

I.

Autonomes Fahren und Unfalldilemmata: Ethischer Problemhorizont und Relevanz

Der erste Teil des Buches widmet sich einer Einführung in das Phänomen des autonomen Fahrens, dessen überblicksartiger Darstellung und Problematisierung sowie der Verortung der zentralen Fragestellung innerhalb des Forschungsdiskurses, zu dessen Weiterentwicklung die Forschungsarbeit beitragen will. In zwei Kapiteln wird auf diese Weise das erste Teilziel der vorliegenden Untersuchung erarbeitet. Besonderes Augenmerk wird dabei auf die Evaluation der Bedeutung und Relevanz von Unfallszenarien sowohl in praktischer als auch theoretischer Hinsicht gelegt.

Als Annäherung an das Anwendungsproblem werden in Kap. 2 Agenda, Ziele und Herausforderungen im weiteren Kontext autonomer Fahrsysteme beschrieben. In Kap. 2.1 werden selbstfahrende Fahrzeuge als Treiber der Mobilitätswende in ihrer technischen, wirtschaftlichen und gesellschaftlichen Dimension konturiert. Ausgehend von einer begrifflichen Reflexion über den zugrundeliegenden Autonomiebegriff werden Motivatoren, Vision und evolutionäre Aspekte der Fahrautomatisierung erläutert. Darauf aufbauend werden in Kap. 2.2 die spezifischen Herausforderungen dargestellt, die

I. Autonomes Fahren und Unfalldilemmata

durch Zusammenhänge und Wechselwirkungen zwischen den vielfältigen Anforderungen an die neue Technologie einerseits und ambivalente Wirkungen derselben andererseits entstehen. Im Anschluss an eine Bestandsaufnahme des gegenwärtigen technischen und regulatorischen Entwicklungsstands wird schließlich kritisch reflektiert, inwiefern das autonome Fahren die an seine Vision geknüpften Erwartungen erfüllen kann.

In Kap. 3 wird der ethische Diskurs des autonomen Fahrens in systematischer Weise rekonstruiert sowie die Bedeutung aufgezeigt, die Unfalldilemmata für theoretische und praktische Fragen des Anwendungskontextes besitzen. Zunächst erfolgt in Kap. 3.1 eine überblicksartige Skizze ethischer Problemfelder und Diskurse, in deren Rahmen die relevanten Literaturströmungen identifiziert, gegeneinander abgegrenzt und in ihrer thematischen Tiefe erläutert werden. Besondere Aufmerksamkeit erfährt dabei die Problematik unvermeidbarer Unfallsituationen. In Kap. 3.2 wird vertiefend auf die Relevanz von Dilemma-Szenarien für die Gestaltung von Unfallalgorithmen eingegangen, welche das Fahrverhalten autonomer Fahrzeuge im Fall einer unabwendbaren Kollision steuern. Anhand der relevanten Forschungsliteratur wird ausführlich argumentiert, dass Entscheidungsprobleme mit dilemmatischen Strukturen im Kontext autonomer Fahrsysteme sowohl theoretisch möglich als auch praktisch existent sind. Zudem wird begründet, weshalb Unfallalgorithmen nicht normierbar sind und welche gesellschaftliche und technische Relevanz sie aufgrund dessen entfalten. In Kap. 3.3 werden schließlich die Ergebnisse von Teil I in einem ersten Zwischenergebnis zusammengefasst.

2. Das Phänomen ›Autonomes Fahren‹: Agenda, Ziele und Herausforderungen

2.1 Selbstfahrende Fahrzeuge als Treiber der Mobilitätswende

2.1.1 Der Autonomiebegriff im Kontext technischer Systeme

Selbstfahrende Fahrzