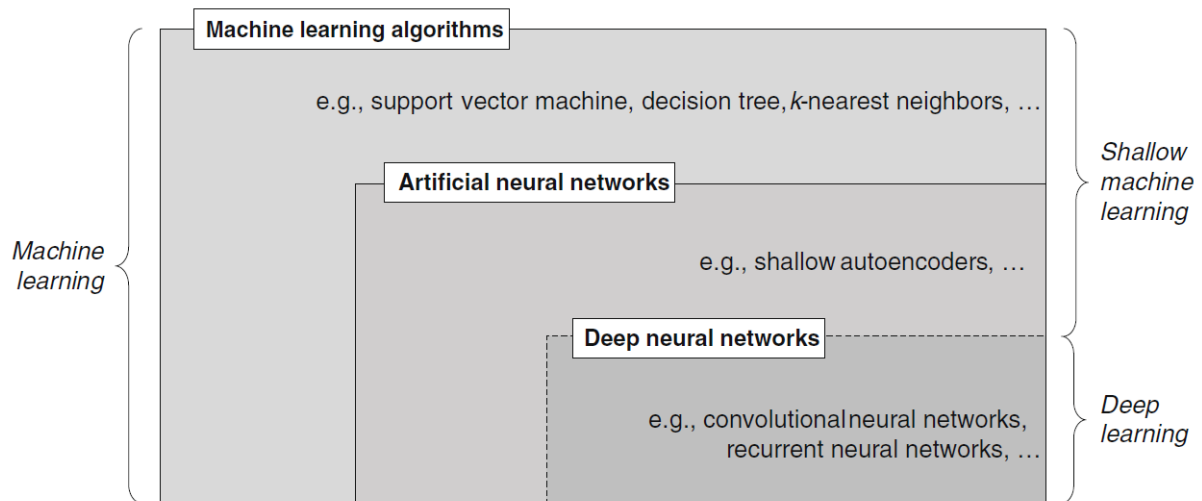


Abbildung 9: Konzepte des Machine Learnings (Janiesch, Zschech & Heinrich 2021:687)

3.6 Zwischenfazit

An dieser Stelle sollen einige Punkte aus den obigen Betrachtungen hervorgehoben werden. In diesem Kapitel wurde zunächst das Fahrverhalten eines Fahrers oder einer Fahrerin betrachtet (Kapitel 3.1). Aus der abstrakten Darstellung lassen sich Aufgaben für die Umsetzung in einem selbstfahrenden Fahrzeug ableiten. Wichtig ist hierbei die Zeit, die ein Fahrschüler oder eine Fahrschülerin benötigt, um ein Fahrzeug sicher zu führen. Ebenfalls relevant sind die notwendige Reaktionsgeschwindigkeit und das Erfassen der Situation in einer konkreten Fahrsituation.

Zweitens wurde die abstrakte Systemsicht eingeführt. Ausgehend vom Blockschaltbild der Fahraufgabe (siehe Abbildung 4: Erweitertes Modell der Fahraufgabe (Donges 2015:19)) können Erkenntnisse aus der Regelungstechnik und Mechanik auf das selbstfahrende Fahrzeug übertragen werden. Wenn ein Fahrer oder eine Fahrerin in Abhängigkeit vom Alter eine unterschiedliche Wahrnehmung der Umgebung aufweist, hat auch die Auswahl der Sensorik für das maschinelle Sehen einen Einfluss auf die sichere Erfassung der Verkehrssituation.

Drittens folgt die Entwicklung eines Fahrzeugs strengen Prozessen (Kapitel 3.4). Andere Branchen verwenden flexiblere und agilere Ansätze, so dass die Automobilindustrie einen „Anpassungsprozess“ eingeleitet hat.

Viertens wurde das maschinelle Lernen vorgestellt (Kapitel 3.5). Mit diesem neuen Ansatz ist eine Beschleunigung der Maschinenprogrammierung möglich geworden, so dass dieser Ansatz als „Enabler“ für die Entwicklung selbstfahrender Fahrzeuge angesehen werden kann.

4 FALLSTUDIEN UND DEBATTE

In diesem Kapitel werden zwei Fallstudien (Case Studies) vorgestellt, deren Untersuchung für die Beantwortung der Forschungsfragen hilfreich ist: Das eine Beispiel ist ein Verkehrsunfall mit einem Fahrzeug der Firma Tesla, das andere ein dokumentierter Unfall mit einem Robotertaxi der Firma Cruise (GM).

4.1 Kollision zwischen einem Tesla und einem Sattelzug

Die Kollision zwischen einem Auto der Marke Tesla und einem Lastwagen-Sattelzug (LKW) wurde vom National Transportation Safety Board (NTSB 2020) mit der Unfallnummer HWY19FH008 dokumentiert. Die folgende Tabelle zeigt die wichtigsten Punkte des NTSB-Berichts:

Unfallnummer	HWY19FH008
Unfalltyp	Kollision auf einer Kreuzung mit einem teilautomatisierten Fahrzeug
Ort	US Highway 441 Delray Beach, in Florida, USA Kreuzung mit der Zuwegung zum Firmengelände „Pero Family Farms“
Datum und Uhrzeit	1. März 2019 um 6:17 Ortszeit
Erstes Fahrzeug	Tesla Model 3, Baujahr 2018
Zweites Fahrzeug	Zugmaschine „International“, Baujahr 2019, mit einem Auflieger der Marke Vanguard aus dem Jahre 2008
Verletzte	Der Fahrer vom Tesla wurde tödlich verletzt. Der Fahrer vom LKW blieb unverletzt.

In die Auswertung flossen zahlreiche Filmaufnahmen ein. Das Firmengelände wurde an der Einfahrt mit einer Überwachungskamera gefilmt. Darüber hinaus stellte die Firma Tesla der Untersuchungskommission die Videoaufzeichnungen des Unfallfahrzeugs sowie weitere interne Fahrzeugdaten zur Verfügung. Die Aufnahmen vom Fahrzeug selbst sind recht umfangreich, da das Modell wie üblich mit zahlreichen Kameras ausgestattet war:

Tesla also provided the NTSB with a series of still images from all eight of the car's cameras. The eight cameras included three forward-facing cameras (long-range, short-range, and wide-angle), two rearward-facing side-angle cameras (mounted under each sideview mirror), two forward-facing side-angle cameras (mounted on each B-pillar), and one rearview camera (mounted above the license plate) (NTSB 2020:3).

Im Folgenden wird der Unfallhergang auf Basis des Unfallberichts der NTSB (2020) zusammengefasst:

Die Anbindung der landwirtschaftlichen Firma Pero Family Farms an den Highway US 441 erfolgt zweiseitig (siehe Abbildung 10). Für Radfahrer gibt es auf beiden Seiten der Straße eigene Fahrspuren. Die Kreuzung ist nicht beleuchtet, laut NTSB reicht das Dämmerungslicht, um „Aktivitäten“ ohne künstliches Licht durchzuführen: „On the day of the crash, civil twilight began at 6:21 a.m., and sunrise was at 6:44 a.m.“ (:5). Die Straßen sind trocken und im Bereich der Unfallstelle sind keine Behinderungen bekannt.

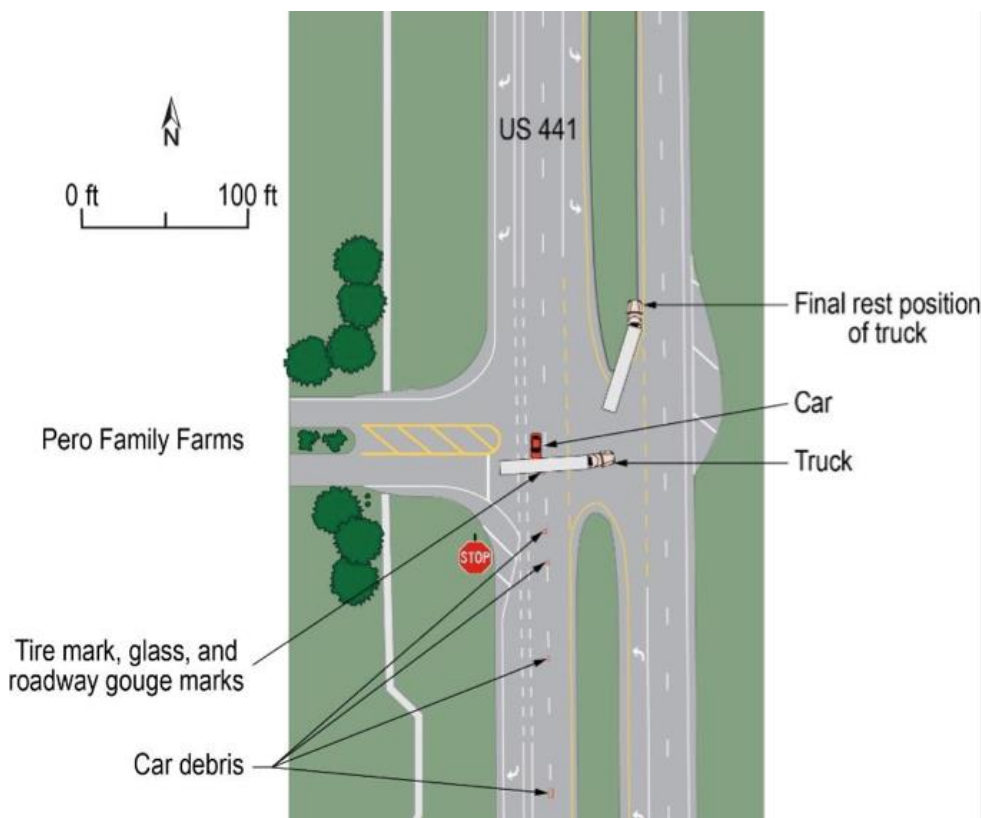


Abbildung 10: Straßenkarte mit Positionen (NTSB 2020:2).

Der Sattelschlepper beginnt am Morgen des 1. März 2019, die landwirtschaftliche Anlage der Firma Pero Family Farms zu verlassen, um auf die Straße US 441 in Richtung Norden zu fahren (siehe Pfeilrichtung 2 in Abbildung 11). Bei Erreichen des Stoppschildes an der Kreuzungseinfahrt verlangsamt er seine Geschwindigkeit, ohne

jedoch vollständig zum Stillstand zu kommen. Nachdem ein Fahrzeug die Kreuzung passiert hatte, beschleunigte der Lkw-Fahrer, um auf die US 441 aufzufahren. Zeitgleich kommt ein Tesla-Fahrer aus der Nord-Richtung (siehe Pfeilrichtung 1). Wie sich später bei der Unfallanalyse herausstellte, war die automatische Fahrfunktion „Autopilot“ aktiviert und der Fahrer hatte die Hände nicht am Lenkrad. Die Geschwindigkeit des Fahrzeugs betrug 69 mph (ca. 111 km/h), erlaubt waren auf diesem Streckenabschnitt nur 55 mph (ca. 88 km/h). Das automatische System erkannte den kreuzenden Sattelschlepper nicht bzw. nicht rechtzeitig, so dass weder der Fahrer noch das automatische System eine (Not-) Bremsung bzw. Verringerung der Geschwindigkeit eingeleitet haben. Dadurch kollidierte der Pkw mit dem Lkw genau in der Mitte des Sattelauflegers. Aufgrund der Wucht des Aufpralls wurde das Dach des Teslas abgeschert und das untere Teilstück vom Auto kam ungefähr 500 m weiter in südlicher Richtung zum Stehen. Als Konsequenz dieser schweren Kollision kam der 50-jährige Autofahrer ums Leben. Der 45-jährige LKW-Fahrer blieb dagegen unverletzt. Die Abbildung 11 zeigt eine Satelliten-Aufnahme vom Unfallort (:5).

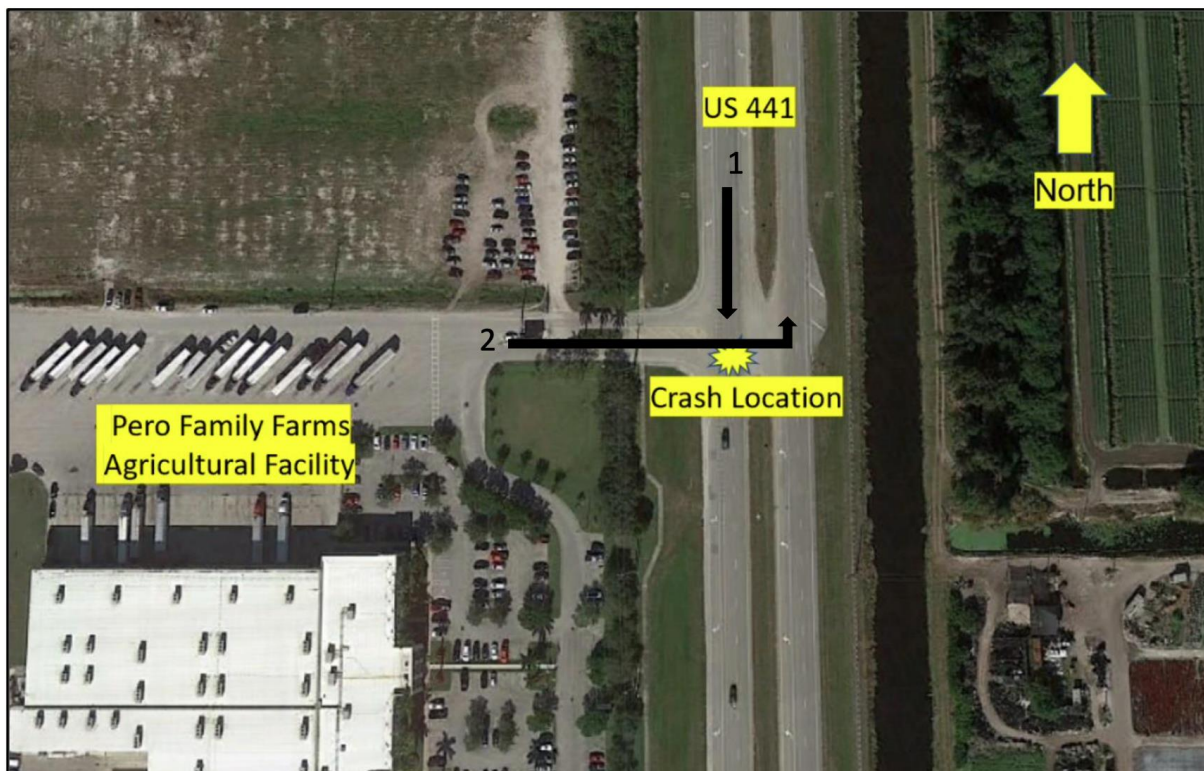


Abbildung 11: Unfallort aus dem Bericht der NTSB (2020:5), ergänzt um Fahrtrichtungen.

Nach dem Unfall sammelten die Untersuchungsbeauftragten Informationen über die beiden Unfallbeteiligten. Der 45-jährige Lkw-Fahrer war im Besitz eines Führerscheins für gewerbliche Fahrer in Florida. Er arbeitete für die Firma FirstFleet, die

wiederum für die Firma Pero Family Farms die Transporttätigkeiten übernommen hatte. Seine Schicht an dem Unfalltag begann um 2:00 Ortszeit. Die Fahrzeiten entsprachen den gesetzlichen Vorgaben. Außerdem konnte keine Ablenkung durch die Benutzung eines Mobiltelefons nachgewiesen werden:

The driver's electronic logs show that he had an adequate sleep opportunity before the crash. Cell phone records show that the driver was not engaged in a phone conversation and was not texting at the time of the crash. (:6)

Zwei gesundheitliche Einschränkungen erschienen in dem Untersuchungsbericht als erwähnenswert: Der Lkw-Fahrer leidet seit seiner Kindheit an Epilepsie, die jedoch mit entsprechenden Medikamenten behandelt wird. Ferner wurden seine Augen im Jahre 2012 mittels refraktiver Chirurgie behandelt. Der Fahrer gab in der Zeugenaussage an, dass er fähig sei „to ‚read with my right eye‘ and ‚see my distance in my left eye““ (:6). Diese auch Monovision genannte Einschränkung führt laut Analysebericht typischerweise zu einer Minimierung der Tiefenwahrnehmung. Bei der ärztlichen Untersuchung des Fahrers im November 2017 wurde dem Fahrer jedoch bescheinigt, dass seine Sehleistung im normalen Bereich liegt (:7). Ein nach dem Unfall von der Firma FirstFleet veranlasster Alkohol- und Drogentest verlief negativ.

Der 50-jährige Tesla-Fahrer war, wie Zeugen aussagten, in einem guten gesundheitlichen Zustand. Auch ihm konnte während der Unfallfahrt keine Ablenkung durch ein Mobiltelefon nachgewiesen werden (:8). Thadani, Lerman, Piper, Siddiqui und Uraizee (2023) hatten zudem herausgefunden, dass der Tesla Fahrer verheiratet und Vater von vier Kindern war. Beruflich war er als Software-Engineer tätig, was als eine gewisse Affinität zur Technik gedeutet wurde. Seine Frau gab an, dass ihr Mann sich vor dem Kauf des Fahrzeuges mit der Technologie des Autopiloten beschäftigt hatte und sich die Tesla Marketingvideos angesehen hätte, in denen der Fahrer auf der Rücksitzbank gesessen hätte und das Fahrzeug autonom gefahren sei (Thadani u.a. 2023:18).

Ablauf des Unfalls aus Sicht des Tesla-Fahrers:

Ca. 5:50 Uhr	Der Tesla Fahrer ist am 1. März 2019 mit seinem Tesla Model 3 aus dem Jahre 2018 auf dem Weg zur Arbeit.
6:16 Uhr Minus 12,3 Sek.	Der Tempomat wird mit einer Höchstgeschwindigkeit von 69 mph (ca. 111 km/h) aktiviert, obwohl die Geschwindigkeitsbegrenzung auf der U.S. 441 55 mph (88 km/h) beträgt. Es sind noch 12,3 Sekunden bis zum Aufprall, bzw. 380 Meter.
Minus 9,9 Sek.	Der Autopilot mit der Version 2.4 wird aktiviert. Es sind noch rund 10 Sekunden bis zum Aufprall, bzw. 308 Meter. Der Autopilot zeigt die übliche Warnmeldung: „Please keep your hands on the wheel. Be prepared to take over at any time.“ (Thadani u.a. 2023)
Minus 8 Sek.	Es sind noch rund 8 Sekunden bis zum Aufprall, bzw. 247 Meter. Mit Bezug auf den Daten-Aufzeichnungen von Tesla (data log „no driver-applied wheel torque“) werden fehlende Fahrerhände am Lenkrad festgestellt (Thadani u.a. 2023).
	Nach NTSB sollte nach ca. 25 Sekunden eine Warnung bzgl. der fehlenden Fahrerwachsamkeit (Fahrerhände) angezeigt werden. Diese Meldung kam jedoch nicht.
	Der LKW überquert die US 411. Der Fahrer wird langsamer, führt jedoch keinen vollständigen Halt am Stoppschild durch.
Minus 2 Sek.	Es sind noch rund 2 Sekunden bis zum Aufprall, bzw. 62 Meter. Die Tesla Front-Kamera filmt den LKW (Abbildung 12), jedoch wird keine Warnung an den Fahrer angezeigt oder eine Bremsung eingeleitet.
0 Sek.	Kollision.



Abbildung 12: Teslas Frontkamera (Thadani u.a., 2023)

4.2 Kollision zwischen einem Robotertaxi und einer Fußgängerin

Die Kollision zwischen einem Taxi der Marke Cruise und einer Fußgängerin führte als Konsequenz zu einer Rückrufaktion der gesamten Fahrzeugflotte (NHTSA 2023). Als Folge mussten 950 Fahrzeuge mit einem Update der Steuerungssoftware ausgestattet werden (Thadani 2023b). Zudem wurde aufgrund des Vorfalls der fahrerlose Testbetrieb der Firma Cruise in San Francisco vorerst untersagt. Im Laufe des Verfahrens hat die Firma Cruise die Anwaltskanzlei Quinn Emanuel mit der Untersuchung des Falls beauftragt (Quinn Emanuel 2024). Im Rahmen dieser Studie wurde auch das technische System seitens der Firma Exponent (2023) analysiert. Der gesamte Untersuchungsbericht wurde der Öffentlichkeit in groben Zügen zur Verfügung gestellt, d.h. Teile, die vertrauliche Informationen über implementierte Algorithmen und Lösungsansätze enthalten, wurden geschwärzt.

In der folgenden Tabelle sind die Eckpunkte des Unfalls aus den zugänglichen Informationen zusammengefasst. Wie sich später herausstellte, war auch ein zweites Fahrzeug, ein Nissan Sentra, an dem Unfall beteiligt.

Unfallnummer	Nicht bekannt
Unfalltyp	Kollision auf einer Kreuzung mit einem vollautonomen Fahrzeug mit einer Fußgängerin
Ort	San Francisco, Kalifornien, USA Kreuzung zwischen der 5th. und Market.
Datum und Uhrzeit	2. Oktober 2023, 21:29 Uhr Ortszeit

Erstes Fahrzeug	Cruise Taxi mit dem Namen „Panini“
Zweites Fahrzeug	Nissan Sentra (Nissan)
Verletzte	Eine Fußgängerin. Die Insassen der Fahrzeuge blieben unverletzt.

Der Unfall ereignete sich an der Kreuzung zwischen der Cyril Magnin Street, wo sie in die 5th Street übergeht, und der Market Street in San Francisco. Die Abbildung 13 zeigt eine annotierte Luftbildaufnahme vom Unfallort (das Bild wurde zu einem anderen Zeitpunkt aufgenommen). Die Fahrrichtung des Cruise Taxis im autonomen Betrieb ist mit der Pfeilrichtung 1 dargestellt. Ferner fährt parallel zum Cruise Taxi ein Nissan Sentra mit der Pfeilrichtung 2 auf die Kreuzung. Die später verletzte Fußgängerin passiert die Kreuzung in Pfeilrichtung 3. Die Unfallstelle ist durch den blauen Rahmen in der Abbildung abgegrenzt.

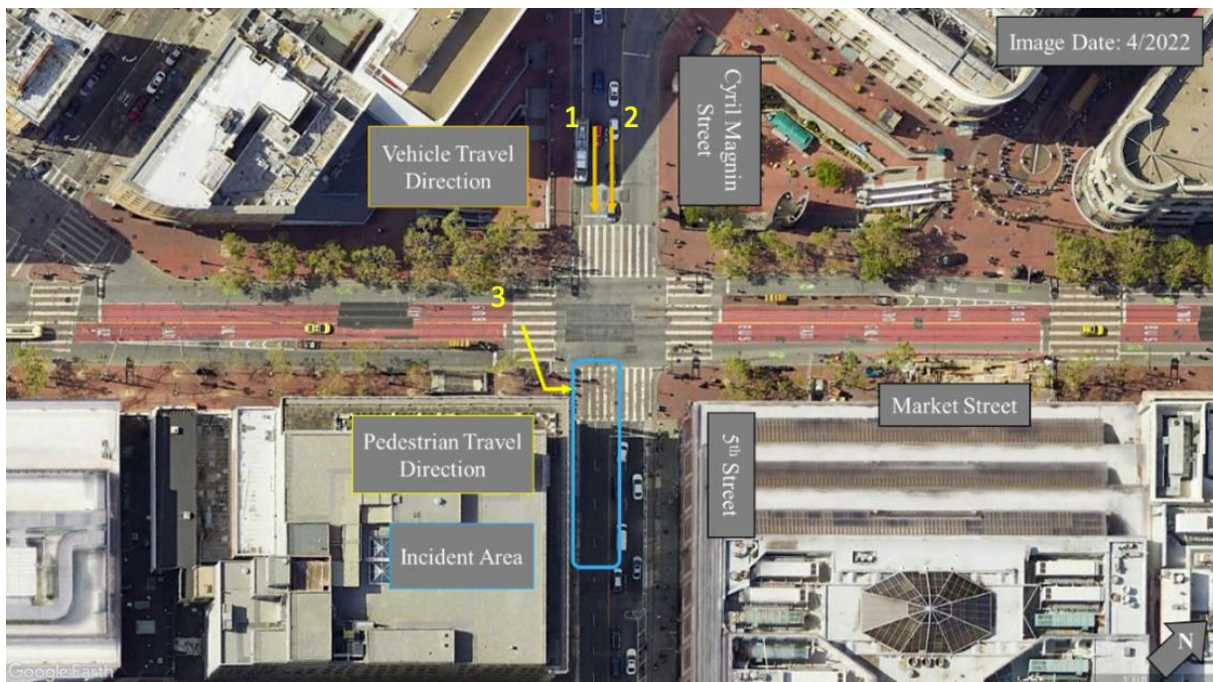


Abbildung 13: Unfallort in San Francisco (Exponent 2023:24), ergänzt um Positionen

Das fahrerlose Cruise Taxi hielt an der roten Ampel an der Kreuzung auf der rechten Spur an (siehe Position 1 in Abbildung 13). Gleichzeitig hält ein Nissan auf der linken Spur neben dem Cruise Taxi an (siehe Position 2).

Nach Umschaltung der Ampel auf „grün“, fahren das Cruise Taxi und der Nissan in gleiche, südliche Richtung. Zu diesem Zeitpunkt überquert eine Fußgängerin den gegenüberliegenden Zebrastreifen an südwestlicher Ecke der Kreuzung in die östliche Richtung und missachtet dabei das Fußgängersignal „Do Not Walk“. Die Fußgängerin überquerte die Fahrspur des Robotertaxis und hielt auf der Fahrspur des Nissans an (d.h. auf der linken Fahrspur). Dabei wurde die Frau von dem Nissan angefahren und infolgedessen auf die Fahrspur des Cruise Taxis geschleudert. Das Taxi erkannte die Situation und führte eine Bremsung durch. Jedoch kam es nicht rechtzeitig zum Stehen und „berührte“ die Fußgängerin. Nach diesem initialen Stopp startete das Taxi erneut eine Fahrt von ungefähr sechs Metern bis zum endgültigen Haltepunkt. Dabei wurde die Fußgängerin vom Taxi mitgezogen, was zu weiteren Verletzungen geführt hatte. Die weiteren Ermittlungen ergaben, dass der Fahrer des Nissan Fahrerflucht begangen hatte.

4.3 Hauptlinien in der ethischen Debatte

Einleitend zu diesem Kapitel sei zunächst auf die mittlerweile langjährige Diskussion um autonome Fahrzeuge und deren möglichen Dilemmasituationen hingewiesen. Auch in diesem Kapitel sind die Dilemmasituationen Gegenstand der Diskussion. Die Debatte hat mit der Zulassung und Verbreitung der ersten autonomen Fahrzeuge an Brisanz gewonnen. Wie Kriebitz, Max & Lütge (2022) feststellen, hat in Deutschland die Änderung der Straßenverkehrsordnung für die Zulassung von autonomen Fahrzeugen nach SAE Level vier die Aufmerksamkeit der Fachöffentlichkeit auf sich gezogen.

Im Folgenden werden die Hauptlinien in der Diskussion aufgegriffen, die im Zusammenhang mit der Forschungsfrage und den konkreten Unfällen in Zusammenhang gebracht werden können.

4.3.1 Ethische Theorien und Leitlinien

Die Debatte um die Anwendung der ethischen Konzepte lässt sich aufteilen in ethische Grundtheorien. Zu nennen wären die deontologischen Theorien (d.h. mit dem Fokus „Plicht“) und die teleologischen Theorien (d.h. mit dem Fokus „Ziel“). Für einen generellen Überblick sei hier auf die Einführung von Pauer-Studer (2020) verwiesen. Die Diskussion in der Wissenschaft zum autonomen Fahren ist sehr breit gefächert und nach einer Analyse der bestehenden Literatur kommen Poszler, Geisslinger, Betz und Lütge zu der Auffassung, dass innerhalb der Gruppe der Wissenschaftler und Praktiker keine Einigkeit besteht, welche ethische Theorie (wenn überhaupt) bei der Anwendung auf die Entscheidungsfindung bei autonomen Fahrzeugen zu befürworten wäre (Poszler, Geisslinger, Betz & Lütge 2023:1).

Für die Entwicklung nutzen Ingenieure gerne Leitlinien. Es würden sich beispielsweise „Asimovs Gesetze der Robotik“ anbieten, die Isaac Asimov im Jahre 1942 veröffentlicht hat. Der Öffentlichkeit wurden diese Regeln durch den Film „I, Robot“ (Regie Alex Proyas, USA, 2004) bekannt. Zu Beginn des Films werden den Zuschauern die Regeln der „neuen technischen Sklaven“ (Nida-Rümelin & Weidenfeld 2020) verdeutlicht. Für die Anwendung in selbstfahrenden Fahrzeugen hat Goodal die Regeln entsprechend umformuliert:

1. An automated vehicle may not injure a human being or, through inaction, allow a human being to come to harm.
2. An automated vehicle must obey orders given it by human beings except where such orders would conflict with the First Law.
3. An automated vehicle must protect its own existence as long as such protection does not conflict with the First or Second Law. (Goodall 2014:7)

Asimov erweiterte später die Regeln noch um ein „nulltes Gesetz“: „Kein Roboter darf die Menschheit schädigen oder durch Untätigkeit zulassen, dass sie geschädigt wird.“ (Nida-Rümelin & Weidenfeld 2020:209). Diese Regeln könnten von Ingenieuren in einem autonomen Fahrzeug berücksichtigt werden. Goodall (2014) äußert hier Zweifel, dass dieser Ansatz jedoch zum Ziel führen wird. Dazu bringt er als Argument, dass ein Fahrzeug keine Notbremsung durchführen könnte, da dadurch Fahrzeuginsassen ein Schleudertrauma bekommen könnten:

[...] the first rule forbids an automated vehicle to allow a human to “come to harm.” A literal interpretation of this rule would prohibit sudden braking, even to avoid a collision, because it would cause whiplash for the vehicle occupants (Goodall 2014:7)

Die Implementierung von konkreten Regeln erscheint aus seiner Sicht als „eher aussichtslose Aufgabe“, wenn neue Unfallarten und Straßentwürfe berücksichtigt werden sollen. Goodall erkennt auf der anderen Seite die Orientierungshilfe von Regeln wie der von Asimov an, findet es jedoch schwierig, diese komplexe menschliche Ethik in einem Fahrzeug umzusetzen (Goodall 2014:8).

Die Forscher Bonnefon, Sharrif und Rahwan (2016) haben Interessierten eine „Moral Machine“ via Webseite¹⁰ zur Verfügung gestellt, mit der jede und jeder die Entscheidungen von intelligenten Maschinen verstehen und besser bewerten kann. Am Ende der Abfragen erfolgt eine Bewertung verschiedener Problemfelder auf einer Skala von „nicht wichtig“ bis „sehr wichtig“. So wird z.B. vom Grundsatz der Gleichbehandlung abgewichen und es werden Präferenzen bezüglich Alter, Sportlichkeit, Geschlecht und sozialem Wert in die Bewertung einbezogen. Nyholm (2018) äußert sich skeptisch zu den Ergebnissen der Moral Machine. Zunächst argumentiert er, dass die Nutzer derzeit noch wenig Erfahrung mit autonomen Fahrzeugen haben und sich ihre Einstellung wahrscheinlich ändern wird, sobald sie

¹⁰ <https://www.moralmachine.net>.

mehr Erfahrungen gesammelt haben. Weiters vermisst Nyholm die Begründung einer Entscheidung, da die Forscher nur die Präferenz aus einer vorgegebenen Liste testen. Aus seiner Sicht ist es jedoch wichtig, die Argumente für oder gegen etwas zu berücksichtigen. Schließlich, so Nyholm, seien Menschen widersprüchlich oder hätten paradoxe Einstellungen: Die Menschen wünschen sich, dass andere schadensminimierende Fahrzeuge benutzen, während sie selbst ein Fahrzeug bevorzugen, das die Insassen schützt (Nyholm 2018:5).

Durch den ehemaligen Bundesminister für Verkehr und digitale Infrastruktur Alexander Dobrindt¹¹ wurde die „Ethik-Kommission für Automatisiertes und Vernetztes Fahren“ eingesetzt. Geleitet wurde die Kommission von dem früheren Bundesverfassungsrichter Dr. Dr. Udo Di Fabio. Ziel der Arbeit sei es, „die notwendigen Leitlinien für das automatisierte und vernetzte Fahren zu erarbeiten“ (BMDV 2017:7). Für diese Arbeit sind die Arbeitsgruppen „unvermeidbare Schadenssituationen“, geleitet von Prof. Dr. Dr. Eric Hilgendorf, und „Verantwortungsreichweite für Software und Infrastruktur“, geleitet von Herrn Prof. Dr. Dr. Kagermann interessant. Die ethischen Aspekte wie die sogenannten Dilemma-Situationen wurden durch Herrn Prof. Dr. Dr. Julian Nida-Rümelin in die Arbeit der Ethik-Kommission eingebracht. Das Ergebnis der Kommission ist ein Rahmenwerk von 20 ethischen Regeln für den automatisierten und vernetzten Straßenverkehr. Im Folgenden sollen einige Aspekte besonders erwähnt werden, die im Licht der beiden Unfallsituationen von Bedeutung sind.

In Regel eins wird der Hauptvorteil von teil- oder vollautomatisierten Verkehrssystemen in der Verbesserung der Sicherheit aller Beteiligten gesehen. Regel zwei nennt als vertretbare Voraussetzung für die Einführung der Automatisierung, dass eine positive Risikobilanz im Vergleich zur menschlichen Fahrleistung ermittelt werden kann. Mit der fünften Regel formuliert die Ethik-Kommission die Forderung, dass die Technik so gestaltet und verbessert werden muss, dass kritische Situationen gar nicht erst entstehen:

[...] dazu gehören auch Dilemma-Situationen, also eine Lage, in der ein automatisiertes Fahrzeug vor der „Entscheidung“ steht, eines von zwei nicht abwägungsfähigen Übeln verwirklichen zu müssen. (BMDV 2017:10)

¹¹ Amtszeit 2013-2017.

In der folgenden sechsten Regel geht die Kommission davon aus, dass die automatische Kollisionsvermeidung gesellschaftlich und ethisch geboten sei. Auf der anderen Seite werden Bedenken geäußert, wenn es zu einer „praktischen Unentrinnbarkeit“ (BMDV 2017:11) kommt, d.h. sich der Mensch dem technischen System unterwerfen muss.

Eine echte dilemmatische Entscheidung, wie die Entscheidung Leben-gegen-Leben, ist aus der Sicht der Ethiker nicht normierbar und in einem System somit im Voraus auch nicht programmierbar (:11) dahingehend, dass es einen sittlich urteilsfähigen, verantwortlichen Fahrzeugführer vorwegnehmen oder ersetzen könnte.

In Deutschland wurde 2021 ein nationaler Rahmen für selbstfahrende Fahrzeuge erlassen und dieses Gesetz beschleunigte laut Kriebitz, Max und Lütge (2022) den Diskurs und hat die „Aufmerksamkeit von politischen Entscheidungsträgern, KI-Ethikern und Rechtsexperten für autonomes Fahren auf sich gezogen“ (Kriebitz, Max & Lütge 2022:1). Die Diskurse zur Ethik des autonomen Fahrens und der künstlichen Intelligenz können sich dabei überlappen, da traditionell die Ethik des autonomen Fahrens der Ethik der KI subsumiert wird (:1). Die besondere Rolle der Ethik des autonomen Fahrens ergibt sich aus den Auswirkungen, die ein fehlerhaftes System auf das Leben oder Eigentum der Menschen haben kann. Die beiden oben dargestellten Unfälle mit dem Tesla und dem Cruise Taxi sind zur Darstellung sehr gut geeignet und gehen dabei durch ihre Komplexität über andere Gedankenexperimente hinaus (zu den Gedankenexperimenten siehe die Ausführungen in Kapitel 4.3.2).

Die Berichte in den Medien veranlassten Rechtsexperten und Ethiker, über neue Rahmenregelungen für das autonome Fahren nachzudenken. Zum Beispiel adressiert das AI4People-Rahmenwerk von Floridi et.al. (Floridi et.al. 2018) fünf übergreifende Leitlinien. Vier der Prinzipien sind identisch mit denen der Medizinethik. Es handelt sich um die Prinzipien des eindeutigen Nutzens (Benefizienz), der Umkehrung, d.h. dass kein anderer Schaden erleiden darf (Nicht-Malefizien), der Gerechtigkeit und schließlich der Entscheidungsfreiheit, d.h. der Autonomie. Für weitere Informationen zur Medizinethik sei hier auf die Arbeit von Beauchamp (2021) verwiesen.

Ein fünftes Prinzip ist die Explizierbarkeit, d.h. die KI soll transparent und nachvollziehbar sein. Diese Rahmenregelungen gehen nach Kriebitz, Max & Lütge auf

bestimmte „Hochrisikobereiche“ ein, die menschliche Aufsicht und Prozesse erfordern, um nachteilige Auswirkungen auf die Rechte des Einzelnen abzumildern (siehe Kriebitz, Max & Lütge 2022:3).

“High risks” are identified when the consequences of AI are deemed irreversible or when an AI solution is likely to reproduce biases that affect vulnerable or historically disadvantaged groups negatively. (Kriebitz, Max & Lütge 2022:3).

Im Jahr 2021 hat die deutsche Bundesregierung die Straßenverkehrsordnung für den Betrieb von autonomen Fahrzeugen erweitert (BMDV 2021). Darin sind einige Eckpunkte für die erste gesetzliche Grundlage in Deutschland festgelegt. Neu ist die Forderung nach einer ständigen Verbindung des Fahrzeugs mit einer Leitstelle (die „technische Aufsicht“). Zudem kann die Leitstelle bei einem Fehlerfall auf das Fahrzeug zugreifen. Des Weiteren ist der Abschluss einer Haftpflichtversicherung für das Personal der Leitstelle Pflicht. Darüber hinaus definierte der Gesetzgeber den neuen Rechtsbegriff eines „risikominimalen Zustandes“:

Risikominimaler Zustand im Sinne dieses Gesetzes ist ein Zustand, in den sich das Kraftfahrzeug mit autonomer Fahrfunktion auf eigene Veranlassung oder auf Veranlassung der Technischen Aufsicht selbständig versetzt, um unter angemessener Beachtung der Verkehrssituation die größtmögliche Verkehrssicherheit für andere Verkehrsteilnehmende und Dritte zu gewährleisten. (BMDV 2021:6)

Und schließlich gibt es die gesetzliche Verpflichtung der Hersteller, Daten während der Fahrt aufzuzeichnen und den Behörden für Ermittlungen zur Verfügung zu stellen.

Bemerkenswert nach Kriebitz, Max und Lütge ist die ausdrückliche Festlegung auf Fahrzeugtypen mit SAE Level 4 und höher (Kriebitz, Max & Lütge 2022:10). Damit wurde die Mischform von Fahrzeugen mit Fahrer oder Fahrerin und einem aktiven autonomen System von der Erweiterung der Verordnung ausgenommen. Nach einer Analyse des deutschen Gesetzes stellen Kriebitz, Max und Lütge fest, dass

[...] das Gesetz das Problem der unvermeidbaren Unfälle nur unzureichend behandelt und das Problem, das in der Spannung zwischen Transparenz- und Sicherheitsprinzipien begründet ist, nicht löst. (Kriebitz, Max & Lütge 2022:2)

Nach ihrer Meinung ist das Gesetz eher als ein „Proof-of-Concept“¹² anzusehen, das Lücken für die Umsetzung lässt. Nach ihrer Meinung gibt es zwei Aspekte, die für

¹² Auch der Machtbarkeitsnachweis ist eine Methode, die dazu dient, zu beweisen, dass eine Idee die gewünschten Ergebnisse erzielen kann.

das autonome Fahren entscheidend sind: Die Hersteller müssen sich um Transparenz bemühen und für Datensicherheit (auch Cybersecurity genannt) sorgen, damit Vertrauen in die neue Technologie wachsen kann (Kriebitz, Max und & Lütge 2022:18).

Immer wieder wird von Auffahrunfällen auf Stauenden berichtet, die häufig zu fatalen Folgen für die beteiligten Verkehrsteilnehmer führen. Aus diesem Grund wurde der sog. Notbremsassistent (Advanced Emergency Braking System, AEBS) entwickelt. Dieses System warnt in einem ersten Schritt den Fahrer oder die Fahrerin vor einer möglichen Kollision. Es wird somit die Unachtsamkeit des Fahrers, der Fahrerin abgemildert, die – wie oben beschrieben – die häufigste Unfallursache darstellt. In einem weiteren Schritt kann eine Notbremsung durchgeführt werden. Auch wenn es dann noch zu einem Aufprall kommt, kann die reduzierte Geschwindigkeit die Wucht des Aufpralls verringern und das Verletzungsrisiko für alle beteiligten Verkehrsteilnehmer und Verkehrsteilnehmerinnen reduzieren. Laut dem Verband der Automobilindustrie müssen ab Mitte 2022/24 alle neuen Fahrzeugtypen, alle erstzugelassenen Pkw und leichte Nutzfahrzeuge mit einem Notbremsassistenten ausgestattet sein (VDA 2024). Die Umsetzung für schwere Nutzfahrzeuge wurde dabei leider ausgeschlossen und soll separat verhandelt werden.

4.3.2 Die Debatte um Dilemma-Situationen

Die wohl bekanntesten Gedankenexperimente in der Ethik sind das Trolley- und das Weichensteller-Problem¹³, bei denen das Dilemma der Unterscheidung zwischen einem oder fünf Unfalltoten vorgenommen werden soll. Das Experiment lässt sich gut kurz erzählen: Eine Bahn ist außer Kontrolle geraten und es besteht die Gefahr, dass fünf Personen, die am Gleis warten, getötet werden. Der Weichensteller erkennt diese Gefahr und könnte durch ein Eingreifen den Zug auf ein anderes Gleis umleiten. Dadurch würde er jedoch eine dort wartende Person töten. Hinter der Geschichte steht die moralische Frage, ob der Tod einer Person in Kauf genommen werden darf, um das Leben von fünf Personen zu retten. Abwandlungen des Gedankenexperimentes existieren ebenfalls mit Vergleichen zwischen älteren Personen und Kindern, einem entgegenkommenden Bus an einer Serpentine oder

¹³ Das Trolley-Problem geht auf Philippa Foot zurück (1967). Das Weichensteller-Problem wurde erstmalig von Welzel (1951) beschrieben.

Motorradfahrern mit unterschiedlicher (Kopf-) Schutzausrüstung (siehe z.B. Gurney 2016). Als weiteres Forschungsbeispiel ist die „Moral Machine“¹⁴ bekannt geworden. Auch in diesem Experiment soll das Dilemma deutlich gemacht werden: Wie entscheidet man in einer Dilemma-Situation, in der es Tote geben wird und man selbst entscheiden muss, wer oder wie viele Menschen in den nächsten Sekunden sterben werden?

In der deutschen Straßenverkehrsordnung gibt es ein Spannungsfeld zwischen zwei Eckpunkten, in dem sich eine Fahrerin und ein Fahrer bewegen müssen: Zum einen gilt der Vertrauensgrundsatz als Rechtsgrundsatz. Zum anderen wird von den Verkehrsteilnehmern eine defensive Fahrweise erwartet¹⁵. Das Vertrauensprinzip ermöglicht nach Gasser (2015) einen flüssigen Verkehrsfluss, da ein Fahrer und eine Fahrerin nicht mit allen möglichen und seltenen Verstößen anderer Verkehrsteilnehmer und Verkehrsteilnehmerinnen rechnen müssen. Andererseits müsse ein Fahrer defensiv fahren und mit Fehlern anderer Verkehrsteilnehmer rechnen. Wie dieses Gebot des defensiven Fahrens im Einzelfall anzuwenden ist, lässt sich laut Gasser aus verschiedenen Urteilen in Deutschland ableiten. Auf kontinentaleuropäischer Ebene kann aus den Urteilen jedoch allenfalls eine Orientierung, nicht aber eine Präjudiz, d.h. eine richtungweisende Rechtsprechung, abgeleitet werden (Gasser 2015:560).

Die Einhaltung der Straßenverkehrsordnung müsste konsequenterweise auch von einem autonomen Fahrzeug mit künstlicher Steuerung gewährleistet werden. Und hier kommt das oben dargestellte Spannungsfeld ins Spiel, denn ohne das Vertrauensprinzip könnte sich ein autonomes Fahrzeug kaum bewegen, insbesondere nicht im Mischbetrieb, d.h. auf einem Fahrweg mit gleichzeitig von Menschen gesteuerten Fahrzeugen. Darüber hinaus wird eine defensive Fahrweise gefordert, die mit Fehlverhalten anderer Verkehrsteilnehmerinnen und Verkehrsteilnehmer rechnet. Gerade dieser Aspekt ist in den untersuchten Fällen jeweils relevant: Im Fall „Tesla“ missachtete der beteiligte Lkw-Fahrer ein Stoppschild. Eine defensive Fahrweise des Tesla hätte die Unfallfolgen vermutlich

¹⁴ <https://www.moralmachine.net/hl/de>.

¹⁵ Siehe Straßenverkehrs-Ordnung (StVO) § 1 Grundregeln, https://www.gesetze-im-internet.de/stvo_2013/__1.html.

deutlich reduziert. Im Fall „Cruise“ haben sogar zwei Personen einen Fehler gemacht: Der Fahrer des Nissan und die Fußgängerin als eigentliches Opfer. Daraus ergibt sich eine Anforderung an die künstliche Intelligenz eines autonomen Fahrzeugs, die Gasser als „Fehlerkompensationsfähigkeit“ bezeichnet (Gasser 2015: 560). Er beschreibt damit ein Verhalten einer maschinellen Steuerung, die beim Auftreten von Fehlern anderer Verkehrsteilnehmer geeignete Maßnahmen zur Gefahrenreduktion einleitet. Im Fall „Cruise“ führte dieses Verhalten zu einem Nothalt des Robotertaxis, was im Nachhinein eine richtige Entscheidung war. Leider wurde dann die Folgeentscheidung getroffen, eine sichere Position anzufahren. Dieses Ziel war zuvor im Programm als „rechter Fahrbahnrand“ definiert worden und diese Entscheidung führte zur Verletzung der Fußgängerin. Im Fall „Telsa“ hat diese Logik, wenn sie überhaupt vom Hersteller umgesetzt wurde, nicht zu rechtzeitigen und erfolgreichen Maßnahmen geführt. Darüber hinaus müssen auch die schwächeren Verkehrsteilnehmerinnen und Verkehrsteilnehmer in die Abwägung der „Handlungsoptionen“ einbezogen werden, was auf der Grundlage des Grundgesetzes geboten ist:

Soweit sich ergibt, dass die maschinelle Umfelderkennung ausreicht, um nichtmotorisierte Verkehrsteilnehmer (Fußgänger und Radfahrer) als solche zu erkennen, wäre ihr Schutz im Rahmen sich ergebender Handlungsvarianten besonders zu berücksichtigen. (Gasser 2015:558)

Um die Entscheidungen einer künstlichen Intelligenz zu programmieren, könnte es hilfreich sein zu verstehen, wie ein Mensch bei einem Unfall über Handlungsoptionen nachdenkt. In Extremsituationen würde der Mensch in Sekundenbruchteilen verschiedene Alternativen moralisch abwägen – so beantworten zumindest Nida-Rümelin und Weidenfeld diese Frage. Der Mensch sei dabei in der Lage, „unter extremen Stress und Zeitnot Entscheidungen zu treffen, auch wenn für die verbale Formulierung von Gründen und Gegengründen keine Zeit bleibt“ (Nida-Rümelin & Weidenfeld 2020:92). Dazu werden dem Menschen bestimmte Alternativen visuell vor Augen geführt, zwischen denen er sich entscheidet. Im Gegensatz dazu muss ein Computersystem auf der Grundlage eines zuvor programmierten Optimierungsalgorithmus eine mathematische Auswahl treffen. Mit den zur Verfügung stehenden Daten und unter Annahme von Wahrscheinlichkeiten könnte eine Auswahl getroffen werden, die die positiven „Konsequenzen seines Handelns maximiert. Digitaler Utilitarismus sozusagen“ (Nida-Rümelin & Weidenfeld 2020:67).

Ein Optimierungsalgorithmus könnte auch verschiedene Werte oder sogar Leben miteinander verrechnen. Nida-Rümelin und Weidenfeld warnen und sprechen von einer „ethischen Nichtverrechenbarkeit“, „weder gegen andere Werte, zum Beispiel ökonomische Vorteile, noch gegen andere Leben“ (Nida-Rümelin & Weidenfeld 2020:93). Dieses sei der Kern der Verfassung der Bundesrepublik Deutschland und ließe sich auch in anderen demokratischen Ordnungen finden. Diese „ethische Nichtverrechenbarkeit“ sei aus ihrer Sicht nur zu erreichen, wenn Dilemmasituationen akzeptiert werden und im Falle eines Auftretens sich die Akteure „zwangsläufig mit Schuld beladen“ (Nida-Rümelin & Weidenfeld 2020:94). Zur Diskussion, ob die Zeit, die für eine Entscheidung benötigt wird (Mensch in Sekundenbruchteilen vs. Entwickler mit Jahren an Entwicklungszeit) einen Unterschied macht, argumentiert Wolkenstein, dass dies nicht der Fall sei (2018). Wolkenstein ist als wissenschaftlicher Mitarbeiter an der Ludwig-Maximilians-Universität München am Institut für Ethik, Geschichte und Theorie der Medizin tätig. Wolkenstein argumentiert, dass die Entscheidung des Menschen auf einer „history of moral education“ (2018:169) beruhe. So wie einem Kind moralisches Handeln anezogen werde, könne sich auch ein Algorithmus im Laufe der Zeit „entwickeln“.

Ein weiterer Diskussionspunkt betrifft die Relevanz des Trolley-Problems. Als Argument mag genannt werden, dass solche Situationen wie das Trolley-Problem nur ein Gedankenexperiment seien und die autonomen Fahrzeuge diese verhindern würden. Forschungen kamen dagegen zu dem Ergebnis, dass eine kollisionsfreie Umgebung für die Robotik und damit auch für autonome Fahrzeuge unrealistisch ist (Goodall 2014, Lin 2015). Eine ähnliche Kritik an dem Gedankenexperiment wird von Tom Michael Gasser, der in der Bundesanstalt für Straßenwesen tätig ist, vertreten: Diese Dilemma-Situation ist aus seiner Sicht im Straßenverkehr eine sehr seltene Ausnahmesituation. Vielmehr würden möglicherweise die Gefahren nicht nur in der Fahrzeugsteuerung zu verorten sein, sondern sie würden dem „Verkehrssystem Straße“ generell innewohnen (Gasser 2015:557).

Aus Sicht von Wolkenstein sollten die Gedankenexperimente wie das Trolley-Problem nicht dazu dienen, konkrete Antworten durch die Auswertung der Ergebnisse zu erhalten, sondern vielmehr dazu genutzt werden, ein tieferes Verständnis des menschlichen Moralsystems zu erforschen (2018:170). Für Wolkenstein hat das Trolley-Problem nämlich eine erschreckende moralische Logik

offenbart, die sich im Laufe der Zeit im Menschen entwickelt hat (Wolkenstein 2018). Als Beispiel nennt er Entscheidungen im Gesundheitswesen, wo in einer Notfallsituation ggf. eine Ersteinschätzung (Triage) vorgenommen werden muss, die dazu führen kann, dass andere Personen erst später behandelt werden und dadurch gesundheitliche Nachteile erleiden könnten. Deshalb werden aus seiner Sicht auch autonome Fahrzeuge vor diesem Dilemma stehen. Aus diesem Grund sei weitere Forschung notwendig, um die menschliche Ethik besser zu verstehen. Die Frage sei dann, wie diese Erkenntnisse in die Entwicklung der künstlichen Intelligenz für autonome Fahrzeuge einfließen könnten.

Neben theoretischen Überlegungen kennen erfahrene Fahrerinnen und Fahrer unzählige kritische Situationen aus eigener Beobachtung oder Erfahrung, wie zum Beispiel Wildunfälle, vor denen an kritischen Stellen sogar mit speziellen Verkehrszeichen gewarnt wird. In den USA ereignen sich jährlich mehr als eine Million Wildunfälle (Lin 2015:74), so dass diese Gefahr auch für ein autonomes Fahrzeug nicht unwahrscheinlich ist. Wie oben dargestellt, muss ein Fahrer oder eine Fahrerinnen in Sekundenbruchteilen die richtige Fahrentscheidung treffen. Ein autonomes Fahrzeug mit künstlicher Intelligenz könnte hingegen durch ein „durchdachtes Entscheidungsfindungsskript“ (Lin 2015:74) den Vorteil haben, in kurzer Zeit optimal zu reagieren.

Dabei muss die künstliche Intelligenz drei verschiedene Fragen beantworten bzw. Stufen in der Entscheidungsfindung durchgehen. Erstens muss geprüft werden, ob das System in der Lage ist, die Verantwortung an den Fahrer zu übergeben. In vielen Unfallszenarien fehlt leider die Zeit, um die Kontrolle an den Fahrer oder die Fahrerinnen zu übergeben. Laut Lin dauert es je nach Ablenkung bis zu 40 Sekunden (Lin 2015:72), bis die Kontrolle über das Fahrzeug übernommen werden kann. Diese Zeit ist deutlich länger als die ein bis zwei Sekunden, die als Reaktionszeit benötigt werden. Wenn die Rückgabe nicht möglich ist, muss geprüft werden, inwieweit eine Bremsung oder eine Vollbremsung möglich wäre. In dieser Entscheidung zur Bremsung müssen Parameter wie die Fahrbahnbeschaffenheit oder die konkrete Situation im rückwärtigen Verkehr bedacht werden. Denn besonders schwere LKWs haben in der Vergangenheit bei Auffahrunfällen zu erheblichen Personenschäden geführt.

Zweitens muss das Objekt klassifiziert werden. Handelt es sich um einen Menschen, ein Tier oder etwas anderes? Die künstliche Intelligenz muss hier eine gute Wahl

treffen. Bei großen Tieren würde es bei einer Kollision zu Risiken für das Fahrzeug und die Insassen führen. Möglicherweise wäre es bei kleineren Tieren sinnvoller, die Route unverändert zu lassen und die Verletzung des Tieres in Kauf zu nehmen. Allerdings haben beispielsweise Haustiere „special places in our hearts“ (Lin 2015:74), so dass ein menschlicher Fahrer oder das autonome System diese Tiere unter Umständen schützen würde bzw. sollte.

Als dritter Punkt ist zu klären, ob die Ausweichmöglichkeit nach links oder nach rechts, ggf. durch eine Bremsung, genutzt werden könnte. Ein Ausweichen nach rechts könnte in einem Graben oder an einem Baum enden. Ein Ausweichen nach links könnte zu einem Frontalzusammenstoß mit dem Gegenverkehr führen. Darüber hinaus beeinflussen die Anzahl der Fahrzeuginsassen und die Art der zusätzlichen Ladung die mögliche Schwere des Unfalls.

Anhand des Beispiels mit dem Wildunfall stellt Lin die Frage, ob das Bremsen der richtige Ansatz ist, um ethische Probleme zu lösen: „[...] it is already unclear that braking should be the safest default option – as a proxy for the most ethical option [...]“ (Lin 2015:75). Selbst wenn Umfelderkennung und künstliche Intelligenz in der Lage wären, alle oben genannten Schritte richtig zu beantworten, wäre die aktuell bevorzugte Bremsung nicht unbedingt die beste Entscheidung. Im Fall „Cruise“ hat das Robotertaxi vor der Fußgängerin einen Nothalt durchgeführt, also genau so, wie es Lin kritisiert hat. Auch hier wäre es evtl. möglich gewesen, ein Ausweichen nach links in den Gegenverkehr oder nach rechts auf den ersten Fahrstreifen in Betracht zu ziehen.

In der weiteren Argumentation sieht Lin eine geringere Schuld bei einem menschlichen Fahrer, wenn in Sekundenbruchteilen und mit begrenzten Informationen Fehlentscheidungen getroffen werden, um die ethisch richtige oder am wenigsten schädliche Entscheidung zu treffen. Auf der anderen Seite sieht Lin aufgrund des Zeitvorteils die Verantwortung bei den Herstellern und Entwicklern von autonomen Fahrzeugen:

[...] programmer and OEM do not operate under the sanctuary of reasonable instincts; they make potentially life-and-death decisions under no truly urgent time-constraint and therefore incur the responsibility of making better decisions than human drivers reacting reflexively in surprise situations. (Lin 2015:75)

In der Milderung der Folgen einer Kollision und dem Abwägen von verschiedenen Handlungsfolgen sieht Lin die ethische Herausforderung. Dazu nutzt er den Begriff „Targeting“ Algorithmus (LIN 2015:73). Dieser Begriff wird auch im militärischen

Bereich zur Auswahl von Zielen verwendet und passt auch für die Zielberechnung für die autonomen Fahrzeuge.

Oben wurde bereits das Argument von Wolkenstein angeführt, dass der Zeitfaktor zwischen Mensch und Maschine bei der Klärung der Verantwortung keine Rolle spielte. Diese Linie steht im Gegensatz zu der Forderung von Lin, die Hersteller hätten aufgrund der zur Verfügung stehenden Entwicklungszeit eine größere Verantwortung, um in einem autonomen Fahrzeug bessere Entscheidungen zu implementieren.

Fraichard und Asama definieren den Begriff „unvermeidlicher Kollisionszustand“ (inevitable collision state) für ein Robotersystem, in dem es, unabhängig von der zukünftigen Trajektorie, zu einer Kollision mit einem ggf. unbekanntem Objekt kommt. Im Umkehrschluss sollte bei der Bahnberechnung jeweils im Voraus sichergestellt werden, dass das Fahrzeug nicht in einen „unvermeidlichen Kollisionszustand“ gebracht wird (Goodall 2014:3).

Ein Rechner kann in Sekundenbruchteilen Handlungsoptionen durchspielen und dann den Weg wählen, der mit dem geringsten Risiko behaftet ist. Dabei ist eine Notbremsung z.B. auf der Autobahn nicht immer die richtige Lösung. Manchmal ist es besser, eine Kombination aus Bremsen und Ausweichen oder sogar Beschleunigen vorzunehmen. Dieses müsste in die Steuerungssoftware implementiert werden. Das bedeutet, die Steuerungssoftware bzw. die künstliche Intelligenz entscheidet ethisch (Nyholm 2018, Goodal 2014).

4.3.3 Die Debatte aus der Sicht ausgewählter Stakeholder

In diesem Kapitel werden die Standpunkte ausgewählter Stakeholder zusammengefasst. Zunächst wird die Sicht der zuständigen Behörden und der Unfallforschung dargestellt. Anschließend werden die Einflussmöglichkeiten eines weiteren wichtigen Akteurs, nämlich der Versicherungswirtschaft, dargestellt.

Seit den 1970er-Jahren wurden verschiedene Maßnahmen ergriffen, um die Anzahl der tödlich verletzten Verkehrsteilnehmenden zu verringern. Zu den Maßnahmen gehören nach Winkle (2015) beispielsweise Änderungen im Straßenbau, Verbesserungen der Sicherheitsausstattung in den Fahrzeugen und in der Rettungskette sowie in der Notfallmedizin. Trotz der steigenden Anzahl von Fahrzeugzulassungen konnte die Anzahl der Verkehrstoten reduziert werden. Auf der anderen Seite sind die Zahlen nach wie vor hoch und die dahinterstehenden

Einzelschicksale lassen sich nur erahnen. Laut Unfallstatistik ereignet sich nach Winkle im Durchschnitt ein tödlicher Verkehrsunfall:

alle 2,7 Stunden in Deutschland

alle 25 Minuten (ca. 34.000 im Jahr) in den USA und

alle 26 Sekunden (mindestens 1.240.000 jährlich) weltweit (Winkle 2015:354).

Bei Statistiken neueren Datums lässt sich noch keine Trendumkehr erkennen. Es sind zum Beispiel im Jahr 2022 nach den Schätzungen der NHTSA (2022) in den USA bis zu 40.000 jährliche Unfallopfer ums Leben gekommen. Unfallforscher können für die Erforschung der Unfallursachen auf Unfalldatenbanken¹⁶ zugreifen. Leider stehen diese Daten nicht immer der Öffentlichkeit zur Verfügung. Die statistische Unfallursachenverteilung nach Winkle ist in Abbildung 14 dargestellt. Diese Verteilung macht deutlich, wie häufig derzeit menschliches Versagen zu Unfällen führt. Unklar ist jedoch, wie sich die Situation in Zukunft entwickeln wird. Denn derzeit ist nicht absehbar, wie sich technische Fehler als neue Fehlerquelle auf die Unfallstatistik auswirken werden.

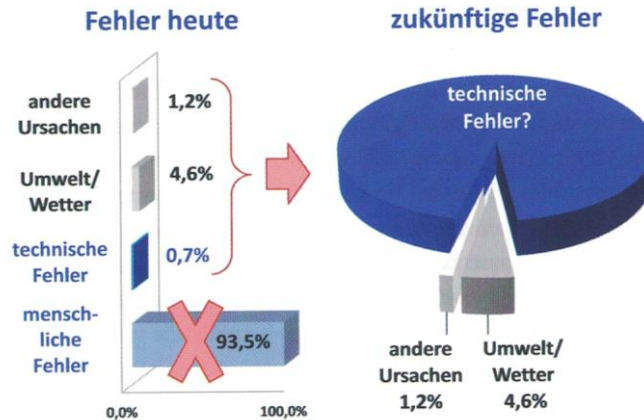


Abbildung 14: Statistische Unfallursachenverteilung (Winkle 2015:396)

Tom Michael Gasser ist bei der Bundesanstalt für Straßenwesen tätig und nimmt in seiner Arbeit (2015) Bezug auf die Rechtsfragen für autonome Fahrzeuge. Passend zur Unfallursachenverteilung formuliert er den Begriff des „Automatisierungsrisikos“ (Gasser 2015:552), das verbleibt, wenn eine maschinelle Fahrzeugsteuerung

¹⁶ Wie zum Beispiel www.gidas.org.

eingeführt wird. Dieses Risiko sei aus seiner Sicht jedoch nicht größer als bei einer menschlichen Fahrzeugführung.

In verschiedenen Studien konnte nachgewiesen werden, dass menschliches Fehlverhalten in den meisten Fällen die Ursache für den Unfall gewesen ist (Smith 2013). Die National Highway Traffic Safety Administration fand in einer Studie 2015 heraus, dass sich 94% der Unfälle auf menschliche Fehler zurückführen lassen. Diese Fehler können in drei Teilgruppen unterteilt werden: In einer gestörten Aufmerksamkeit, in einer falschen Entscheidung und nicht zuletzt in einer zu geringen Fähigkeit zum Beherrschen des Fahrzeuges oder den aktuellen geforderten Fahrfähigkeiten. Laut NHTSA war in 41 Prozent der Fälle eine zu geringe Aufmerksamkeit des Fahrers bzw. der Fahrerin die häufigste Ursache für einen Unfall (NHTSA 2015:2).

Wie oben genannt wurde, liegt die Hauptursache der Unfälle in einem Fehlverhalten des Menschen (siehe NHTSA 2015). Jedoch kommen neue Sicherheitsprobleme durch selbstfahrende Fahrzeuge hinzu. Wie beim Fall Tesla ersichtlich, hat das System technisch versagt. Da die selbstfahrenden Fahrzeuge über Internet mit einem Server verbunden sind, sind sie auch anfällig für Hackerangriffe. Somit müssen die Hersteller während der Entwicklung die Informationssicherheit (oder engl. Cybersecurity) berücksichtigen. Diese „nicht-traditionellen Sicherheitsprobleme“ haben sich laut Kriebitz, Max und Lütge (2022:6) zu einem neuen Forschungsgebiet entwickelt.

Nach einer Verordnung müssen Unfälle mit automatisierten Fahrfunktionen gemeldet werden. Hier kann ein Anstieg der Unfälle mit Tesla Fahrzeugen festgestellt werden, insbesondere seitdem der Autopilot auch in Städten und Wohngebieten aktiviert werden kann (Siddiqui & Merrill 2023:8). Laut einer Analyse von Thadani, Lerman, Piper, Siddiqui und Uraizee (2023:3) wurden 2014 mehr als 700 Unfälle seit der Einführung des Tesla Autopiloten registriert, davon zumindest 19 mit schwerwiegenden Folgen. Auf der anderen Seite müssen die Unfallzahlen mit der absoluten Anzahl von schwerwiegenden Unfällen betrachtet werden. Die National Highway Traffic Safety Administration schätzt in ihrem Bericht für das Jahr 2022, dass „42,795 people died in motor vehicle traffic crashes“ (NHTSA 2022).

Die Fahrzeughersteller bringen immer weitere Fahrassistenten und Selbstfahr-Funktionen auf dem Markt und der Weg zu einem vollautomatischen Fahrzeug schreitet immer weiter voran. Es stellt sich die Frage, ob diese neuartigen Funktionen zu einer Verbesserung der Verkehrssicherheit beitragen. Es ist ferner unklar, ob die Öffentlichkeit die Bereitschaft zeigen wird, sich von einer künstlichen Intelligenz fahren zu lassen.

Ein wichtiger Stakeholder in der Betrachtung und Etablierung von autonomen Fahrfunktionen ist die Versicherungswirtschaft. Da jedes Fahrzeug – zumindest in Deutschland – eine Haftpflichtversicherung besitzen muss, haben die Versicherungen u.a. durch die Preisgestaltung ihrer Policen eine Lenkungsmöglichkeit. Mit besonders günstigen Prämien könnten sie den ergänzenden autonomen Funktionen zu einem Durchbruch verhelfen. Es wäre eine einfache kaufmännische Rechnung, dass sich z.B. die Mehrkosten für eine Zusatzausstattung durch ein KI-System durch eine geringere Versicherungspolice nach einiger Zeit amortisieren. Auf der anderen Seite könnte sich eine Versicherung zu höheren Prämien entschließen, da die Reparaturen eines Fahrzeugs mit ADAS-Ausstattung deutlich teurer werden könnten. Dies lässt sich leicht nachvollziehen durch die zahlreichen, kostspieligen Sensoren, die bei einem Crash möglicherweise ersetzt werden müssten. Eine höhere Prämie wäre bei dieser Argumentation ein zusätzliches Argument gegen ein Fahrzeug mit ADAS-Komfortfunktionen.

Die Versicherer stehen neben der Risikobewertung vor Fragen der ethischen und gesellschaftlichen Bewertung. Könnte beispielsweise eine ADAS-Zusatzfunktion ein Leben retten? Eine andere Frage könnte sein, ob Menschen durch neuartige Versicherungsmodelle diskriminiert werden.

Um die Entwicklung von Fahrzeugen der Stufe SAE Level 3 weiter voranzutreiben, werden Stück für Stück die Erkenntnisse aus dem Bereich der Künstlichen Intelligenz in dem Fahrzeugbau integriert. Poszler und Geißlinger (2021) nennen die drei Themengebiete der Objekt-Erkennung, die Vorhersage des Verhaltens anderer Verkehrsteilnehmenden und die Entscheidungsfindung als wesentliche Gebiete, in denen die KI schon eingesetzt wird (Poszler & Geißlinger 2021:2). Aus ihrer Sicht schreitet die Entwicklung so schnell voran, dass ethische Überlegungen „hinterherhinken“ (:2).

4.3.4 Debatte mit Fokus „Tesla Autopilot“

Die National Highway Traffic Safety Administration hat nach einem Update der Tesla Steuerungssoftware die Unfälle mit Tesla Fahrzeugen noch einmal untersucht. Dabei äußert die Behörde Bedenken bzgl. der Zeit für die Rückübertragung der Fahr-Verantwortung an den Fahrer bzw. die Fahrerin. Nach ihrer Meinung könnte ein aufmerksamer Fahrer bzw. Fahrerin die Unfälle nicht in jedem Fall vermeiden, jedoch bliebe Zeit für das Abschwächen der Folgen eines Unfalls:

In more than half (59) of these crashes, the hazard was visible five or more seconds prior to the impact, with a subset of 19 exhibiting a hazard visible for over 10 seconds prior to the collision. For events unfolding faster, such as those where the hazard may have first been seen less than two seconds prior to the crash, an attentive driver's timely actions could have mitigated the severity of a crash even if the driver may not have been able to avoid the crash altogether. (NHTSA 2024:4)

Die automatischen Fahrfunktionen in einem Tesla werden mit dem Produktnamen Autopilot angeboten (Tesla 2024:87ff). Darunter wird eine Reihe von Funktionen zusammengefasst, die das Fahren sicherer und stressfreier machen soll. Diese Erweiterung wird laut Handbuch in allen neuen Modellen der Tesla Modellpalette integriert (:87). Die Dokumentation weist den Fahrer, die Fahrerin darauf hin, dass das System immer zu überwachen sei und dass der Nutzer jederzeit in der Lage sein muss, die Kontrolle über das Fahrzeug zu übernehmen. Damit sieht sich Tesla selbst nach dem SAE-Standard als Level 2 Anbieter, d.h. der Fahrer und die Fahrerin tragen die Verantwortung, auch wenn der Autopilot aktiviert wurde.

Neben der Basis-Funktionalität werden Full-Self-Driving-Funktionen als Betaversion angeboten. Für das maschinelle Sehen bzw. die Umfelderkennung stützt sich Tesla beim Modell 3 auf Videokameras ab. Insgesamt sind in dem Modell sieben Kameras verbaut (:18). Dazu kommt noch Kartenmaterial und ein Global Positioning System (GPS) zum Einsatz. Dabei ist der Fahrer oder die Fahrerin auch hier verpflichtet, das Kartenmaterial auf dem aktuellen Stand zu halten. Im Handbuch gibt Tesla zahlreiche Einschränkungen des Systems an: Das System kann möglicherweise andere Verkehrsteilnehmer übersehen, wenn das Fahrzeug mit einer Geschwindigkeit über 50 mph (80 km/h) fährt und ein anderer Verkehrsteilnehmer zu stehen gekommen ist oder sehr langsam fährt (:116). Ferner besteht die Möglichkeit, dass der Tesla auf Objekte oder Fahrzeuge reagiert, die auf der Fahrbahn gar nicht vorhanden sind. Dies kann zu einer unnötigen oder unangemessenen Verlangsamung des Fahrzeugs führen (:116). Einschränkungen hat das System laut

Hersteller zum einen bei schlechten Sichtverhältnissen (z.B. starker Regen oder schlecht beleuchtete Fahrbahn bei Nacht) und zum anderen bei starkem Lichteinfall (z.B. durch entgegenkommende Fahrzeuge oder direkte Sonneneinstrahlung). Zur Leistungsfähigkeit des Systems gibt es unterschiedliche Ansichten. Im Jahr 2016 zitierte die Washington Post den Tesla-Chef Elon Musk (CEO) zur Self-Driving Funktionalität der Tesla Modelle:

[...] he already believes the technology has proved itself, even though Tesla introduced the self-driving software into its fleet of vehicles only late last year and even that didn't turn the vehicles into fully autonomous vehicles. (Frankel 2016).

Zudem gab Musk zu dem Zeitpunkt einen Ausblick in die Zukunft, denn in den nächsten 24 bis 36 Monaten „Tesla will be able to drive virtually all roads at a safety level significantly better than humans“ (Frankel 2016).

In einem Artikel von Siddiqui und Merrill (2023) wird auf eine erhöhte Anzahl von Unfällen mit dem Fahrerassistenzsystem von Tesla eingegangen. Laut Meinung von Tesla würde die Einführung mit dem Argument verteidigt, dass der Nutzen den Schaden überwiegt (Siddiqui & Merrill 2023:2). Und dann geht Elon Musk in seiner Argumentation noch einen Schritt weiter und sieht sich moralisch verpflichtet, die Funktion auszurollen, auch wenn er dafür angeklagt oder sogar verklagt wird:

At the point of which you believe that adding autonomy reduces injury and death, I think you have a moral obligation to deploy it even though you're going to get sued and blamed by a lot of people. [...] Because the people whose lives you saved don't know that their lives were saved. And the people who do occasionally die or get injured, they definitely know — or their state does. (Siddiqui & Merrill 2023:3)

Philip Koopman, ein international bekannter Experte mit mehr als 25 Jahren Erfahrung im Bereich der Sicherheit selbstfahrender Autos, kritisiert die Funktion des Tesla Autopiloten (AP) als zu unzureichend umgesetzt:

The safety issue with AP (and Tesla) in a nutshell is summarized as: "A comparison of Tesla's design choices to those of L2 peers identified Tesla as an industry outlier in its approach to L2 technology by mismatching a weak driver engagement system with Autopilot's permissive operating capabilities." (Koopman 2024)

Nach der Perspektive aus der Sicht des Herstellers wechselt nun die Betrachtung in die Ebene der Nutzer. Das eigentliche Unfallopfer ist der Fahrer des Tesla Model 3. Der Fahrer wurde uns als technikbegeisterter Mensch vorgestellt, so wie seine Frau im Untersuchungsbericht dargestellt hat. Aus den Daten lässt sich nicht ermitteln, wie

die Frau und die Familie mit dem Verlust ihres Ehemannes und Vaters umgehen können. Die Frau gab an, ihr Mann „hätte die Werbefilme für Tesla gesehen“. Vermutlich war der Mann von den in den Werbefilmen gezeigten Möglichkeiten so begeistert, dass er den Umfang des automatisierten Fahrens im Fahrzeug nicht hinterfragte bzw. mit seinem Laienwissen nicht nachvollziehen konnte. Ein Verschulden liegt beim Unfallfahrer selbst: Zum einen hat er auf der Strecke eine überhöhte Geschwindigkeit eingestellt, zum anderen war das autonome Fahren für diesen Streckenabschnitt nicht freigegeben, konnte aber trotzdem im Auto aktiviert werden. Im Normalfall ohne autonomes Fahren hätte er wahrscheinlich eine Strafe wegen Geschwindigkeitsüberschreitung mit Unfallfolge bekommen. Im konkreten Fall war die automatische Fahrfunktion verantwortlich. In dieser kurzen Übernahmezeit hätte das Fahrzeug die Gefahrensituation erkennen und bei Unlösbarkeit die Verantwortung an den Fahrer zurückgeben müssen. Aufgrund der Unfallanalyse ist dieses jedoch nicht geschehen. Weder der Fahrer noch das Automatik-System hatten Maßnahmen zur Unfallverhinderung eingeleitet. Die Fehler der Technik nutzt laut Thadani, Lerman, Piper, Siddiqui und Uraizee (2023) auch der Anwalt der Opfer-Familie und argumentiert, die Technologie hätte wiederholt versagt, als sie nicht bremste oder als sie keine Warnung vor dem Sattelschlepper gegeben habe.

Ein weiterer Stakeholder ist der LKW-Fahrer, der an dem Unfall beteiligt war. Der Berufskraftfahrer nutzte den Verkehrsraum, um seiner vertraglichen Verpflichtung als Fahrer nachzukommen. Wie aus dem Untersuchungsbericht ersichtlich ist, wurden die Fahrer regelmäßig medizinisch untersucht. Der Fahrer ist zum Unfallzeitpunkt medizinisch behandelt, so dass eine Teilnahme am Straßenverkehr von den Behörden freigegeben wurde. Dem Fahrer unterlief morgens ein kleiner Fahrfehler, d.h. an der Kreuzung hätte er an einem Stoppschild kurz anhalten müssen. Die Untersuchung ergab, dass er das Fahrzeug verlangsamte und dann nach Überprüfung der Straße wieder beschleunigte, um die Kreuzung zu passieren. Mit etwas mehr Umsicht, vielleicht bei besserer Gesundheit, hätte er den Unfall verhindern können. Von weiteren Umständen wurde nichts berichtet.

Der Untersuchungsbericht NTSB (2020:15) fasst die wahrscheinlichen Ursachen für den Unfall folgendermaßen zusammen: Als erster Punkt wird das fehlerhafte Verhalten des LKW-Fahrers genannt, der dem Tesla an der Kreuzung die Vorfahrt hätte

gewähren müssen. Als zweiter Punkt wird die Funktionsweise des Tesla Autopiloten genannt, der für einen gewissen Zeitbereich eine Abwesenheit des Unfallfahrers erlaubte. Und zuletzt noch die Versäumnisse des Unternehmens, den Einsatz des Autopiloten auf die Bedingungen zu beschränken, für die er entwickelt wurde. Erschwerend kommt hinzu, dass der National Highway Traffic Safety Administration (NHTSA) keine Verifikationsmethode zum Testen von Level-2-Fähigkeiten zur Verfügung stand (NTSB 2020:15). Laut NTSB wurde deshalb eine Verbesserungsempfehlung mit der Kennung „H-17-38“ an das NHTSA gerichtet, um diesen Umstand zu verbessern (:13). Dieser hatte zum Report-Zeitpunkt den Status „Open – Unacceptable Response“ (:13).

4.3.5 Debatte mit Fokus „Cruise“ Robotertaxi

In der Fallstudie „Cruise“ ist eine Fußgängerin das Opfer. Innerhalb der Großstadt San Francisco hatte sie regelwidrig eine Straße überquert und war dabei von zwei Fahrzeugen erfasst worden. Das erste Auto wurde traditionell von einem Fahrer gesteuert und erfasste die Fußgängerin laut Untersuchungsbericht frontal. Danach wurde sie auf die zweite Fahrspur geschleudert und wurde dort zusätzlich durch das Robotertaxi verletzt.

Die Leitung der Firma Cruise ging recht offen mit dem Unfall um. Sie veranlasste eine Untersuchung, veröffentlichte den Bericht (Exponent 2023) und übernahm die Verantwortung für den Fall. In der Untersuchung wurde das Verhalten der Mitarbeitenden am Unfalltag untersucht. Über die Leitstelle wurden sie informiert, dass das Fahrzeug einen Gegenstand in Fahrtrichtung festgestellt hatte und dann die Parkposition eingenommen hatte. Sie hatten zu dem Zeitpunkt keine Information darüber, dass eine Fußgängerin durch das Fahrzeug verletzt wurde.

Der Stadtrat von San Francisco hatte dem Testbetrieb von Robotertaxi zugestimmt. Aufgrund des Verkehrsunfalls und wahrscheinlich auch aufgrund der öffentlichen Debatte in den Medien (Thadani 2023b) hatte der Stadtrat die Erlaubnis wieder zurückgezogen.

5 THEOLOGISCHE GRUNDLEGUNG

Die theologische Ethik spielt in dieser Arbeit eine besondere Rolle. Ich beziehe mich in dieser Untersuchung vor allem auf die Verantwortungsethiken von Dietrich Bonhoeffer und H. Richard Niebuhr, die im Folgenden dargestellt und diskutiert werden. Die Darstellung wird ergänzt durch das von Emil Brunner formulierte christliche Menschenbild, das implizit in eine Situationsbeurteilung einfließen kann.

Als Ergänzung sei zu Beginn gesagt, dass Bonhoeffer und Niebuhr nicht als die einzigen relevanten Befürworter für ein ethisches Konzept der Verantwortung angesehen werden. Angesichts der Debatte um Risiken in der technologischen Entwicklung zog Hans Jonas mit seinem Buch „Das Prinzip Verantwortung: Versuch einer Ethik für die technologische Zivilisation“ (Jonas 1979) die Aufmerksamkeit auf sich. Für eine weitere Lektüre sei hier auf die Arbeit von Jonas und Rezipienten wie de Villiers (2006) verwiesen.

5.1 Bonhoeffers Konzept der Verantwortung

In diesem Absatz wird Bonhoeffers Verständnis von Verantwortung vorgestellt. Zunächst wird die beeindruckende Biografie Bonhoeffers in Teilen nachgezeichnet. Um die Relevanz seines Werkes für die heutige Zeit zu unterstreichen, werden ausgewählte zeitgenössische Autoren, die sich auf Bonhoeffer beziehen, zitiert. Anschließend wird der Begriff der Verantwortung unter Einbeziehung aktueller Beiträge diskutiert.

5.1.1 Biografie von Dietrich Bonhoeffer

In diesem Kapitel wird eine biografische Skizze Bonhoeffers gegeben, um die Einflüsse auf seinen Entwurf der Verantwortungsethik besser einordnen zu können. Dietrich Bonhoeffer wurde am 4. Februar 1906 in Breslau geboren¹⁷. Er war das sechste von insgesamt acht Kindern in einer bürgerlichen Intellektuellenfamilie. Ab dem sechsten Lebensjahr lebte die Familie in Berlin, da seinem Vater Karl Bonhoeffer die Leitung der dort ansässigen Universitäts-Nervenklinik übertragen wurde. Die Erziehung im Hause Bonhoeffer war geprägt von einem harmonischen und

¹⁷ Die wesentlichen biografischen Eckpunkte zur Biografie Bonhoeffers wurden aus dem Dietrich-Bonhoeffer-Portal (Rosenau 2024) entnommen sowie aus der Biografie von Eberhard Bethge (Bethge 2017).

erfüllten Miteinander. Wie Eberhard Bethge in der Biografie schreibt, lag das Hauptziel der Erziehung darin, „die Kinder zu verantwortlichen Menschen heranzubilden“ (Bethge 2017:14). Von den Geschwistern wurde erwartet, dass sie auf die Gefühle und Bedürfnisse der anderen Rücksicht nahmen. Und so wurde aus Sicht von Bethge dieses „Bedenken-der-anderen“ eine wichtige Komponente in Bonhoeffers Theologie (:14).

Nachdem er mit siebzehn Jahren das Abitur bestanden hatte, begann er ein Theologiestudium. Das Theologische Examen konnte er im Winter 1927/28 abschließen. Danach reichte er seine Doktorarbeit ein, die im Jahr 1930 veröffentlicht wurde. Sein beruflicher Werdegang war für diese Zeit durch Auslandsaufenthalte in Spanien, England und sogar in den USA geprägt. Die sogenannte Bekennende Kirche in Deutschland bat Bonhoeffer 1935, ein Predigerseminar in Finkenwalde zu leiten und Bonhoeffer nahm die Anfrage an. Zunächst konnte das Seminar öffentlich stattfinden. Nachdem der Staat die Aktivitäten stark behinderte, wurde das Seminar im Untergrund weitergeführt. Für die Teilnehmer und für Bonhoeffer selbst wurde diese Zeit als besonders fruchtbringend erlebt. In einem Brief spricht Bonhoeffer später von seiner „beruflich und menschlich ausgefülltesten Zeit“ (:62). Auf dem Seminar lernte er auch seinen Freund Eberhard Bethge kennen, der bei der Verbreitung der Arbeit zur Ethik später eine Schlüsselrolle einnehmen sollte.

Die „Ethik“ vom Bonhoeffer wurde erst posthum veröffentlicht. Sein Schwager, Hans von Dohnanyi, gehörte zur „Abwehr“ und hatte Dietrich Bonhoeffer eine Tätigkeit in derselben vermittelt. Der offizielle Auftrag bestand darin, Spionagetätigkeiten gegen das Dritte Reich aufzu-decken. Gleichzeitig waren jedoch die führenden Männer laut Zimmerling (Zimmerling 2021:7) konspirativ gegen Hitler aktiv. Aus diesem Grund wurde Bonhoeffer zu einem Doppelagenten:

Seine theologischen Überlegungen mussten sich dadurch noch einmal ganz neu und anders als bisher bewähren, nämlich außerhalb von Theologie und Kirche in einem weltlichen Tätigkeitsbereich auf dem Feld des gesellschafts-politischen Engagements, im aktiven Widerstand gegen den National-sozialismus. (:8)

Seine Arbeit an der „Ethik“ wurde nach dem missglückten Anschlag auf Hitler durch seine Verhaftung gestoppt. Sein Freund Bethge fand Teile seiner Arbeit zur Ethik am Tag nach der Verhaftung noch auf seinem Schreibtisch liegen und rettete somit die Gedanken Bonhoeffers. Das Manuskript überstand die Kriegszeit. Nach dem Krieg

wurden die einzelnen Teilarbeiten, Notizen und Entwürfe 1949 als Buch veröffentlicht (:9-11).

5.1.2 Rezeption des Bonhoefferschen Verantwortungsbegriffs

Die Relevanz von Bonhoeffers Konzept der Verantwortung wird durch die Tatsache unterstrichen, dass es für zeitgenössische Autoren immer noch von Bedeutung ist. Hierzu sollen beispielhaft drei Autoren und deren Arbeiten genannt werden.

Als erstes soll auf den Theologen Wolfgang Huber eingegangen werden. In seinem Buch „Menschen, Götter und Maschinen. Eine Ethik der Digitalisierung“ nimmt er Bezug auf die Verantwortungsethik Bonhoeffers. Huber schreibt, dass im Gegensatz zu einem Ende der Geschichte, das Gott vorbehalten ist, die Verantwortung des Menschen innerhalb der geschichtlichen Zeit liegt (Huber 2022:36). Der Mensch müsse auf dem Weg zu neuen Möglichkeiten „Chancen und Risiken gleichermaßen in den Blick nehmen“ (Huber 2022:36). Bonhoeffer hat hierzu die Unterscheidung zwischen den „letzten und den vorletzten Dingen“ (Bonhoeffer 2023:137ff) geprägt und dies ist auch für Huber von maßgeblicher Bedeutung.

Als zweiter Autor ist der in Belgien lehrende van den Heuvel zu nennen. Er ist Professor für Systematische Theologie an der ETF (Evangelisch Theologische Fakultät) in Leuven (van den Heuvel 2024b). Seine Thesen zu „'Who is Jesus Christ for us Today?' Bonhoeffer's Ethics of Responsibility as a Promising Paradigm for Contemporary Public Theology“ hat van den Heuvel in der Eröffnungsfeier für das Studienjahr 2024/2025 in Leuven vorgestellt (van den Heuvel 2024).

Als drittes Beispiel soll die für diese Arbeit interessante Forschung von Verhaegen „Integrating Theological Ethics of Responsibility with Teleological and Deontological Ethics for Embodied AI“ (Verhaegen 2021) genannt sein. Seine Arbeit wurde von van den Heuvel begleitet.

Im folgenden Kapitel werden die Eckpunkte des Bonhoefferschen Verantwortungsbegriffs vorgestellt., die van den Heuvel treffend als „Matrix of responsibility“ (van den Heuvel 2024:13) bezeichnet hat.

5.1.3 Matrix der Verantwortung

Bonhoeffers Bezugspunkt aller Betrachtungen zum ethischen Handeln ist die Anerkennung der Wirklichkeit des Schöpfers. In der Menschwerdung seines Sohnes Jesus Christus hat er sich den Menschen offenbart. Diese „Christus-Zentriertheit“ wird deutlich, wenn Bonhoeffer schreibt:

Nicht Ideale, Programme, nicht Gewissen, Pflicht, Verantwortung, Tugend, sondern ganz allein die vollkommene Liebe Gottes vermag der Wirklichkeit zu begegnen und sie zu überwinden. (Bonhoeffer 2021:69)

Aufgrund Jesu Erlösungstat kann der Mensch „befreit“ für den anderen handeln. Der Mensch ist von Gott angenommen. Der Theologe Christoph Fleischer sagt dazu, dass wir Gott durch unser Leben „Antwort geben“:

Ich stehe zugleich für Christus vor den Menschen und für die Menschen vor Gott. Der Begriff der Verantwortung ist hier also nicht allgemein ethisch definiert, sondern theologisch: Dieses Leben als Antwort auf das Leben Jesu Christi nennen wir ‚Verantwortung‘. (Fleischer 2009)

Laut van den Heuvel (2024:13) arbeitet Bonhoeffer das Konzept der Verantwortung am deutlichsten in seinem Manuskript „Die Geschichte und das Gute“ (Bonhoeffer 2021:245-256) heraus. Bonhoeffer schreibt, dass sich das verantwortliche Leben durch eine doppelte Struktur beschreiben lässt: Da ist zunächst die Bindung des Lebens an den Menschen und an Gott. Und auf der anderen Seite besteht die Freiheit des eigenen Lebens (:256). Schwerpunkt der Arbeit von Bonhoeffer ist laut van den Heuvel weder die kirchliche Ethik noch die damals vorherrschende Ethik von Immanuel Kant, sondern die Frage, wie die christliche Ethik zum Gemeinwohl beitragen kann („responsibility for the common good“, van den Heuvel 2024:12). Darüber hinaus stellt van den Heuvel fest, dass in der damaligen einzigartigen Situation die bisherigen ethischen Ansätze nicht geeignet waren, die Zusammenhänge angemessen zu erfassen (:13).

Als erste Dimension der Matrix der Verantwortung steht die Stellvertretung. Was ist damit gemeint? Als erste Aktivität der Stellvertretung kann die Erlösungstat von Jesus Christus für den Menschen angesehen werden. Daraus wird im zweiten Schritt eine Ausrichtung des Lebens des Menschen auf Christus und dies kann so zu einem stellvertretenden Leben werden, das sich für den Nächsten einsetzen kann. In „Die Geschichte und das Gute“ bringt Bonhoeffer (2023:245-256) den Zusammenhang

zwischen Verantwortung und Stellvertretung auf dem Punkt, wenn er schreibt, dass jemand „real an die Stelle eines anderen tritt“:

Dass Verantwortung auf Stellvertretung beruht, geht am deutlichsten aus jenen Verhältnissen hervor, in denen der Mensch unmittelbar genötigt ist, an der Stelle anderer Menschen zu handeln, also etwa als Vater, als Staatsmann, als Lehrmeister. Der Vater handelt an der Stelle der Kinder, indem er für sie arbeitet, für sie sorgt, eintritt, kämpft, leidet. Er tritt damit real an ihre Stelle. (Bonhoeffer 2021:151)

Es lässt sich aus heutiger Sicht nachspüren, wie Bonhoeffer durch sein eigenes Leben und Wirken dem Ruf zur Verantwortung gefolgt ist und wie er die Stellvertretung „in vollkommener Hingabe des eigenen Lebens an den anderen Menschen“ (:258) umgesetzt hat. Verglichen mit dem heute wahrnehmbaren ausgeprägten Individualismus sagt Bonhoeffer, dass selbst der einsam lebende Mensch sich nicht der Verantwortung und das heißt der Stellvertretung entziehen kann. Und dies unabhängig davon, in welchem Umfang Verantwortung getragen wird, als Einzelner, für eine Gemeinschaft oder für eine ganze Gemeinschaftsgruppe (:157). Darüber hinaus ist es eine Kunst, weder sich selbst noch einen anderen Menschen zu verabsolutieren, wie von der Heuvel mit Bezug auf Bonhoeffer (:259) zusammenfasst:

[...] it is the art of neither absolutizing yourself or the other person - because through that absolutization, responsibility is lost. Instead, responsibility is the careful, difficult negotiation between justified interests, both of myself and those of others. (van den Heuvel 2024:14)

Die Übernahme der Verantwortung muss sich nach Bonhoeffer an die Wirklichkeit anpassen. Hierzu führt Bonhoeffer den Begriff der „Wirklichkeitsgemäßheit“ ein und dies ist ein weiterer Aspekt in der Matrix der Verantwortung. Bonhoeffer betont die Auseinandersetzung mit der Realität, wie sie ist und er beschreibt das Verhalten eines Verantwortlichen so:

Sein Verhalten liegt nicht von vornherein und ein für alle Mal, also prinzipiell fest, sondern es entsteht mit der gegebenen Situation. Er hat kein absolut gültiges Prinzip zur Verfügung, das er fanatisch gegen jeden Widerstand der Wirklichkeit durchzuführen hätte, sondern er sucht das in der gegebenen Situation Notwendige, „Gebotene“ zu erfassen und zu tun. (Bonhoeffer 2021:260) Auch an dieser Stelle wird Bonhoeffers Zentrierung auf Christus deutlich, wenn er sagt, „die Wirklichkeit ohne den Wirklichen verstehen zu wollen, bedeutet in einer Abstraktion leben, [...]“, bedeutet Vorbeileben an der Wirklichkeit (:153). Nach Fleischer entfaltet Bonhoeffer keinen ständig geltenden Anspruch des Sollens, sondern eine situationsabhängige

Entscheidung, die sich auch als „christologisch orientierte Situationsethik“ (Fleischer 2009) beschreiben lässt:

Weil es nicht um die Durchführung irgendeines grenzenlosen Prinzips geht, darum muss in der gegebenen Situation beobachtet, abgewogen, gewertet, entschieden werden, alles in der Begrenzung menschlicher Erkenntnis überhaupt. Es muss der Blick in die nächste Zukunft gewagt, es müssen die Folgen des Handelns ernstlich bedacht werden, ebenso wie eine Prüfung der eigenen Motive, des eigenen Herzens, versucht werden muss. Nicht die Welt aus den Angeln zu heben, sondern am gegebenen Ort das im Blick auf die Wirklichkeit Notwendige zu tun, kann die Aufgabe sein. (Bonhoeffer 2021:161)

Ein Handeln in dieser Wirklichkeit macht sich „die Erfahrung vieler Generationen errungene Weisheit“ (:272) zunutze. Aus Bonhoeffers Sicht verachtet kein Staatsmann ungestraft diese Gesetzmäßigkeiten. Auf der anderen Seite schränkt Bonhoeffer den „Verantwortungshorizont“ ein. Aus seiner Sicht müsste der Mensch sich seine eigene Endlichkeit („Geschöpflichkeit“) bewusst machen und es wäre nicht die Aufgabe, „die Welt zu überspringen und aus ihr das Reich Gottes zu machen“ (:156). Um ein korrektes Handeln in der Wirklichkeit zu ermöglichen, muss der und die Handelnde ein Kontextwissen haben bzw. aufbauen. Bonhoeffer spricht von einem „Wesensgesetz“, das jeder Sache innewohnt und es sei gleichgültig, „[...] ob es sich um eine vorgefundene Naturgegebenheit oder um ein Erzeugnis des menschlichen Geistes, ob es sich um eine materielle oder ideelle Größe handelt“. (:164)

Huber richtet sein Augenmerk in der Verantwortungsethik auf den zeitlichen Verlauf der Ereignisse. Das Ende der Geschichte sei Gott vorbehalten, die Menschen müssten innerhalb der geschichtlichen Zeit die Verantwortung übernehmen, um das „Gegebene zu bewahren und Neues zu ermöglichen“ (Huber 2022:36). Für Huber ist es wichtig, Chancen und Risiken gleichermaßen in den Blick zu nehmen. Wie oben von Bonhoeffer zitiert, muss ein „Blick in die nächste Zukunft gewagt“ (Bonhoeffer 2021:161) werden. Huber argumentiert, bei „ethischen Überlegungen geht es stets um einen Übergang vom Schlechteren zum Besseren“ (Huber 2022:38). Dazu bedarf es einer größeren Zahl von möglichen Optionen und die Nutzung neuer Möglichkeiten müssten an dem verantwortungsethischen Prinzip der „vorausschauenden Vorsicht – das precautionary principle“ gebunden werden (:39).

Ein weiterer Punkt zur Matrix der Verantwortung ist das Gewissen. Damit wird die Instanz zur eigenen Reflexion von Entscheidungen umschrieben. Nach Reuter gehen in das auf eigene Handlungen bezogene reflexive Selbstbewusstsein unterschiedliche Elemente ein: Im allgemeinen Sprachgebrauch wird hiermit die Unterscheidung von richtigen und falschen Handlungen durch das moralische Selbstbewusstsein beschrieben. Damit können Handlungen in der konkreten Situation, in der Vergangenheit oder in der Zukunft gemeint sein. Hier dazu zwei prominente Beispiele. Im 1. Korintherbrief nimmt Paulus die Personen in Schutz, die ein rigides Normbewusstsein haben und Götzenopferfleisch nicht essen wollen. Ein anderes bekanntes Beispiel ist die Verteidigungsrede von Luther 1521 in Worms, als er sich „auf sein in Gottes Wort gefangenes Gewissen berief, weil gegen das Gewissen zu handeln weder sicher noch heilsam ist“ (Reuter 2015:77).

Eine zweite zentrale Aufgabe des Gewissens ist nach Reuter die „Bewahrung der personalen Identität und Integrität“ (:77). Auf eine unbestechliche Weise meldet das „Identitätsbewusstsein“ einer Person, dass sie sich durch ihr Handeln mit sich selbst entzweit. Diese Argumentation nimmt auch Bonhoeffer in seinen Ansatz der Verantwortungsethik mit hinein:

Die Verantwortung für den Nächsten hat hier ihre Grenze in der Unantastbarkeit des Gewissensrufes. Eine Verantwortung, die zu einem Handeln wider das Gewissen zwingt, würde sich selbst verurteilen. (Bonhoeffer 2021:169)

Ein drittes Element des Gewissens ist die „Fähigkeit zum fallbezogenen Urteilen und Entscheiden“. Im Gegensatz zu dem untrüglichen Urteil über sich selbst kann sich nach Reuter das situative Gewissensurteil auch irren. Wichtig an dieser Stelle sei die Feststellung, dass „das Gewissen nicht per se die Stimme Gottes in uns“ sei (Reuter 2015:78).

Nachdem oben die Kriterien Stellvertretung, Wirklichkeitsgemäßheit und Gewissen vorgestellt wurden, folgt nun als vierter und letzter Punkt in der Matrix der Verantwortung die Bereitschaft zur Schuldübernahme. Für Bonhoeffer nimmt ein Verantwortlicher oder eine Verantwortliche Schuld auf sich und steht für diese ein. Dabei hofft die verantwortliche Person in der Beziehung zu Gott auf Gnade und auf ein freisprechendes Gewissen:

Wer in Verantwortung Schuld auf sich nimmt – und kein Verantwortlicher kann dem entgehen –, der rechnet sich selbst und keinem anderen diese Schuld zu

und steht für sie ein, verantwortet sie. Er tut es nicht in dem frevelnden Übermut seiner Macht, sondern in der Erkenntnis zu dieser Freiheit genötigt und in ihr auf Gnade angewiesen zu sein. (Bonhoeffer 2021:175)

Die nachfolgende Skizze (siehe Abbildung 15) gibt einen grafischen Überblick über die von Bonhoeffer entwickelte Ethik.

Ethik Bonhoeffers:

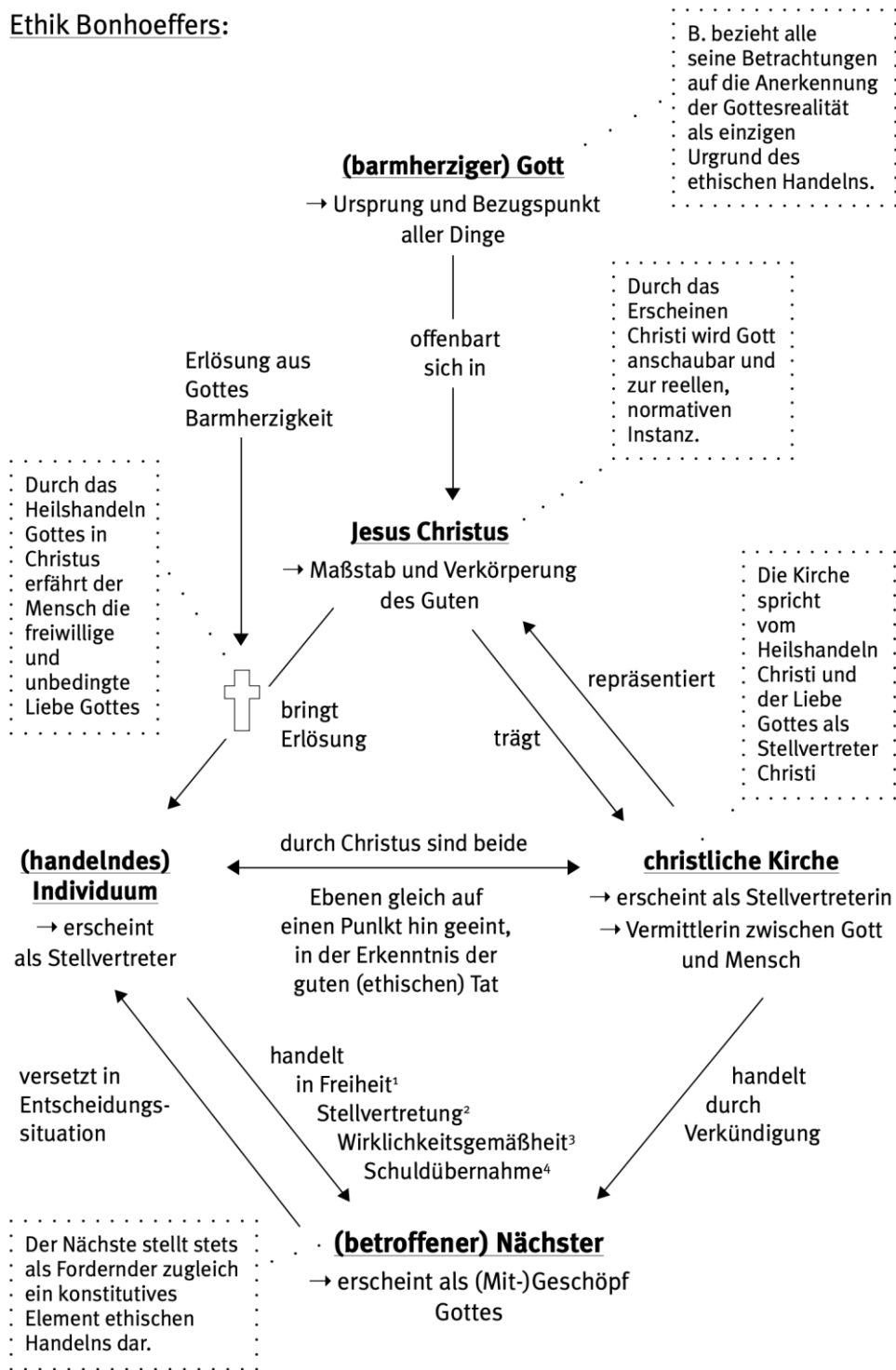


Abbildung 15: Die Ethik Bonhoeffers (Biewald & Beckmann 2007)

5.1.4 Zwischenfazit zu Bonhoeffer

Im Unterschied zu anderen Ethikkonzepten sieht Bonhoeffer den Menschen in der alleinigen Bindung zu Gott zur gehorsamen und verantwortlichen Tat berufen. Es ist eine Ethik, die nach dem Willen Gottes fragt, der zur Nähe der Weltwirklichkeit interpretiert werden muss. Es ist eine christologisch orientierte Situationsethik. Laut Fleischer „darf (sie) allerdings nicht als Gesinnungsethik angesehen werden, sondern versteht sich selbst in diesem Zusammenhang als Verantwortungsethik“ (Fleischer 2009:22). Den Willen Gottes zu erkennen, ist jedoch nicht leicht und nach Fleischer hält die christliche Ethik die Frage offen, „wie menschliches Handeln als Entsprechung des Willens Gottes aussehen kann, ohne den Schwerpunkt auf die urteilende Kritik, sondern eher auf das vorbildliche Handeln zu legen“ (:22). Ein geeignetes Wissen über den Kontext, Bonhoeffer nennt es in seiner Ethik das Wesensgesetz, das einer Sache innewohnt, ist für ein erfolgreiches Handeln notwendig. Offen bleibt die Frage, wie verantwortliches Handeln nach dem Willen Gottes aussehen könnte. Bonhoeffer ist davon überzeugt, dass Gott einem den Wunsch eröffnen wird und macht seinen Lesern und Leserinnen Mut, wenn er schreibt, „dass Gott dem, der ihn demütig fragt, seinen Willen gewiss zu erkennen gibt“ (Bonhoeffer 2021:326).

5.2 Niebuhrs Konzept der Verantwortung

In diesem Kapitel wird das Verständnis zur Verantwortung von H. Richard Niebuhr vorgestellt. Zunächst wird sein Leben kurz vorgestellt. Anschließend wird der zentrale Ansatz von Niebuhr diskutiert und durch aktuelle Beiträge ergänzt.

5.2.1 Biografie von H. Richard Niebuhr

Georg Kalinna ist in seiner Arbeit (Kalinna 2021) auf die Verantwortungsethik von H. Richard Niebuhr eingegangen. In seinem Buch „Der Mensch als antwortendes Wesen – Gedanken zur gegenwärtigen Verantwortungsethik“ (Kalinna 2021) wurde eine erstmalige Veröffentlichung der Dankesrede Niebuhrs anlässlich der Verleihung der Ehrendoktorwürde der Bonner Universität im Jahr 1959 vorgenommen. Sie ist eine von mehreren Versionen der Gedanken von Niebuhr, die „über Jahre im Lehrbetrieb gereift, von ihm jedoch noch nicht in eine Buchform gebracht worden sind“ (:33). Bis zu seinem frühen Tod am 5. Juli 1962 lehrte H. Richard Niebuhr an der Yale Divinity School den „Course in Christian Ethics“. Die wesentlichen

Gedanken von Niebuhr wurden in dem posthum erschienenen Buch „The Responsible Self“ (Niebuhr 1963) zusammengefasst.

Im Geleitwort zu Kalinna bemerkt der Soziologe Hans Joas einen „der merkwürdigsten Rezeptionsdefizite“ (Kalina 2021:7) und das weitgehende Desinteresse an den Arbeiten der Brüder Niebuhr in Deutschland. Dabei gehören die Brüder Reinhold (1892–1971) und H. Richard Niebuhr (1894–1962) aus amerikanischer Sicht zu den wichtigsten Theologen des zwanzigsten Jahrhunderts (:1). Dies sei umso mehr erstaunlich, da Niebuhrs aufgrund ihrer Herkunft und Sprachkenntnisse für den „deutsch-amerikanischen intellektuellen Dialog geradezu prädestiniert waren“¹⁸. Beispielhaft für das geringe Interesse soll an dieser Stelle das Grundlagenbuch von Reuter im „Handbuch der Evangelischen Ethik“ genannt sein, in dem die Verantwortungsethik von Niebuhr nur mit einem kurzen Verweis (2015:80) erwähnt wird.

Da H. Richard Niebuhrs Arbeit in englischer Sprache verfasst wurde, ist es wichtig zu verstehen, wie die verwendeten Begrifflichkeiten in ihrem Kontext verwendet wurden. Der englische Ausdruck „responsible“ wird meist verwendet, wenn man etwas als „gut“ oder „gerecht“ bezeichnen möchte. In seinem Vortrag „Der Sinn der Verantwortlichkeit“ weist Niebuhr darauf hin, dass „responsible“ früher im Sinne von „entsprechend“ (2021:43) verwendet wurde. Seit dem 19. Jahrhundert scheint es gebräuchlich geworden zu sein, „verantwortlich sein“ als „gehorsam oder auf ein Ziel ausgerichtet sein“ zu definieren (:43).

5.2.2 Zentraler verantwortungsethischer Ansatz

Für Niebuhr fällt die christliche Ethik in zwei Themengebiete. Zum einen geht es um das aktive und ein bibelorientiertes Leben eines Menschen vor Gott:

[...] den Dienst im Hören und im Gehorsam zu den Geboten, im Eingehen und Halten von Versprechen, im Trachten nach ewigen Gütern, in der Antwort auf Gottes schöpferische, richtende und erlösende Taten. (:37)

Zum anderen behandelt die christliche Ethik das Leben des heutigen Menschen mit den Fragen: „Was soll ich tun?“ und „Warum und wie soll ich es tun?“ (Niebuhr 2021:37). Aus der Sicht von Niebuhr sind beide Pole relevant und mit Blick auf

¹⁸ Ihre Sprachkenntnisse hatten sie durch ihr Elternhaus, einem deutsch-amerikanischen Pfarrhaus, erworben.

heutige „Einsatzmöglichkeiten“ kann man sich nicht „ganz ausschließlich auf die Bibel oder auf das christliche Bewusstsein stützen“ (:37).

Die Frage „Was soll ich tun?“ lässt sich nach Niebuhr nur dann beantworten, wenn zuvor gefragt wurde „Was ist das letzte Ziel meines Lebens?“ Und hier würde die Schwierigkeit in der Deutung des höchsten Gutes liegen. Nach Niebuhr hat man schon oft versucht zu zeigen, „dass die Bibel die Ethik des Trachtens nach einem höchsten Gut herausstelle“ (:39). Nach seiner Sicht enthält die Bibel in der Tat Spuren dieser Ethik, jedoch ist sie viel mehr als ein „Ethos des Trachtens“. Für ihn ist die Bibel eher als eschatologisches und weniger als teleologisches Buch zu sehen – sie habe „Gottes Ziel und weniger das Ziel des Menschen im Auge“ (:39).

Für den Christen von heute ist nach Niebuhr das Wichtigste, nach „den Zielen, Wünschen und dem Willen dessen, der unser Schicksal bestimmt“ (:39) zu fragen. Der Mensch dürfe eigene, kurzfristige Ziele auswählen und hätte dadurch auch ein gewisses Maß an Freiheit, um diese Ziele zu erreichen. Es steht nach Niebuhr jedoch nicht in der Macht des Menschen, „das Reich Gottes auf Erden zu schaffen, Gott zu verherrlichen, unsere eigene Seligkeit oder Vollkommenheit zu erreichen“ (:39).

In der „Verwirklichung der modernen Existenz“ (:41) muss der Mensch unzählige Entscheidungen vollziehen. In dieses System passt nach Niebuhr nicht der Ansatz der Kasuistik, d.h. dass für alle Fälle ein Leitsystem mit Regeln vorliegt. Ein anderer Ansatz ist die Verbindung der Intuition mit dem Willen Gottes. Nach Niebuhr wissen Vertreter des „Intuitionismus“ jeweils ganz genau, „was Gott in jedem Augenblick von uns fordert“ (:42). Niebuhr sieht diese beiden Methoden als unzureichend und nicht hilfreich an,

[...] denn sie entnimmt ja das handelnde Ich ganz beträchtlich aus den Geschichtszusammenhängen und aus der unmittelbaren, sozialen, zeitbegrenzten Situation des vor Gott lebenden Selbst (:42).

Eine prägnante Vorgehensweise von Niebuhr ist es, die Gesetzmäßigkeiten zur Verantwortung mithilfe von Metaphern zu beschreiben. Eine Metapher zur Erklärung der teleologischen Theorie ist der Baumeister oder engl. „man-the-maker“ (Kalinna 2021:56). Damit ist ein Mensch gemeint, der sein Handeln vorher entwirft und dann durchführt. Eine weitere Metapher für das Verstehen der Deontologie ist der Bürger

oder die Bürgerin oder wie Niebuhr es nannte „man-the-citizen“, also jemand, der oder die zur Befolgung oder dem Erlass von Gesetzen in Erscheinung tritt.

Zum teleologischen Handlungsverständnis macht Niebuhr drei Kritikpunkte geltend. Als erstes richtet sich das teleologische Konzept auf ein Ziel in der Zukunft. Für Niebuhr ist der entscheidende Moment jedoch die Gegenwart und diese sei weit mehr als eine „Durchgangsstation zu einem zu erreichenden zukünftigen Ziel“ (Kalinna 2021:56). Die Gegenwart sei der „kritische Augenblick, der alle Folgen für die Zukunft in sich trägt“ (Niebuhr 2021:39). Ein zweiter Kritikpunkt an der teleologischen Theorie ist die Entscheidung, ob man in der konkreten Situation das Lassen oder das Eingreifen wählen sollte. Oft sei das Handeln des Menschen „eine situationsgebundene Notlösung“ (:39). Als dritten Kritikpunkt hält Niebuhr das teleologische Handlungsverständnis im Bezug auf den biblischen Ethos für unzureichend:

Es gehe den biblischen Traditionen nicht so sehr darum, ein wie auch immer geartetes höchstes Gut oder Ziel zu erstreben, sondern nach einem Ziel zu fragen, das dem Menschen *extra se* gegenübersteht: dem Ziel, das Gott für den Menschen hat (Kalinna 2021:56).

Auch am deontologischen Konzept übt Niebuhr in dreifacher Weise Kritik (siehe Kalinna 2021:56ff). Erstens wird die Einhaltung von Rechten und Pflichten dem Wesen menschlichen Handelns nicht gerecht. In einer jeweiligen Situation würden dem Menschen die Möglichkeiten für „Entscheidungsspielräume“ und Spontanität zur Verfügung stehen. Eine zweite Kritik bezieht sich auf die „soziale Dimension menschlichen Handelns“. Damit meint Niebuhr beispielsweise die Anforderungen, denen Menschen an ihren konkreten Plätzen wie z.B. im Beruf ausgesetzt sind. Der Rückgriff auf ein von der sozialen Situation abgekoppeltes Gesetz ist darum aus der Sicht von Niebuhr nicht hilfreich. Und drittens sei für Niebuhr „die Gebotsethik zu formalistisch“ (Kalinna 2021:57).

Niebuhr unterteilt die Verantwortung in vier Elemente und beginnt mit dem ersten Element, der Tatsache, dass das „Ich antwortet“. Wenn der Mensch handelt, so steht er in einer „Wechselbeziehung“. Persönliches Handeln ist für Niebuhr „antwortendes Handeln“, d.h. „es ist die Reaktion auf eine Aktion“ (Niebuhr 2021:43). Der Mensch steht „inmitten eines naturgegebenen und gesellschaftlichen Spannungsfeldes“ (:44), in dem auf den Menschen eingewirkt wird oder wo er abgestoßen wird. Damit kann

der Mensch bei Niebuhr als „responsives“ (d.h. antwortendes) Wesen angesehen werden.

Das zweite Element ist die Abhängigkeit der Antwort von der „Deutung der Handlung, die es erleidet“. Das heißt, eine Handlung wird gedeutet und durchdacht und mit anderen Handlungen verglichen. Das Ziel ist, eine sinnvolle Handlung zu erkennen. Niebuhr schreibt, dass der Mensch die Geschehnisse „als Teil eines größeren Vorganges oder als Glieder einer Kette von Ereignissen“ deutet. Und als Konsequenz dieser Deutung wird unsere Antwort („response“) beeinflusst. Dieses Beobachten und Deuten lasse sich zwischen Nationen, in Firmen oder auch in der Familie erkennen. Niebuhr bringt als Beispiel die Formung des Charakters eines Kindes. Dieser wird weniger durch die Erziehung der Eltern geformt „als durch die Deutung, die das Kind der Haltung der Eltern gibt“ (zum Beispiel Annahme oder Ablehnung) (:47). Neben dem Leben in der Gesellschaft deutet der Mensch auch die Ereignisse in der Natur, wie Hitze und Kälte oder Gesundheit und Krankheit. Der Mensch ordnet nach Niebuhr seine Beobachtungen in verschiedene Kategorien ein, z.B. ob ein Ereignis ein Zufall sei, sich auf ein maschinelles Funktionieren abstützt oder ob auch Gnade im Spiel war. Diese Deutungen kann der Mensch nur teilweise bewusst kontrollieren, darauf weist Niebuhr hin. In der Verantwortung versucht der Mensch die Frage nach dem „Was soll ich tun?“ zu beantworten, indem er zunächst die Frage nach dem „Was geht hier vor sich?“ oder „Was erleide ich?“ zu beantworten versucht.

Ein drittes Element ist für Niebuhr das „Rechenschaft-Geben“, damit meint er, der Mensch handelt in „Erwartung einer Antwort“. Das Bild aus dem juristischen Kontext kann hier zum Verständnis dienen. Dieses Element schaut auf der Zeitachse rückwärts oder auch vorwärts in Erwartung von Einwänden oder Korrekturen. Der oder die Handelnde steht in der Verantwortung, wenn er oder sie mögliche Reaktionen in Kauf nimmt und steht so in einer „Wechselwirkung“. Ferner können die zur Verfügung stehenden „Zeitspannen“ für eine Handlung oder Gegenhandlung Einfluss nehmen auf die Antwort des Menschen.

Ein viertes Element verantwortlichen Handelns ist das Eingebundensein in eine Gemeinschaft. Die Verantwortung und die Existenz in der Gemeinschaft sind für

Niebuhr in vielerlei Hinsicht miteinander verknüpft. Gemeinschaft ist dabei nicht nur die direkte Beziehung zum Gegenüber. Sie kann auch dadurch entstehen, dass gemeinsam „Mächte und Ereignisse“ erlebt werden, die zusammen auf die Gemeinschaft einwirken. Der oder die Handelnde kann Dinge nur unter Berücksichtigung einer gemeinsamen Begrifflichkeit erkennen. Diese Begrifflichkeit, so Niebuhr, hat der oder die Handelnde aus der Gemeinschaft, d.h. von den „Mitwissenden“ (Niebuhr 2021:50).

Zusammenfassend gesagt, definiert Niebuhr die Verantwortlichkeit (responsibility) als eine der Situation angemessene Reaktion:

The idea of an agent's action as response to an action upon him in accordance with his interpretation of the latter action and with his expectation of response to his response; and all of this is in a continuing community of agents. (Niebuhr 1963:65)

5.2.3 Verortung der Theorie Niebuhrs

Responsivität ist für Niebuhr ein Merkmal der Verantwortlichkeit. Damit, so Kalinna, ist Niebuhr nicht der Einzige, der sich als Repräsentant theologischer Ethik der Responsivität des menschlichen Daseins bezeichnet. Zu nennen wären die bis heute bedeutenden Beiträge von Bonhoeffer und auch Karl Barth. Bonhoeffer schreibt in seiner Ethik zum Thema Responsivität Folgendes:

Wir leben, indem wir auf das in Jesus Christus an uns gerichtete Wort Gottes Antwort geben. Weil es ein auf unser ganzes Leben gerichtetes Wort ist, darum kann auch die Antwort nur eine ganze, mit dem ganzen Leben, wie es sich jeweils handelnd realisiert, gegeben sein. (Bonhoeffer 2023:253)

Wie lässt sich Niebuhrs Theorie von Bonhoeffer abgrenzen oder haben beide identische Konzepte entwickelt? Am obigen Beispiel lässt sich die offenkundige „christozentrische Ausrichtung von Bonhoeffers Verantwortungsbegriff“ (Kalinna 2021:64) erkennen. Auf der anderen Seite eröffnen die Überlegungen Niebuhrs in Bezug zum menschlichen Handeln die Möglichkeit, für eine nichtchristliche Betrachtungsweise „anschlussfähig“ (:64) zu bleiben. Und so kommt Kalinna zu der Feststellung, dass Niebuhr „eine eigenständige responsivitätstheoretisch geprägte Form von Verantwortungsethik entworfen“ (:64) hat.

5.3 Brunners Begriff der Verantwortlichkeit

Bezugnehmend auf die Verantwortungsmatrix nach Bonhoeffer soll in diesem Kapitel auf einem weiteren Aspekt eingegangen werden, der einen Einfluss auf die Bewertung einer konkreten Entscheidungssituation hat: das Menschenbild. Denn in der Interaktion zwischen Handelnden und betroffenen Mitmenschen (siehe dazu auch Abbildung 15) steht unterschwellig die Frage, welchen „Wert“ dem Mitmenschen als Gegenüber beigemessen wird. In Rahmen dieser Arbeit soll die Definition von Emil Brunner vorgestellt werden.

Emil Brunner (1889–1966) war Professor für Systematische und Praktische Theologie an der Universität in Zürich. Er veröffentlichte 1937 die erste selbständige theologische Anthropologie (Kessler 2025). Brunner konkretisiert das klassische Bild von der Gottesebenbildlichkeit des Menschen („imago Dei“) zu einem Verständnis von „Gerufen sein“ und „Antwort geben“. Dazu schreibt er:

Dass Gott die Kreatur ins Dasein „ruft“ (1. Mos 1, Röm 14,17, 2 Kor 4,6), trifft einzig beim Menschen im wörtlichen Sinne zu. Die Art dieser doppelseitigen Bezogenheit heißt von Gott aus „Anruf“, vom Menschen aus „Antwort“. Das Sein des Menschen ist so nach seinem Kern verstanden: verantwortliches Sein. (Brunner 1958:22)

Durch das „verantwortliche Sein“ ergibt sich für Brunner eine Substanz des Menschseins:

In der christlichen Lehre vom Menschen geht es um die wahre Erkenntnis des verantwortlichen Seins. Wer das Wesen der Verantwortlichkeit verstanden hat, der hat das Wesen des Menschen verstanden. Die Verantwortlichkeit ist nicht ein Attribut; sie ist die Substanz des Menschseins [...]. So ist auch das Wissen um Verantwortlichkeit das, was jeden Menschen zum Menschen macht. (Brunner 1958:14)

So lässt sich Brunners Begriff der Verantwortlichkeit folgendermaßen zusammenfassen: Zum einen ist „das Wissen um die Verantwortlichkeit das, was jeden Menschen zum Menschen macht“ (Brunner 1937:39). Und zum anderen ist „wahre Verantwortlichkeit dasselbe wie die wahre Menschlichkeit“ (:40).

5.4 Schlussfolgerung

In diesem Kapitel wurde die Verantwortungsethik nach Bonhoeffer und Niebuhr vorgestellt. Dabei wurde deutlich, dass die Theorien sich nicht klar abgrenzen lassen und die beiden wichtigen Vertreter Unterschiede, aber auch Gemeinsamkeiten aufzeigen.

Für Bonhoeffer ist die Verantwortung ein „Leben als Antwort auf das Leben Jesu Christi“ (Bonhoeffer 2023:254). Ferner geht es Bonhoeffer um die „ganzheitliche Verpflichtung, um eine Bindung des ganzen Lebens an den in Jesus Christus offenbarten Gott“ (Kalinna 2021:64):

Ich verantworte unter Einsatz des Lebens mit Worten das, was durch Jesus Christus geschehen ist. Ich verantworte also primär nicht mich selbst, mein Tun, ich rechtfertige mich nicht selbst, sondern ich verantworte Jesus Christus und damit allerdings auch den mir von ihm gewordenen Auftrag. (Bonhoeffer 2023:255)

Nur Bonhoeffer geht auf die Schuldhaftigkeit des Menschen ein. Dieser Punkt ist besonders bemerkenswert und hilfreich für eine ethische Entscheidungsfindung.

Huber (2022:36) sieht eine Verbindung zwischen Bonhoeffers „Letztem und Vorletztem“ und der „responsorischen Verantwortung“ von Niebuhr. Nach Huber kommt in der responsorischen Verantwortung „zur Geltung, dass sich das menschliche Leben in Beziehungen vollzieht“ (:36). Hierbei haben die Beziehungen zu Gott, zur Lebenswelt, zu den Mitmenschen und zu sich selbst laut Huber eine hervorgehobene Bedeutung.

Und nicht zuletzt wurde in diesem Kapitel aufgezeigt, wie das angenommene Menschenbild als implizit beeinflussendes Element den Handelnden bzw. die Handelnde in der Rolle der Stellvertretung bzw. in Bezug auf die responsorische Verantwortung Einfluss nimmt. Das Menschenbild kann somit als Basis oder Fundament für die ethischen Theorien gesehen werden.

6 THEOLOGISCH-ETHISCHE ENTSCHEIDUNGSFINDUNG

Nachdem im vorangegangenen Kapitel das Konzept der Verantwortungsmatrix vorgestellt wurde, soll im Folgenden dieser Ansatz genutzt werden, um eine theologisch-ethische Entscheidungsfindung in der Bearbeitung der Kernfragen zu wagen. Zunächst werden die Stakeholder vorgestellt. Es folgt die Benennung der moralischen Frage und die Diskussion der wesentlichen Argumente.

6.1 Kontextsensibilität und Wirklichkeitsgemäßheit

Für die Analyse der moralischen Situation braucht es eine Offenheit zum Erfassen einer Situation. Bonhoeffer nannte es das Wesensgesetz (Bonhoeffer 2021:164), das es zu erkennen gilt. Die vorherigen Kapitel in dieser Arbeit sind dazu gedacht, einen Kontext aufzubauen. Zu dieser Kontextsensibilität gehört auch das Sammeln der beteiligten Personen und Interessengruppen. In diesem Zusammenhang wird auch gerne die engl. Bezeichnung Stakeholder verwendet (Bleisch et.al. 2021:32), ein aus der Betriebswirtschaftslehre¹⁹ entlehnter Begriff.

Stakeholder können sich unterschiedlich verhalten: Sie können direkt beteiligt sein oder sich auch in einer Organisation zusammengeschlossen haben. Nach Ellmann, Behrend, Hübner und Weitlinger können „Stakeholder betroffen sein, Einfluss oder Macht ausüben, sie können Promotoren oder auch potenzielle Konfliktverursacher (Opponenten)“ sein (Ellmann, Behrend, Hübner & Weitlinger 2012:77). Im Allgemeinen versucht jede Interessengruppe, ihre eigenen Wünsche durchzusetzen. Es wäre unfair, Personen oder Gruppen nicht zu Wort kommen zu lassen, die von einer bestimmten Handlung oder Regelung betroffen sind (Bleisch et al. 2021:33).

¹⁹ Seitens eines Unternehmens sind die Stakeholder für den Erfolg des Unternehmens von besonderer Bedeutung. Rohrschneider und Spang (2012:133) geben folgende Beispiele von Stakeholdern an:

Kunden, wie z.B. Auftraggeber, Nutzer und Betreiber, Geldgeber.
Mitarbeiter, wie z.B. Projektleiter und -mitarbeiter, Vorgesetzte, Controller.
Eigentümer des Unternehmens, in dem das Projekt realisiert wird.
Zulieferer, wie z.B. Lieferanten, Dienstleister, Berater, Versicherer.
Gesellschaft, wie z.B. Menschen und Institutionen aus dem Umfeld, Behörden, Politik, Überwachungsorgane und Medien.

Im Fall „Tesla Autopilot“ lassen sich die Interessengruppen wie in Tabelle 1 dargestellt unterteilen.

Interessengruppe	Fall Tesla
Kunden	der Fahrer des Fahrzeugs Model 3, in diesem Fall gleichzeitig das Unfallopfer (und weitere Kunden)
Eigentümer des Unternehmens	Die verantwortlichen Leiter der Firma (Elon Musk), die Mehrheitsanteilseigner
Mitarbeiter des Unternehmens	Die zahlreichen Mitarbeiter aus den verschiedenen Abteilungen, neben der Entwicklung beispielsweise auch aus den Einheiten Marketing, Dokumentation, Rechtsabteilung usw.
Nutzer und Nutzerinnen des öffentlichen Verkehrsraums	hier der beteiligte LKW-Fahrer und andere, die mit Teslas in Kontakt kommen
Gesellschaft im weiteren Sinne	Angehörige der Unfallopfer Überwachungsorgane Politik und Medien

Tabelle 1: Stakeholder im Fall Tesla

Für das „Cruise Robotertaxi“ lassen sich die Interessengruppen wie in Tabelle 2 dargestellt unterteilen.

Interessengruppe	Fall Cruise
Kunden	Die Nutzer und Nutzerinnen von Robotertaxis
Eigentümer des Unternehmens	Die verantwortlichen Leiter der Firma, die Mehrheitsanteilseigner
Mitarbeiter des Unternehmens	Die zahlreichen Mitarbeiter aus den verschiedenen Abteilungen, neben der Entwicklung beispielsweise auch aus den Einheiten Marketing, Dokumentation, Rechtsabteilung usw.
Nutzer und Nutzerinnen des	Fahrer des Nissans Die verletzte Fußgängerin

öffentlichen Verkehrsraums	und alle anderen Teilnehmer im Verkehrsraum
Gesellschaft im weiteren Sinne	Angehörige von Unfallopfern Überwachungsorgane Politik und Medien

Tabelle 2: Stakeholder im Fall Cruise

6.2 Klärung der moralischen Frage

In diesem Kapitel soll die moralische Frage definiert werden, um moralische von nicht-moralischen Fragestellungen zu trennen. Bleisch et al. (2021:45) führen dazu aus, dass Fragen nach Konventionen, Bräuchen oder Anstandsregeln oder technische oder instrumentelle Angelegenheiten wie eine Gebrauchsanweisung Fragen der normativen Art betreffen, es sich hierbei jedoch nicht um Fragen der Moral handelt. Die Ethik fragt danach, was zu tun „geboten“, „verboten“ oder „erlaubt“ ist. Bleisch et al. sprechen in diesem Zusammenhang von den „drei deontischen Operatoren“ (Bleisch et al. 2021:44). Daraus ergeben sich wieder Vorschriften, Regeln oder Maximen. Moralische Normen sollen dabei eine allgemeine Verbindlichkeit und Unparteilichkeit aufweisen. „Wir schulden einander gewisse Dinge ganz unabhängig von persönlichen Interessen oder Sympathien“ (Bleisch et al. 2021:47). Ferner benötigen moralische Normen eine „Begründungsbedürftigkeit“, d.h. sie sind nur dann legitim, wenn sie sich „jeder vernünftigen Person nachvollziehbar begründen lassen“. Wenn es darum geht, moralische Normen einzufordern, dann geht es in der Regel um fundamentale Werte und Interessen. Damit sind existentielle Bedürfnisse²⁰ oder das Recht auf Autonomie, d.h. wichtige Entscheidungen selber fällen zu können, gemeint.

Die moralische Schlüsselfrage, die sich für diese Forschungsarbeit stellt, kann daher wie folgt formuliert werden: Ist es aus theologischer Sicht zu verantworten, autonome Fahrzeuge fahren zu lassen?

²⁰ In diesem Zusammenhang sei auf die Maslowsche Bedürfnishierarchie, oder auch Bedürfnispyramide, von dem US-amerikanischen Psychologen Abraham Maslow (1908–1970) verwiesen.

6.3 Argumente der Stakeholder

Laut einer Analyse von Thadani, Lerman, Piper, Siddiqui und Uraizee (2023:3) wurden seit der Einführung des Tesla Autopiloten im Jahr 2014 mehr als 700 Unfälle registriert, davon zumindest 19 mit schwerwiegenden Folgen. Auch das Unfallfahrzeug im Fall Tesla gehört zu dieser Gruppe, in denen die automatisch Fahrfunktion namens „Autopilot“²¹ verbaut wurde. Auf einer Pressekonferenz im Jahr 2016 gab der Tesla Geschäftsführer Elon Musk bekannt, dass die Autopilot-Software im Moment besser sei als ein menschlicher Fahrer (Frankel 2016:1). Und weiter führte Musk aus, dass die Software laufend verbessert würde „by analyzing data from the hundreds of millions of miles driven by current owners“ (Frankel 2016:1). In Sinne der Verkehrssicherheit wäre es sehr zu begrüßen, wenn eine Technologie wie der Tesla Autopilot zu einer Verringerung der Unfallzahlen führen würde.

Das eigentliche Opfer des Unfalls mit dem Tesla Model 3 ist der Fahrer Jeremy Banner,. Auf der Firmenseite erläutert Tesla die Funktionalität des Autopiloten mittels eines Imagevideos. Auch Banner hatte sich laut Ehefrau diese Filme angesehen. In dem nach wie vor aktuellen Video wird ein Fahrer gezeigt, der durch die Straßen von Kalifornien fährt, ohne dabei das Lenkrad zu nutzen:

“The person in the driver’s seat is only there for legal reasons,” the video says. “He is not doing anything. The car is driving itself.” (Thadani et. al. 2023:18).

Ferner beschreibt Tesla auf der Firmenseite den Funktionsumfang des Autopiloten wie folgt:

Full Self-Driving Capability. All new Tesla cars have the hardware needed in the future for full self-driving in almost all circumstances. The system is designed to be able to conduct short and long distance trips with no action required by the person in the driver’s seat.

The future use of these features without supervision is dependent on achieving reliability far in excess of human drivers as demonstrated by billions of miles of experience, as well as regulatory approval, which may take longer in some jurisdictions. As these self-driving capabilities are introduced, your car will be continuously upgraded through over-the-air software updates. (Tesla 2024b)

Aus der obigen Beschreibung lässt sich vermuten, dass das Fahrzeug in der Lage sei, Fahrten ohne Zutun eines Fahrers oder einer Fahrerin durchzuführen. Im zweiten Teil der Beschreibung findet sich eine interessante Formulierung: Zunächst wird auf

²¹ Siehe <https://www.tesla.com/autopilot>.

die zukünftige Nutzung der Funktion ohne Aufsicht eingegangen, die davon abhängt, dass die Zuverlässigkeit weit über die von menschlichen Fahrern hinausgeht. Diese Vorgehensweise wäre aus christlich-ethischer Betrachtung absolut wünschenswert. Danach kommt der Einschub „[...] as demonstrated by billions of miles of experience, as well as regulatory approval [...]“ (Tesla 2024b), bei dem der Eindruck entsteht, dass die Qualität durch Milliarden von Kilometern schon bewiesen wurde und die behördliche Freigabe in einigen Ländern noch etwas länger dauern könnte. Zumindest könnte es ein Verbraucher oder eine Verbraucherin in dieser Weise interpretieren. Vermutlich geht es jedoch eher darum, dass die sichere Funktion in Zukunft noch bewiesen werden muss, durch „Milliarden von unfallfrei gefahrenen Kilometern“.

Eine Quelle für relevante Informationen ist das Benutzerhandbuch zum Tesla Model 3. Es soll im Folgenden geklärt werden, inwieweit der Unfallfahrer hier notwendige Informationen hätte finden können. Die in dem oben genannten Video gezeigte Funktionalität wird im amerikanischen Handbuch als „Full Self-Driving (Beta)“ (Tesla 2024:87) bezeichnet. In der europäischen bzw. deutschen Version fehlt die Nennung einer Beta-Version²² im Funktionsumfang des Fahrzeuges. Es scheinen hier Zweifel berechtigt zu sein, ob der Verkauf und die Verwendung von ungetesteter Software für selbstfahrende Fahrzeuge²³ für den deutschen oder europäischen Straßenverkehr überhaupt erlaubt wäre.

Auf der Einführungsseite zum Autopiloten gibt das System eine Warnung aus, dass der Fahrer oder die Fahrerin zu jeder Zeit in der Lage sein soll, die Kontrolle über das selbstfahrende Fahrzeug wieder zu übernehmen.

Keep your hands on the steering wheel at all times and be mindful of road conditions, surrounding traffic, and other road users (such as pedestrians and

²² Beta-Version ist ein Begriff aus der Software-Entwicklung. In diesem Entwicklungsstadium können Funktionalitäten enthalten sein, die noch nicht auf ihre Qualität, d.h. die sichere Verwendung überprüft wurden. Ein weiterer Ansatz kann sein, „echte Nutzer und Nutzerinnen“ für ein Beta-Testing-Programm zu gewinnen, um so Statistiken und Informationen über ihr Nutzerverhalten zu sammeln. Dieser Ansatz ist nach Humble & Farley ein „evolutionary approach to the adoption of features which is very effective“ (Humble & Farley 2011:90).

²³ In Deutschland benötigt ein Fahrer oder eine Fahrerin eines Testfahrzeuges einen speziellen Prototypenführerschein als Ergänzung zum allgemein notwendigen Führerschein. Siehe z.B. <https://www.fahrzeugetprobung.org/2017/05/18/pruefgelaendefuehrerscheine-prototypenfuehrerscheine-fuer-pkw/>, abgerufen am 23.04.2024.

cyclists). Pay attention to the road at all times and always be prepared to take immediate action. (Tesla 2024:87)

Wie sich hier sehen lässt, besteht eine Diskrepanz zwischen den Informationen, die auf der Webseite von Tesla zur Verfügung gestellt werden und die möglicherweise die Anschaffung eines Fahrzeuges beeinflussen und den Informationen, die im Handbuch bereitgestellt werden. Weitere für diese Arbeit wichtige Informationen finden sich in der Dokumentation zu den Limitierungen und Einschränkungen des Tesla Autopiloten (Tesla 2024:116ff). Hier soll auf drei Punkte eingegangen werden. Als erstes sieht Tesla die Verantwortung bei der Nutzung des Autopiloten beim Fahrer oder der Fahrerin: “It is the driver’s responsibility to stay alert, drive safely, and be in control of the vehicle at all times” (Tesla 2024:116). Ein weiterer zweiter Punkt ist die Objekterkennung in Abhängigkeit von der Fahrgeschwindigkeit. Tesla warnt den Nutzer und die Nutzerin, dass bei hohen Geschwindigkeiten, d.h. ab 80 km/h andere Verkehrsteilnehmer (Objekte) nicht sicher erkannt werden und das selbstfahrende Fahrzeug auch nicht bremst, wenn andere Objekte sich nur sehr langsam bewegen oder nur partiell in der Sichtachse zum Fahrweg sichtbar sind (Tesla 2024:116). Ein dritter Punkt mit großem Einfluss auf die Funktion des Autopiloten sind die Licht- und Sichtverhältnisse zur jeweiligen Fahrsituation. Tesla gibt ein paar Beispiele, in denen sein System möglicherweise fehlerhaft funktioniert. Dabei gibt Tesla an, dass diese Liste von möglichen Einschränkungen nicht vollständig sei und “it is the driver’s responsibility to be in control of Model 3 at all times” (Tesla 2024:117). Hier ist eine Auswahl von Situationen, die besonders häufig Fehler aufweisen könnten:

The road has sharp curves or significant changes in elevation. Road signs and signals are unclear, ambiguous, or poorly maintained. Visibility is poor (due to heavy rain, snow, hail, etc. or poorly lit roadways at night). You are driving in a tunnel or next to a highway divider that interferes with the view of the camera(s). Bright light (such as from oncoming headlights or direct sunlight) interferes with the view of the camera(s). (Tesla 2024:116)

Zusammengefasst besteht das Risiko für den Fahrer oder die Fahrerin, dass Objekte vom System unter bestimmten Bedingungen nicht erkannt werden. Inwieweit die Informationen über die konkrete Fahrsituation unzureichend für das selbstfahrende Fahrzeug sind, sagt das System dem Fahrenden nicht.

In dem Artikel „The final 11 seconds of a fatal Tesla Autopilot crash – A reconstruction of the wreck shows how human error and emerging technology can

collide with deadly results “ von Thadani et al. (2023) wurden Fachleute um eine Einschätzung gebeten. Ein Interviewpartner war Professor Philip Koopmann von der Carnegie-Mellon-Universität. Er forscht seit über 25 Jahren zu autonomen Fahrzeugen. Aus seiner Sicht darf sich das autonome Fahren nur dann aktivieren, wenn das System auch ausreichend dafür in der Lage ist, die Steuerung zu übernehmen:

“If a system turns on, then at least some users will conclude it must be intended to work there,” Koopman said. “Because they think if it wasn’t intended to work there, it wouldn’t turn on.” (Thadani et. al. 2023:21)

Eine weitere Einschätzung gab Professor Andrew Maynard von der Arizona State University ab. Nach seiner Meinung vertrauen die Kunden einfach der Technologie, da sie nicht die Zeit oder das Verständnis hätten, die Feinheiten zu verstehen – „so at the end they trust the company to protect them“ (Thadani et. al. 2023:21).

Der Untersuchungsbericht NTSB (2020:15) fasst die wahrscheinlichen Ursachen für den Unfall folgendermaßen zusammen: Als erster Punkt wird das fehlerhafte Verhalten des LKW-Fahrers genannt, der dem Tesla an der Kreuzung die Vorfahrt hätte gewähren müssen. Als zweiter Punkt wird die Funktionsweise des Tesla Autopiloten genannt, der für einen gewissen Zeitbereich eine geistige Abwesenheit des Unfallfahrers erlaubte. Und zuletzt noch die Versäumnisse des Unternehmens, den Einsatz des Autopiloten auf die Bedingungen zu beschränken, für die er entwickelt wurde.

Erschwerend kommt hinzu, dass dem National Highway Traffic Safety Administration (NHTSA) keine Verifikationsmethode zum Testen von Level-2-Fähigkeiten (nach der SAE-Einstufung) zur Verfügung stand (NTSB 2020:15). Laut NTSB wurde deshalb eine Verbesserungsempfehlung mit der Kennung „H-17-38“ an das NHTSA gerichtet, um diesen Umstand zu verbessern (:13). Dieser hatte zum Report-Zeitpunkt den Status „Open – Unacceptable Response“ (:13).

Ein weiteres Themenfeld in der Argumentation ist die Kritik an der Vermarktungspraxis der in den Teslafahrzeugen angebotenen Autopilot-Funktion. In diesem Zusammenhang hat sich das amerikanische „Center for Auto Safety“ zu Wort gemeldet. Das Center wurde 1970 gegründet und ist nach eigenen Angaben eine unabhängige, gemeinnützige Verbraucherschutzorganisation in den USA, die sich u.a. für die Verbesserung der Sicherheit von Fahrzeugen einsetzt. In dem offenen

und zehnteiligen Brief an die Staatsanwaltschaft bittet die Vereinigung um eine Untersuchung bezüglich der „irreführenden und trügerischen Darstellungen von Tesla in Bezug auf die Sicherheit und Fähigkeiten seiner Autopilot-Funktion“ (Levine 2019). Mit Bezug auf das National Transportation Safety Board (NTSB) und eigenen Analysen geht das „Center for Auto Safety“ (Center) davon aus, dass ein mangelhaftes Verständnis für die Funktion des Autopiloten zu den Unfällen beigetragen hat und dieses falsche Verständnis würde u.a. durch die Vermarktungspraxis der Firma Tesla begünstigt:

After studying the first of these fatal accidents, the National Transportation Safety Board (NTSB) determined that over-reliance on and a lack of understanding of the Autopilot feature can lead to death. There can no longer be any question that these deceptive practices have made it reasonable for Tesla owners to believe that a Tesla with Autopilot is an autonomous vehicle capable of self-driving. To be clear, it is not. (Levine 2019:2)

Eine weiterer Kritikpunkt ist die Einstufung der Tesla-Fahrzeuge in einem höheren SAE-Level-Schema.

Tesla continues to be the only automaker to describe its Level 2 vehicles as “self-driving,” and the name of its driver assistance suite of features, Autopilot, connotes full autonomy. (Levine 2019:2)

Eine Pflichtaufgabe für Fahrerinnen und Fahrer für die Nutzung von Fahrassistenzsystemen nach SAE-Stufe zwei ist das fortwährende Überwachen des Systems, z.B. muss der Fahrer oder die Fahrerin den Lenker während der Fahrt laufend berühren. Im Falle des Autopiloten kritisiert das Center, dass eine Fahrt ohne das Halten des Lenkrades für einen längeren Zeitraum möglich sei. Dadurch würde der Verbraucher zu der Annahme verleitet, dass der Autopilot das Fahrzeug zu einem selbstfahrenden Fahrzeug macht (Levine 2019:2). Das Center listet in seinem Brief eine Reihe von Beispielen auf, in denen aufgrund von fehlerhaftem Verständnis die Fahrer oder Fahrerinnen die Kontrolle für die Fahrt an den Autopiloten übergeben haben und das System nicht überwachten. So wird von Levine auch das in dieser Arbeit untersuchte Unglück beschrieben und am Ende kommt er zu dem Schluss, dass die Nutzung eine konstante Überwachung benötigt:

As with previous incidents, if the driver of this Tesla had understood that the Autopilot system required constant attention to the driving environment, he might still be alive today (Levine 2019:6).

Eine Schuld liegt bei dem Tesla-Fahrer selbst: Zum einen hat er auf der Strecke eine überhöhte Geschwindigkeit eingestellt und zum anderen war das autonome Fahren für diesen Streckenabschnitt nicht freigegeben, es ließ sich jedoch trotzdem im Auto aktivieren. In diesem kurzen Zeitbereich der Übernahme hätte das Fahrzeug die gefährliche Situation erkennen und bei Unlösbarkeit die Verantwortung wieder an den Fahrer zurückgeben müssen. Aufgrund der Unfallanalyse ist dieses jedoch nicht geschehen. Weder der Fahrer noch das Automatik-System hatten Maßnahmen zur Unfallverhinderung eingeleitet. Die Fehler der Technik nutzt laut Thadani, Lerman, Piper, Siddiqui und Uraizee (2023) auch der Anwalt der Opferfamilie und argumentiert, die Technologie hätte wiederholt versagt, als sie nicht bremste oder als sie keine Warnung vor dem Sattelschlepper gegeben habe.

Ein Stakeholder ist der 45-jährige LKW-Fahrer, der an dem Unfall beteiligt war. Dem Fahrer unterlief morgens ein Fahrfehler, d.h. an der Kreuzung hätte er an einem Stoppschild kurz anhalten müssen. Die Untersuchung ergab, dass er das Fahrzeug verlangsamte und dann nach Überprüfung der Straße wieder beschleunigte, um die Kreuzung zu passieren. Aus der Befragung nach dem Unfall kann man erkennen, wie er mit sich und der Situation haderte. Er hätte die Kreuzung in der Vergangenheit schon oft passiert und hätte nach seiner Meinung auch diesmal genügend Zeit zum Passieren gehabt:

“I’ve done it a dozen times,” the driver said of his fateful left turn. “And I clearly thought I had plenty of time. I mean, it was dark, and the cars looked like they was back further than what they was.” “Yeah,” the investigator said. “And, I mean, it’s just something I’m —,” the driver said. “It’s okay, it’s okay,” the investigator responded. “Yeah, take your time,” another investigator said. “Just,” the driver said, pausing again. “It’s something I’m going to have to live with.” (Thadani et. al. 2023:24)

Ein Unfallopfer oder -verursacher zu sein, kann zu langwierigen Beeinträchtigungen der Person führen. Von außen kann man nur erahnen, welche Einflüsse dies auf den LKW-Fahrer haben mag. Aus den Formulierungen im Interview lässt sich sein Verantwortungsgefühl und Schuldbewusstsein erahnen.

6.4 Ein Blick in die Luftfahrt

An dieser Stelle der Diskussion soll auf ein anderes Gebiet eingegangen werden, in dem der Gedanken einer übergeordneten Aufsicht für automatische Systeme schon in der Verkehrstechnik erfolgreich umgesetzt wurde: in der Luftfahrt. Die

Sicherheitsanforderungen sind in diesem Gebiet sehr streng, da verständlicherweise ein Ausfall katastrophale Folgen haben könnte. Wir kommen nun zu einem bekannten Unfall.

Der Pilot steuert oder fliegt das Flugzeug. Dabei hat seine Rolle einen eher überwachenden Charakter. Die Maschinen fliegen nahezu selbständig durch den verbauten „Autopiloten“. Unterstützt wird der Pilot oder die Pilotin bei großen Flugzeugen durch einen zweiten Piloten oder eine Pilotin und zusätzlich durch das Konzept von Fluglotsen, die für einen jeweiligen Flughafen den Flugverkehr überwachen. Sowohl Lotsen als auch Piloten werden auf „ihrem System“ ausgebildet und trainiert. Das bedeutet, ein Pilot oder eine Pilotin wird mit den Eigenheiten des jeweiligen Modells vertraut gemacht.

Im Allgemeinen gilt somit die Reise mit einem Flugzeug als sehr sicher. Hierzu haben die umfangreichen Vorkehrungen, wie die in der Entwicklung einbezogenen Sicherheitsstandards und Zertifizierungen, das geschulte Personal in der Luft und die Überwachung vom Boden jeweils ihren Beitrag geleistet. Und trotz allem kann es zu einem unvorhergesehenen Ereignis kommen. Der Öffentlichkeit bekannt wurde der Flug 1549 der US Airways aufgrund einer Notwasserung auf dem Hudson River am 15. Januar 2009, bei dem alle Passagiere überlebten. Flugkapitän war Chesley Sullenberger. Seine Geschichte wurde in dem Film „Sully“ 2016 verfilmt²⁴.

Der Unfallhergang wurde, wie allgemein in diesen Fällen üblich, genauestens untersucht. Das National Transportation Safety Board hat zu diesem Unfall einen 213-seitigen Bericht „Loss of Thrust in Both Engines After Encountering a Flock of Birds and Subsequent Ditching on the Hudson River“ vorgelegt (NTSB 2010). Der erfahrene Pilot entschied sich für eine außergewöhnliche Maßnahme – die Notlandung des Flugzeuges auf dem Hudson River. Dies war nicht die einzige zielführende Entscheidung des Piloten. Um ein Beispiel für das verantwortliche Handeln und die Schuldübernahme in der konkreten Situation zu beleuchten, soll hier auf zwei Entscheidungen des Piloten eingegangen werden: Erstens die Überprüfung der Antriebe und zweitens sein Verhalten bei der Evakuierung.

²⁴ Quelle: https://de.wikipedia.org/wiki/US-Airways-Flug_1549, zugegriffen am 20.03.2025.

Die Maßnahmen während des Antriebsausfalls waren folgende: Laut dem Untersuchungsbericht der NTSB übernahm der Pilot zwölf Sekunden nach dem Vogelschlag die Kontrolle über das Flugzeug (NTSB 2010:87). Weitere fünf Sekunden später begann die Crew mit der Überprüfung der dreiseitigen Handlungsanweisung „Engine Dual Failure Checklist“. Obwohl die Flugbesatzung die Überprüfung erst zu einem Drittel durchgeführt hatte, startete der Kapitän die zusätzliche Energieversorgung „APU“ (Auxiliary Power Unit). Dies war laut der Checkliste zu dem Zeitpunkt noch nicht vorgesehen und der Pilot wich dadurch von den Standardvorgaben ab. Das NTSB sieht diese Entscheidung im Nachhinein betrachtet als sehr wichtig an, da dadurch das Flugzeug während der Notlandung steuerbar blieb:

Starting the APU early in the accident sequence proved to be critical because it improved the outcome of the ditching by ensuring that electrical power was available to the airplane. Further, if the captain had not started the APU, the airplane would not have remained in normal law mode. This critical step would not have been completed if the flight crew had simply followed the order of the items in the checklist (NTSB 2010:88).

Die Evakuierung erschien rückblickend betrachtet recht geordnet. Und das, obwohl das Wasser im hinteren Teil des Flugzeugs schnell anstieg und es zu improvisierten Anweisungen der Flugbegleiter kam (NTSB 2010:106). Am Schluss der Evakuierung stieg der Flugkapitän als Letzter aus dem sinkenden Flugzeug. Zeugenberichten zufolge war er noch einmal durch die Maschine gegangen, um sicherzugehen, dass alle Passagiere das Flugzeug verlassen hatten (Pitzke 2009). Dieses Beispiel folgt der sprichwörtlichen Regel „der Kapitän verlässt als letzter das Schiff“ als Beispiel für eine volle Verantwortungsübernahme, die ein erhöhtes Risiko für die eigene Gesundheit in Kauf genommen hat.

6.5 Evaluation der Argumente

6.5.1 Fall Tesla

Im Fall Tesla haben wir vor allem zwei wichtige Beteiligte am Unfallgeschehen: Der Tesla Fahrer und der Fahrer des LKWs. Obgleich der Fokus auf dem autonom fahrenden Fahrzeug und der fatalen Fehleinschätzung der Situation liegt, scheint aus der Sicht des Bonhoefferschen Verständnisses von Verantwortung der LKW-Fahrer

interessant zu sein. Bei genauerer Betrachtung lernen wir seine körperlichen Einschränkungen kennen, d.h. zum einen die Schwierigkeiten, Distanzen einzuschätzen und auf der anderen Seite die behandelte Epilepsie. Die Diskussion um die autonomen Fahrzeuge hat ihn zuvor möglicherweise nicht erreicht. Und trotzdem ist er ein Opfer dieser neuen Technologie geworden, da er aufgrund seiner Fehleinschätzung zu früh in die Kreuzung eingefahren ist und es so zu dem fatalen Crash gekommen ist.

Die Situation ließe sich mit Bonhoeffers Perspektive der Schuldübernahme beschreiben, wenn der Verantwortliche Schuld auf sich lädt und z.B. einen Unbeteiligten verletzt. An dieser Stelle lässt sich fragen, wie sorgfältig und überlegt der Teslafahrer gehandelt hat. Er war auf dem Weg zur Arbeit, die Strecke hat er somit vermutlich täglich befahren. Als begeisterter Teslafahrer hat er vermutlich die Funktionen des Wagens ausprobiert. Laut Unfallbericht ist er mit erhöhter Geschwindigkeit gefahren. Aufgrund der täglichen Wegstrecke lässt sich annehmen, dass er die Geschwindigkeitsvorschriften auf der Wegstrecke kannte und hier bewusst eine Überschreitung eingegangen ist. Über die Gründe lässt sich nur mutmaßen. Vielleicht war es zu der Zeit sehr ruhig auf der Straße oder vielleicht hatte er eine wichtige Aufgabe bei seiner Arbeit zu erledigen und er wollte rechtzeitig am Ziel ankommen. Laut Untersuchungsbericht waren seine Hände während des Unfalls nicht am Lenkrad. Dies könnte darauf hindeuten, dass der Fahrer mit anderen Dingen beschäftigt war und seine Aufmerksamkeit nicht dem Verkehrsgeschehen galt. Erstaunlich am Fahrzeug sind folgende Punkte. Zunächst muss das Fahrzeug die aktuelle zulässige Höchstgeschwindigkeit auf der Straße kennen. Wenn das Fahrzeug autonom fahren kann, ist es erstaunlich, dass eine höhere als die zulässige Geschwindigkeit eingestellt werden kann. Zweitens gibt das Handbuch eine Reihe von Einschränkungen bezüglich der Sicht bei Dämmerung an. Auch hier ist es verwunderlich, dass in dem System keine Überwachung der konkret verfügbaren Möglichkeiten des Systems vorhanden ist. D.h., das System hat offensichtlich keine „Selbsteinschätzung“, ob es in der gegenwärtigen Situation und insbesondere der gegenwärtigen Sichtverhältnisse in der Lage ist, die Kontrolle zu übernehmen. Und drittens konnte das System nicht überprüfen, ob der Fahrer das System ausreichend überwacht. Dabei ist die Überprüfung, ob der Fahrer eine Hand am Lenkrad hat, ein sehr rudimentärer Test. Aus diesen drei Punkten lässt sich

festhalten, dass das Fahrzeug sehr wenig „Wissen“ über seine aktuelle Umgebung und zusätzlich dem Verhalten des Fahrers hatte.

Dem Teslafahrer kann man den Vorwurf machen, dass er eine Geschwindigkeitsübertretung initiiert hat. Es fällt auf, dass Fahrer und Fahrzeug aus der Sicht der Verantwortung keine Einheit bilden bzw. sich nicht ergänzen. Wider besseren Wissens fährt der Fahrer zu schnell und das Fahrzeug übernimmt im Automatikmodus diese Einstellung. Aufgrund der Kenntnis des Kartenmaterials hätte das Auto „wissen“ müssen, welche Einstellungen an der Straße erlaubt gewesen wären. Und dabei ist dies nur eine „statische Information“, die keine oder geringe Anforderungen an das maschinelle Sehen stellt. Eine weitere wichtige Erkenntnis ist, dass das Fahrzeug nicht erkennt, wenn es aufgrund von Umwelteinflüssen nicht in der Lage ist, die Steuerung sicher zu übernehmen. Das Handbuch beschreibt eine Vielzahl von Einschränkungen. Beispielsweise sind die Lichtverhältnisse eine solche Einschränkung. Für eine verantwortliche Umsetzung sollte das Fahrzeug die Selbstfahrfunktionen nur dann anbieten, wenn die Witterungsbedingungen und die Straßenlage solches auch zulassen. Der Fahrer muss davon ausgehen, dass er ein getestetes Fahrzeug nutzt. Der Fahrer ist im guten Glauben, dass er die Aufgabe an das Fahrzeug „delegiert hat“ und widmet sich anderen Dingen oder ist abgelenkt. Der Fahrer hatte im Nachhinein betrachtet vermutlich keine genaue Vorstellung davon, was das automatische System in der Lage ist zu leisten. Er vertraut der Werbung und hat möglicherweise die Einschränkungen nicht realisiert. Es kommt erschwerend hinzu, dass die Funktionen laufend vom Hersteller geändert werden. D.h., theoretisch denkbar wäre die Situation, dass kurz zuvor ein Update installiert wurde und der Fahrer nun die neue Funktionalität testen wollte.

6.5.2 Fall Cruise

Aus dem Cruisebeispiel lässt sich sehr gut das menschliche Fehlverhalten erkennen: Die Schuld der Fußgängerin als auch das fehlerhafte Verhalten des ersten Unfallfahrers. Ferner wird deutlich, wie unsicher die Situation aus der Perspektive der technischen Aufsicht erkannt wurde. Möglicherweise fehlte hier die Erfahrung, da eine neue Technologie eingeführt wurde und entsprechende Ausbildungswege noch etabliert werden müssen. Ein anderer Aspekt könnte sein, dass dem Leitstand nicht alle Informationen zur Verfügung standen. Denn nicht alle Verkehrsteilnehmer wer-

den erkannt und senden Informationen an den Leitstand bzw. an die Fahraufsicht. Im Gegensatz dazu werden Flugzeuge durch ihre Position und Kennzeichnung den jeweiligen Fluglotsen automatisch mitgeteilt. Dadurch kann in der Flugüberwachung automatisch ein abstraktes Bild der realen Situation für die Fluglotsen erstellt werden. Im Gegensatz dazu haben bei dem Unfall in San Francisco die Mitarbeiter und Mitarbeiterinnen nur Teile der Informationen vorliegen. Sie waren somit nur bedingt in der Lage, eine gute Entscheidung im Sinne der Menschen zu treffen.

Ein weiteres offensichtliches Beispiel für die Schuld ist die unterbliebene Erste Hilfe durch den ersten Fahrer. Doch was wäre passiert, wenn der Fahrer versucht hätte, auszusteigen und der verletzten Frau zu helfen? Es lassen sich nur Vermutungen anstellen, ob das nachfolgende selbstfahrende Fahrzeug die Situation richtig erkannt hätte. Hieraus lässt sich eine Rückfrage an die Entwickler und Entwicklerinnen (und deren Gewissen) formulieren: Wurden entsprechende Situationen in den Trainingsdaten der künstlichen Intelligenz berücksichtigt?

Gegenüber dem ersten Fahrer hat das Robotertaxi nahezu vorbildlich gehandelt. Es stoppte die Fahrt, nachdem die Fußgängerin auf der Straße erkannt wurde. Das selbstfahrende Fahrzeug konnte auf das fehlerhafte Verhalten der Fußgängerin reagieren. An dieser Stelle wäre eine Erste-Hilfe-Maßnahme unter Zuhilfenahme der Passagiere des Taxis ein wertvolles Verhalten gewesen. Leider hatten die Entwickler und Entwicklerinnen als Notmaßnahme das „sicherere Erreichen des rechten Straßenrandes“ als einzige Möglichkeit vorgesehen. Aus der Sicht Bonhoeffers haben sie dadurch Schuld auf sich geladen, da sie diese Unfallsituation unter Laborbedingungen nicht durchdacht haben.

7 BEANTWORTUNG DER FORSCHUNGSFRAGE

Ausgangspunkt für diese Forschungsarbeit bildet eine Untersuchung zu zwei realen Unfällen mit selbstfahrenden Fahrzeugen. Die Fälle wurden so gewählt, weil ausreichend Material für die Auswertung publiziert wurde. Zum anderen erfolgte die Auswahl aufgrund unterschiedlicher Fahrzeugeigenschaften: Im Fall Tesla kann das autonome Fahren während der Fahrt vom Fahrer oder der Fahrerin aktiviert und auch deaktiviert werden. Im Fall Cruise besteht diese Option nicht, da bei diesem Typ kein Fahrer oder keine Fahrerin vorgesehen ist.

Drei Beobachtungen gleich vorweg: Erstens fällt bei der Betrachtung der ethischen Debatte auf, dass die Diskussion vor allem auf der Basis von Gedankenexperimenten vollzogen wird. Zweitens, die christliche Verantwortungsethik spielt in der Diskussion so gut wie keine Rolle. Darum zielt die Forschung dieser Arbeit darauf ab, einen Beitrag aus dieser Perspektive zu leisten. Und als dritte Beobachtung stehen die Unfälle im grundsätzlichen Widerspruch zu der Annahme, dass durch eine weitere Automatisierung im Straßenverkehr Opferzahlen reduziert werden können.

Für diese Arbeit wurde folgende Forschungsfragen gewählt:

Hauptforschungsfrage:

Ist es aus theologischer Sicht zu verantworten, autonome Fahrzeuge fahren zu lassen?

Dazu ergänzend die Sub-Forschungsfragen:

1. Welche Umstände haben zu einem Unfall mit einem autonomen Fahrzeug geführt?
2. Wie ist die Nutzung von selbstfahrenden Fahrzeugen aus der Perspektive der christlichen Verantwortungsethik zu bewerten?

Um dem interdisziplinären Charakter dieser Forschungsarbeit gerecht zu werden, wurde methodisch zunächst der Forschungsstand erläutert und eine Grundlegung für

die technische Seite vorgenommen. Mit diesem Wissen konnten die Umstände der zwei Unfälle genauer untersucht werden und damit wurde die erste Sub-Forschungsfrage beantwortet. Hierbei ist ein für die Debatte wichtiges Ergebnis zum Vorschein gekommen: Bei beiden Unfällen führte eine geradezu tragische Verkettung von Fehlern zu einer schweren Verletzung eines Verkehrsteilnehmers oder einer Verkehrsteilnehmerin. Ferner zeigte sich das schuldhafte Verhalten der Menschen, die wider besseres Wissen die gesetzlichen Regelungen des Straßenverkehrs missachtet haben. Dagegen basieren Gedankenexperimente häufig auf einer Situation, die nicht diese Tiefe an Verkettungen in den Ereignissen enthält. Zusätzlich zur Verkettung von Fehlern spielt der zeitliche Ablauf im einstelligen Sekundenbereich eine entscheidende Rolle.

Aufbauend auf die Erkenntnisse aus den Fallstudien wurde eine theologische Grundlegung vorgenommen. Zielsetzung dabei war, einen theologischen Blickwinkel für die Beantwortung der zweiten Sub-Forschungsfrage zu legen. In diesem Kapitel kamen die verantwortungsethischen Ansätze von Bonhoeffer und Niebuhr zu Wort. Die Relevanz ihrer Arbeiten wurde ferner durch zeitgenössische Beiträge unterstrichen. Zudem wurde die Betrachtung durch einen Beitrag von Brunner abgerundet. Darauf aufbauend kamen die einzelnen Stakeholder mit ihren Argumenten zu Wort.

Auf diesem bis hierher geschilderten Weg sind somit Beiträge für die Beantwortung der Forschungsfragen entstanden. Grundsätzlich ist das Leben und die Gesundheit aller Verkehrsteilnehmer und -teilnehmerinnen zu schützen. Dies ist das Ziel. Aus dem Blickwinkel der Verantwortungsmatrix wäre es zu begrüßen, wenn selbstfahrende Fahrzeuge dazu beitragen würden. Jedoch beschreibt die Verantwortungsmatrix auch das schuldhafte Handeln. Im Fall Tesla wurde beispielsweise bei überhöhtem Tempo die Automatik aktiviert. Zusätzlich hat es der Unfallfahrer versäumt, das System auf korrekte Funktion zu überwachen.

Implizit ist es zu weiteren Beobachtungen gekommen. In der technischen Grundlegung wurde die durch die Kybernetik beschriebene Systemsicht beschrieben. Einzelne Systeme können dabei wieder zu einem übergeordneten System zusammengefasst werden. Diese Systemsicht kann bei der Betrachtung der beiden

Unfälle genutzt werden: Zum einen fand die Kollision im öffentlichen Verkehrsraum statt. Die andere Besonderheit ist das Zusammenspiel zwischen einem autonom fahrenden Fahrzeug und einem klassischen Verkehrsteilnehmer, d.h., es fand ein sogenannter Mischbetrieb statt. Möglicherweise wäre eine erhöhte Sicherheit zu erzielen, wenn alle Fahrzeuge mit einer Technologie ausgestattet würden, um ihre Existenz und ihre Absichten dem direkten Umfeld mitzuteilen. Dieser Mischbetrieb als erweitertes System könnte somit eine Hürde für die Einführung von selbstfahrenden Fahrzeugen darstellen.

Ein weiterer Punkt nach der Analyse der Argumente scheint ein unklares Verständnis über die Leistungsfähigkeit autonomer Fahrzeuge und deren Nutzung zu sein. Hier müsste dringend nachgebessert werden, zum Beispiel durch eine Erweiterung der Führerscheinausbildung oder durch ein Einschreiten der staatlichen Stellen, wenn eine Diskrepanz zwischen aktuellem Funktionsumfang und angedachter Zielbeschreibung die Verbraucher zu Fehleinschätzungen führt.

Die untersuchten Fälle widerlegen die These, dass ein automatisches System in jedem Fall bessere Ergebnisse liefert. Das „gesamte“ Teslasystem aus Produktmarketing und Produkt mit seinen Einschränkungen und Freiheitsgraden erzeugten beim Endanwender im untersuchten Fall ein fehlerhaftes Bild über die Leistungsfähigkeit des Systems. Dadurch entsteht für den Fahrer oder die Fahrerin oder beteiligte Personen eine zusätzliche Gefährdung, anstatt den propagierten gesteigerten Sicherheitsgewinn zu erzielen.

Zur Beantwortung der Forschungsfrage wurden die Ansätze zur Verantwortungsethik von Bonhoeffer und Niebuhr verwendet. Niebuhr geht von einem Antwort geben aus, bzw. eine Antwort auf eine gedachte Reaktion. Dies ist im Straßenverkehr ein übliches Verfahren. Verkehrsteilnehmer und Verkehrsteilnehmerinnen müssen sich verständigen, beispielsweise durch technische Hilfsmittel, durch Blickkontakt oder zum Beispiel durch Handzeichen. Im Falle des Teslafahrers blieb eine Reaktion aus. Auch bei dem Robotertaxi scheint diese Art der Kommunikation nicht gegeben zu sein. Das Taxi hat zwar löblicherweise nach dem Erkennen des Hindernisses einen Nothalt durchgeführt. Ein erfahrener Fahrer oder Fahrerin könnte die Situation im Vorfeld jedoch besser einschätzen, die Unfallsituation schon frühzeitig erkennen und

auch anderen Verkehrsteilnehmer warnen. In beiden untersuchten Fällen fehlte die Art der menschlichen Kommunikation aufgrund der Übertragung der Entscheidungsgewalt auf die Maschine. Es wurde in beiden Fällen nicht gelöst, wie der im Straßenverkehr notwendige „Blickkontakt“ zu anderen Teilnehmern realisiert werden sollte.

Wie gezeigt, definiert Bonhoeffer die Eckpunkte Wirklichkeitsgemäßheit, Stellvertretung, Gewissen und Schuldübernahme in seiner Matrix der Verantwortung. Bonhoeffers Ansatz basiert auf einem reflektierten Handeln, wie er es im Dritten Reich als Beispiel vollzogen hat. Im Tesla-Fall vollzieht der Tesla-Fahrer folgenschwere Fehler. Es fällt im Nachhinein schwer, hier von einem reflektierten Handeln zu sprechen, die der Fahrer in Verantwortung für die anderen Menschen vollzieht. Im Cruise Fall könnte man durch das Nothalten ein verantwortliches Verhalten der Entwickler dieses Systems erkennen, die versucht haben, eine mögliche Gefahrenabwehr in das System zu implementieren. In dem untersuchten Cruise-Fahrzeug waren diese Maßnahmen leider unzureichend vorhanden.

Und so wäre das selbstfahrende Fahrzeug für die Wirklichkeit besonders förderlich, wenn dadurch die Leben der Bürger geschützt werden würden. Die untersuchten Fälle zeigten technische Unzulänglichkeiten auf und ferner fehlendes Wissen bei den Nutzern des Systems. In Bezug auf Lebensschutz muss hier eine Nachbesserung eingefordert werden.

Zusammengefasst lässt sich somit für die Beantwortung der Hauptforschungsfrage folgendes formulieren: In beiden untersuchten Fällen ist es aus Sicht der christlichen Verantwortungsethik nicht vertretbar, die Fahrzeuge autonom fahren zu lassen. Der Versuch, den Menschen durch eine Automatisierung bzw. KI zu ersetzen, hat in den beiden untersuchten Fällen zu schweren Folgen für die beteiligten Personen geführt.

Und abschließend soll noch ein Ausblick gewagt werden: Es scheint auf der Basis dieser Arbeit angebracht zu sein, ein Paradigmen-Wechsel im Lösungsansatz zu vollziehen: Anstatt von einem „Ersetzen des Menschen“ zu sprechen, sollte von einer „Ergänzung durch KI“ gesprochen werden. Der Mensch wird in der Fahraufgabe durch eine KI in seinen Fähigkeiten unterstützt. In dem Zusammenschluss würde

etwas entstehen, das den vorher gekannten Fähigkeiten überlegen ist. In diesem Zusammenschluss würde der Mensch dann eine neue Rolle erhalten. Ähnlich wie der Pilot im Flugzeug müsste der Mensch in die Lage versetzt werden, das System zu überwachen. Dazu braucht es neue geeignete Maßnahmen und Kenntnisse. Und dies lässt Raum für weitere Forschungen und Debatten.

8 LITERATURVERZEICHNIS

- Anjum, S. K. & Wolff, C., 2021, "Agile Principles in Automotive Software Development: Analysis of Potential Levers," in IEEE European Technology and Engineering. Management Summit (E-TEMS), 141–147, von doi: 10.1109/E-TEMS51171.2021.9524860.
- Aristoteles, 2015, *Nikomachische Ethik*, Reclam, Stuttgart.
- Arkin, R. C., 2000, *Behavior-based robotics*, MIT Press, Cambridge, Mass., London.
- Asimov, I., 1942, "Runaround" in *Astounding Science Fiction*, 94–103.
- Asimov, I., 2015, *Ich, der Roboter*, Heyne, München.
- Barentsen, J., Bunnell, A., Ebsworth, J., Henning, K., Little, B. A., Peckham, J., Plutschinski, T. (Hg.), 2021, *The 4th Industrial Revolution from an Ethical Perspective*, Sallux, Amersfort.
- Bartneck, C., Lütge, C., Wagner, A. & Welsh, S., 2019, *Ethik in KI und Robotik*, Hanser, München.
- BBC, 2023, *Cruise self-driving cars investigated after two accidents*, angesehen am 03.11.2023, von <https://www.bbc.com/news/technology-67133409>.
- Beauchamp, T. L., 2021, „Der ‚Vier-Prinzipien‘-Ansatz in der Medizinethik“ in Biller-Andorno, N., Monteverde, S., Krones, T. & Eichinger, T. (eds) *Medizinethik. Grundlagentexte zur Angewandten Ethik*. Springer VS, Wiesbaden.
- Bedford-Strohm, J., Höhne, F. & Zeyher-Quattlender, J. (Hg.), 2019, *Digitaler Strukturwandel der Öffentlichkeit. Interdisziplinäre Perspektiven auf politische Partizipation im Wandel*, Nomos, Baden-Baden.
- Bethge, E., 2017, *Dietrich Bonhoeffer (Rowohlt Monographie)*, Rowohlt, Hamburg, Kindle Edition.
- Biewald, R. & Beckmann, J., 2007, *Verantwortliches Leben, Eine theologisch-ethische Spurensuche*, angesehen am 05.08.2024, von <https://www.dietrich-bonhoeffer.net/fileadmin/media/projekte/werkbuch-6/werkbuch-6.pdf>.
- Bleisch, B., Huppenbauer, M. & Baumberger, C., 2021, *Ethische Entscheidungsfindung. Ein Handbuch für die Praxis*, Nomos, Baden-Baden.
- Bonhoeffer, D., 2006, *Ethik*, Gütersloher Verlagshaus, Gütersloh.
- Bonhoeffer, D., 2023, *Ethik*, 7. Auflage der Taschenbuchausgabe, Gütersloher Verlagshaus, Gütersloh.
- Bonhoeffer, D. & Zimmerling, P. (Hg.), 2021, *Freiheit zum Leben: Ausgewählte Texte zur Ethik*, Brunnen, Gießen.
- Bonnefon, J.-F., Sharrif, A., & Rahwan, I., 2016, „The social dilemma of autonomous vehicles“ in *Science*, 352(6293), 1573–1576.
- Brunner, E., 1937, *Der Mensch im Widerspruch*, Furche, Berlin.
- Brunner, E., 1958, *Gott und sein Rebell. Eine theologische Anthropologie*, Rowohlt, Hamburg.
- Budelacci, O., 2022, *Mensch, Maschine, Identität. Ethik der Künstlichen Intelligenz*, Schwabe, Basel.
- Bundesministerium für Digitales und Verkehr (BMDV), 2017, *Ethik-Kommission Automatisiertes und Vernetztes Fahren*, angesehen am 20.07.2024, von https://bmdv.bund.de/SharedDocs/DE/Publikationen/DG/bericht-der-ethik-kommission.pdf?__blob=publicationFile.
- Bundesministerium für Digitales und Verkehr (BMDV), 2021, *Entwurf eines Gesetzes zur Änderung des Straßenverkehrsgesetzes und des*

- Pflichtversicherungsgesetzes – Gesetz zum autonomen Fahren*, angesehen am 15.06.2024, von https://bmdv.bund.de/SharedDocs/DE/Anlage/Gesetze/Gesetze-19/gesetz-aenderung-strassenverkehrsgesetz-pflichtversicherungsgesetz-autonomes-fahren.pdf?__blob=publicationFile.
- Crawford, K., 2021, *Atlas of AI. Power, Politics, and the Planetary Costs of Artificial Intelligence*, Yale University Press, New Haven, London.
- Czichos, H., 2019, *Mechatronik, Grundlagen und Anwendungen technischer Systeme*, 4. Auflage, Springer, Berlin.
- De Villiers, D. E., 2002, Who will bear moral responsibility? in *Communicatio* 28 (1): 16–21.
- De Villiers, D. E., 2006, Prospects of a Christian ethics of responsibility (part 1): An assessment of an American version in *Verbum et Ecclesia*, Vol 27, No 2, angesehen am 28.02.2025, von <https://doi.org/10.4102/ve.v27i2.159>.
- Dieterle, E., Holland, B. & Dede, C., 2021, The Cyclical Effects of Ethical Decisions Involving Big Data and Digital Learning Platforms, in *The Ethical Use of Data in Education: Promoting Responsible Policies and Practices*. Teachers College Press, New York.
- Donges, E., 1977, *Experimentelle Untersuchung und regelungstechnische Modellierung des Lenkverhaltens von Kraftfahrern bei simulierter Straßenfahrt*. Doctoral dissertation, Technische Hochschule Darmstadt.
- Donges, E., 2015, „Fahrerverhaltensmodelle“ in Winner, H., Hakuli, S., Lotz, F. & Singer, C. (Hg.), *Handbuch Fahrerassistenzsysteme, Grundlagen, Komponenten und Systeme für aktive Sicherheit und Komfort*, 3. Auflage, Springer, Wiesbaden, 17–26.
- Eltahan, M. & Elsayed, K., 2024, Enhancing Autonomous Driving By Exploiting Thermal Object Detection Through Feature Fusion in *International Journal of Intelligent Transportation System Research*, volume 22, 146–158.
- Ertel, W., 2016, *Grundkurs Künstliche Intelligenz: Eine praxisorientierte Einführung*, Springer, Wiesbaden, Kindle-Edition.
- Exponent, 2023, Cruise AV SF Incident – Pedestrian Collision, Technical Root Cause Analysis in Quinn Emanuel 2024, *Report on the October 2th Accident*, 111–195.
- Fahrenberg, J., 2017, *Annahmen über den Menschen. Menschenbilder aus psychologischer, biologischer, religiöser und interkultureller Sicht*, 5. Auflage, Asanger, Heidelberg und Kröning.
- Fleischer, C., 2009, *Die Ethik Dietrich Bonhoeffers*, angesehen am 05.08.2024, von <https://www.theomag.de/59/cf17.htm>.
- Fischer, J., 1998, *Handlungsfelder angewandter Ethik. Eine theologische Orientierung*, Kohlhammer, Stuttgart.
- Fischer, J., 2002, *Theologische Ethik. Grundwissen und Orientierung*, Kohlhammer, Stuttgart.
- Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Lütge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P. & Vayena, E., 2018, “AI4People—An Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations”, in *Minds and Machines*. 28, 689–707.
- Föllinger, O., Konigorski, U., Lohmann, B, et al., 2022, *Regelungstechnik, Einführung in die Methoden und ihre Anwendung*, 13. Auflage, VDE, Berlin.
- Frankel, T. C., 2016, “Elon Musk says Tesla’s autopilot is already ‘probably’ better than human drivers”, in *The Washington Post*, angesehen am 21.03.2025, von

- <https://www.washingtonpost.com/news/the-switch/wp/2016/01/11/elon-musk-says-teslas-autopilot-is-already-probably-better-than-human-drivers/>
- Fraichard, T & Asama, H, 2003, Inevitable collision states. A step towards safer robots?, in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, Las Vegas, 388–393. <https://doi.org/10.1109/IROS.2003>
- Gasser, T.M., 2015, Grundlegende und spezielle Rechtsfragen für autonome Fahrzeuge, in Maurer, M., Gerdes, J. C., Lenz, B., & Winner, H. (Hg.), *Autonomes fahren: Technische, rechtliche und gesellschaftliche Aspekte*, Springer, Berlin, 543–574.
- Gerdes, J. C., & Thornton, S. M., 2015, Implementable ethics for autonomous vehicles, in Maurer, M., Gerdes, J. C., Lenz, B., & Winner, H. (Hg.), *Autonomous driving. Technical, legal and social aspects*, Springer, Berlin, 687–706.
- Goodall, N. J., 2014, “Ethical Decision Making during Automated Vehicle Crashes”, *Transportation Research Record*, 2424(1), von <https://doi.org/10.3141/2424-07>, angesehen am 05.04.2025, 58–65.
- Goodall, N. J., 2016, „Away from trolleys and toward risk-management”, in *Applied Artificial Intelligence*, 30(8), 810–821.
- Grimm, P, Keber, T.O. & Zöllner, O. (Hg.), 2019, *Digitale Ethik, Leben in vernetzten Welten*, Reclam, Ditzingen.
- Grunwald, A., 2019, Der unterlegene Mensch, Die Zukunft der Menschheit im Angesicht von Algorithmen, Künstlicher Intelligenz und Robotern, riva, München.
- Gurney, J. K., 2016, „Crashing into the unknown: An examination of crash-optimization algorithms through the two lanes of ethics and law”, in *Albany Law Review*, 79(1), 183–267.
- Haas, N. & Sessler, M.-L., 2022, „Integration moralischer Anforderungen in den agilen Entwicklungsprozess KI-basierter Anwendungen am Beispiel von Scrum in *HMD Prax. Wirtsch.* 59, 667–682.
- Haupt, R., Schmitz, S. (ed.), 2019, *Digitalisierung: Datenhype mit Werteverlust? Ethische Perspektiven für eine Schlüsseltechnologie*, Hänssler, Holzgerlingen.
- Hays, R. B., 1996, *The Moral Vision of the New Testament: Community, Cross, New Creation. A Contemporary Introduction to New Testament Ethics*, HarperCollins, New York.
- Henning, K., 2021, “Artificial Intelligence is a Gamechanger of the Way of Living and Working”, in *The 4th Industrial Revolution from an Ethical Perspective*, Sallux, Amersfoort.
- Huber, W., 2022, *Menschen, Götter und Maschinen. Eine Ethik der Digitalisierung*, C.H. Beck, München.
- Huber, W., 2021, *Dietrich Bonhoeffer: Auf dem Weg zur Freiheit. Ein Porträt*, C.H. Beck, München.
- Huber, W., 2024, *Lebenslauf*, angesehen am 01.02.2025, von <http://www.wolfganghuber.info/leben.html>.
- Huber, W., Meireis, T. & Reuter, H.-R. (Hg.), 2015, *Handbuch der Evangelischen Ethik*, C.H. Beck, München.
- Humble, J. & Farley, D., 2011, *Continuous Delivery: Reliable Software Releases through Build, Test, and Deployment Automation*, Pearson International, Boston.
- IEEE, 2021, *Standard Model Process for Addressing Ethical Concerns during System Design*, IEEE Std 7000-2021, 1–82.

- Janiesch, C., Zschech P. & Heinrich, K., 2021, "Machine learning and deep learning" in *Electron Markets* 31, 685–695.
- Jonas, H., 1979, *Das Prinzip Verantwortung. Versuch einer Ethik für die technologische Zivilisation*, Suhrkamp, Frankfurt am Main.
- Jonas, H., 1984, *The Imperative of Responsibility*, University of Chicago Press, Chicago.
- Kalinna, G., 2021, *Der Mensch als antwortendes Wesen. Gedanken zur gegenwärtigen Verantwortungsethik. Mit einem Vortrag von H. Richard Niebuhr*, TVZ, Zürich.
- Kessler, V., 2019a, „Was ist der Mensch“, *ACF-Seminar Christliche Anthropologie*, Gummersbach.
- Kessler, V., 2019b, „Ein freier Wille?“, *ACF-Seminar Christliche Anthropologie*, Gummersbach.
- Kessler, V., 2025, „Emil Brunner: Der Mensch im Widerspruch“, *ACF-Seminar Christliche Anthropologie*, Gummersbach.
- Kirchschlaeger, P. G., 2021, *Digital Transformation and Ethics, Ethical Considerations on the Robotization and Automation of Society and the Economy and the Use of Artificial Intelligence*, Nomos, Baden-Baden.
- Klaus, G. & Liebscher, H., 1974, *Was ist, was soll Kybernetik?*, Urania, Leipzig.
- Koopman, P., 2024, *Comments on the NHTSA 2024 report*, angesehen am 08.06.2024, von https://www.linkedin.com/posts/philip-koopman-0631a4116_nhtsa-ea22002-investigation-close-out-activity-7189707310976446464--6sc.
- Kriebitz, A., Max, R. & Lütge, C., 2022, "The German Act on Autonomous Driving: why ethics still matters", in *Philosophy & Technology* 35, 1–13.
- Küppers, E. W. U., 2023, *A Transdisciplinary Introduction to the World of Cybernetics: Basics, Models, Theories and Practical Examples*, Springer, Wiesbaden.
- Levine, J., 2019, *Request for Investigation of Deceptive and Unfair Practices in Advertising and Marketing of the "Autopilot" Feature Offered in Tesla Motor Vehicles*, angesehen am 23.08.2024, von <https://www.autosafety.org/wp-content/uploads/2019/07/Center-for-Auto-Safety-Tesla-Autopilot-Letter-to-UT-Atty.-Gen.-FINAL.pdf>.
- Ladd, J., 2019, Computers and Moral Responsibility: A Framework for an Ethical Analysis, in Gould, C. C., 2019, *The Information Web: Ethical and Social Implications of Computer Networking*, Routledge, London.
- Lin, Y., 2003, *How is Ethics of Responsibility Possible?* Dissertation, Freie Universität Berlin, <http://dx.doi.org/10.17169/refubium-13578>.
- Lin, P., 2015, „Why ethics matters for autonomous cars“, in M. Maurer, J. C. Gerdes, B. Lenz, & H. Winner (Hg.), *Autonomes fahren: Technische, rechtliche und gesellschaftliche Aspekte*, Springer, Berlin, 69–85.
- Maurer, M., Lenz, B., Gerdes, J. C., Winner, H. (Hg.), 2015, *Autonomes Fahren. Technische, rechtliche und gesellschaftliche Aspekte*. Springer, Berlin Heidelberg.
- Misselhorn, C., 2019, *Grundfragen der Maschinenethik*, Reclam, Ditzingen.
- Moreno, E. 2024, *Auto Safety Regulator Investigating Tesla Recall of Autopilot*, angesehen am 08.06.2024 von <https://www.nytimes.com/2024/04/26/business/tesla-autopilot-recall-nhtsa-investigation.html>

- Mühling, M., 2012, *Systematische Theologie: Ethik*, Vandenhoeck & Ruprecht, Göttingen.
- National Highway Traffic Safety Administration (NHTSA), 2015, *Critical Reasons for Crashes Investigated in the National Motor Vehicle Crash Causation Survey*, angesehen am 08.06.2024, von <https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812115>.
- National Highway Traffic Safety Administration (NHTSA), 2022, NHTSA Estimates for 2022 Show Roadway Fatalities Remain Flat After Two Years of Dramatic Increases, angesehen am 11.04.2024, von <https://www.nhtsa.gov/press-releases/traffic-crash-death-estimates-2022>.
- National Highway Traffic Safety Administration (NHTSA), 2023, *Safety Recall Report 23E-086*, angesehen am 28.03.2024, von <https://static.nhtsa.gov/odi/rcl/2023/RCLRPT-23E086-7725.PDF>.
- National Transportation Safety Board (NTSB), 2010, *Aircraft Accident Report NTSB/AAR-10 /03*, angesehen am 20.03.2025, von <https://www.nts.gov/investigations/AccidentReports/Reports/AAR1003.pdf>.
- National Transportation Safety Board (NTSB), 2020, *Highway Accident Brief HWY19FH008*, angesehen am 25.03.2024, von <https://www.nts.gov/investigations/AccidentReports/Reports/HAB2001.pdf>.
- Nida-Rümelin, J. & Weidenfeld, N., 2020, *Digitaler Humanismus. Eine Ethik für das Zeitalter der Künstlichen Intelligenz*, Piper, München.
- Niebuhr, H. R., 1963, *The Responsible Self, An Assay in Christian Moral Philosophy*, Harper & Row, New York.
- Niebuhr, H.R., 2021, „Der Sinn der Verantwortung“, in Kalinna, G., *Der Mensch als antwortendes Wesen. Gedanken zur gegenwärtigen Verantwortungsethik. Mit einem Vortrag von H. Richard Niebuhr*, TVZ, Zürich, 37–54.
- Nyholm, S., 2018, „The ethics of crashes with self-driving cars: A roadmap“, in *Philosophy Compass*, 13, angesehen am 26.06.2024, von <https://doi.org/10.1111/phc3.12507>.
- Pannaga, N., Ganesh, N., Gupta, R., 2013, „Mechatronics—an introduction to mechatronics“, in *International Journal of Engineering Research & Technology*, Vol. 2 Issue 8, 128–134.
- Pauder-Studer, H., 2020, *Einführung in die Ethik*, A. Franke, Tübingen.
- Peckham, J., 2021, „The Ethical Implications of 4IR“, in *The 4th Industrial Revolution from an Ethical Perspective*. Sallux, Amersfort.
- Pieper, A., 2017, *Einführung in die Ethik*, A. Franke, Tübingen.
- Pitzke, M., 2009, New York feiert das Hudson-Wunder in *Der Spiegel (online)*, 16. Januar 2009, angesehen am 20.03.2024, von <https://www.spiegel.de/panorama/spektakulaere-airbus-notlandung-new-york-feiert-das-hudson-wunder-a-601587.html>.
- Posocco, L., 2017, „France“, in *The Dieselgate*, Springer International, Basel.
- Poszler, F. & Geisslinger, M., 2021, AI and Autonomous Driving: Key ethical considerations, angesehen am 25.03.2024, von https://ieai.mcts.tum.de/wp-content/uploads/2021/02/ResearchBrief_February2021_AutonomousVehicles_FINAL.pdf.
- Poszler, F., Geisslinger, M., Betz, J. & Lütge, C., 2023, “Applying ethical theories to the decision-making of self-driving vehicles: A systematic review and integration of the literature“, in *Technology in Society*, Volume 75, angesehen am 25.03.2024, von <https://doi.org/10.1016/j.techsoc.2023.102350>.

- Quinn Emanuel, 2024, Report on the October 2th Accident, angesehen am 28.03.2024, von https://assets.ctfassets.net/95kuvdv8zn1v/1mb55pLYkkXVn0nXxEXz7w/9fb0e4938a89dc5cc09bf39e86ce5b9c/2024.01.24_Quinn_Emanuel_Report_re_Cruise.pdf.
- Rasmussen, J., 1983, "Skills, rules, and knowledge; signals, signs, and symbols, and other distinctions in human performance models" in *IEEE Transactions on Systems, Man, and Cybernetics*, vol. SMC-13, 257–266, New York.
- di Rattalma, M. F., 2017, *The Dieselgate*, Springer International, Basel.
- Rausch, A., Broy, M., 2008, „Die V-Modell XT Grundlagen“ in Höhn, R., Höppner, S., *Das V-Modell XT - Grundlagen, Methodik und Anwendungen*, Springer, Berlin, 1–17.
- Reuter, H.-R., 2015, Grundlagen und Methoden der Ethik, in Huber, W., Meireis, T., Reuter, H.-R. (Hg.), *Handbuch der Evangelischen Ethik*, C.H. Beck, München.
- Rosenau, H. (Hg.), 2024, *Dietrich Bonhoeffer – Biografie*, angesehen am 24.08.2024, von <https://www.dietrich-bonhoeffer.net/biografie/>
- Russell, S., Norvig, P., 2010, *Artificial Intelligence: A Modern Approach*, Prentice Hall, Hoboken
- Russell, S., Norvig, P., 2012, *Künstliche Intelligenz, Ein moderner Ansatz*. 3. Auflage, Pearson, München.
- SAE, 2021, *Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles*, angesehen am 02.10.2023, von https://www.sae.org/standards/content/j3016_202104.
- Schoberth, W., 2006, *Einführung in die theologische Anthropologie*, Wissenschaftliche Buchgesellschaft, Darmstadt.
- Searle, J. R., 1980, "Minds, brains, and programs", in *Behavioral and Brain Sciences* 3 (3), 417–457.
- Siddiqui, F. & Merrill, J. B., 2023, "17 fatalities, 736 crashes: The shocking toll of Tesla's Autopilot", angesehen am 11.06.2024, von <https://www.washingtonpost.com/technology/2023/06/10/tesla-autopilot-crashes-elon-musk/>.
- Smith, B., 2013, *Human Error as a Cause of Vehicle Crashes*, angesehen am 25.03.2024, von <http://cyberlaw.stanford.edu/blog/2013/12/human-error-cause-vehicle-crashes>.
- Stadelmann, T., 2019, „Wie maschinelles Lernen den Markt verändert“, in Haupt, R., Schmitz, S. (Hg.), *Digitalisierung: Datenhype mit Werteverlust? Ethische Perspektiven für eine Schlüsseltechnologie*, Hänssler, Holzgerlingen, 67–80.
- Stiller, C., Bachmann, A., Geiger, A., 2015, Maschinelles Sehen, in *Handbuch Fahrerassistenzsysteme, Grundlagen, Komponenten und Systeme für aktive Sicherheit und Komfort*, Springer, Wiesbaden.
- Tesla, 2024, *Model 3 Owner's Manual*, angesehen am 03.02.2024, von https://www.tesla.com/ownersmanual/model3/en_us/.
- Tesla, 2024b, Autopilot – Future of Driving, angesehen am 21.08.2024, von <https://www.tesla.com/autopilot>
- Thadani, T., Lerman, R., Piper, I., Siddiqui F. & Uraizee, I., 2023, *The final 11 seconds of a fatal Tesla Autopilot crash*, angesehen am 25.03.2024, von <https://www.washingtonpost.com/technology/interactive/2023/tesla-autopilot-crash-analysis/>
- Thadani, T., 2023b, *Cruise recalls all its driverless cars after pedestrian hit and dragged*, angesehen am 28.03.2024, von

- <https://www.washingtonpost.com/technology/2023/11/08/cruise-crash-driverless-recall/>
- Thomas, G., 2021, *How to Do Your Case Study*, SAGE, Thousand Oaks.
- Van den Heuvel, S. C., 2024, 'Who is Jesus Christ for us Today?' *Bonhoeffer's Ethics of Responsibility as a Promising Paradigm for Contemporary Public Theology*, ETF, Leuven.
- Van den Heuvel, S. C., 2024b, *Education & Selected Bibliography*, angesehen am 01.02.2025, von <https://www.etf.edu/en/staff/steven-van-den-heuvel/>.
- Veldsman, D. P., 2019, „Embracing the Eye of the Apple: On anthropology, theology and technology“, *HTS Teologiese Studies/Theological Studies* 75(1), angesehen am 10. Oktober 2023, von <https://doi.org/10.4102/hts.v75i1.5498>.
- Verband der Automobilindustrie VDA, 2024, *Auf dem Weg zur Vision Zero: keine schweren Unfälle mehr im Verkehr*, angesehen am 25.03.2024, von <https://www.vda.de/de/themen/innovationen/sicherheit/notbremsassistent>.
- Verhaegen, M., 2021, *Integrating Theological Ethics of Responsibility with Teleological and Deontological Ethics for Embodied AI*, MTh thesis, Evangelische Theologische Faculteit, Leuven.
- Weber, M., 2005, *Gesammelte Politische Schriften, Potsdamer Internet-Ausgabe*, angesehen am 18.11.2023, von <http://www.uni-potsdam.de/u/paed/pia/index.htm>.
- Weber, M., 1919, Politik als Beruf, in *Gesammelte Politische Schriften. Potsdamer Internet-Ausgabe*, angesehen am 18.11.2023, von <http://www.uni-potsdam.de/u/paed/pia/index.htm>.
- Weber-Lewerenz, B., 2022, *Accents of added value in construction 4.0. Ethical observations in dealing with digitization and AI*, Springer, Wiesbaden.
- Wiener, N., 1963, *Kybernetik, Regelung und Nachrichtenübertragung im Lebewesen und in der Maschine*, Econ, Düsseldorf.
- Winkle, T., 2015, „Sicherheitspotential automatisierter Fahrzeuge: Erkenntnisse aus der Unfallforschung“, in Maurer, M., Gerdes, J. C., Lenz, B. & Winner, H. (Hg.), *Autonomes fahren: Technische, rechtliche und gesellschaftliche Aspekte*. Berlin, Heidelberg, Springer, 351–376.
- Winner, H., Hakuli, S., Lotz, F. & Singer, C., 2015, *Handbuch Fahrerassistenzsysteme, Grundlagen, Komponenten und Systeme für aktive Sicherheit und Komfort*, Springer, Wiesbaden.
- Wolff, H. W., 2018, *Anthropologie des Alten Testaments*, 2. Auflage, Gütersloher Verlagshaus, Gütersloh.
- Wolkenstein, A., 2018, “What has the Trolley Dilemma ever done for us (and what will it do in the future)? On some recent debates about the ethics of self-driving cars,” in *Ethics and Information Technology*, 20, 163–173.
- Yin, R. K., 2018, *Case Study Research and Applications: Design and Methods*, SAGE, Thousand Oaks.
- Zimmerling, P., 2021, „Einführung“ in Bonhoeffer, D. & Zimmerling, P. (Hg.), *Freiheit zum Leben: Ausgewählte Texte zur Ethik*, Brunnen, Gießen, 7–17.

9 ANHANG

9.1 Abkürzungen

AD Autonomous Driving (Autonomes Fahren)

AI Artificial Intelligence (siehe KI)

AV Autonomous Vehicle (Autonomes Fahrzeug)

IEEE Institute of Electric and Electronic Engineers

KI Künstliche Intelligenz

ML machine learning (maschinelles Lernen)

9.2 Abbildungen

Abbildung 1 Schema ethische Entscheidungsfindung (Bleisch et.al., 2021)	16
Abbildung 2: Multiple-Case Study nach Yin (2018)	17
Abbildung 3: Leistungsstufen von qualifizierten menschlichen Bedienern (Rasmussen 1983:258)	26
Abbildung 4: Erweitertes Modell der Fahraufgabe (Donges 2015:19)	29
Abbildung 5: Blockschaltbilder einer Regelung und einer Steuerung (Czichos 2019:110)	32
Abbildung 6: Kybernetisches System mit doppelter Rückkopplung (Küppers 2023: 27)	33
Abbildung 7: Grundprinzip mechatronischer Systeme (Czichos 2019:23)	35
Abbildung 8: Vorgehensmodell V-Modell (Rausch & Broy 2008:17).....	39
Abbildung 9: Konzepte des Machine Learnings (Janiesch, Zschech & Heinrich 2021:687)	43
Abbildung 10: Straßenkarte mit Positionen (NTSB 2020:2).....	46
Abbildung 11: Unfallort aus dem Bericht der NTSB (2020:5), ergänzt um Fahrtrichtungen.	47
Abbildung 12: Teslas Frontkamera (Thadani u.a., 2023).....	50
Abbildung 13: Unfallort in San Francisco (Exponent 2023:24), ergänzt um Positionen	51
Abbildung 14: Statistische Unfallursachenverteilung (Winkle 2015:396)	65
Abbildung 15: Die Ethik Bonhoeffers (Biewald & Beckmann 2007)	80

9.3 Tabellen

Tabelle 1: Stakeholder im Fall Tesla	90
Tabelle 2: Stakeholder im Fall Cruise	91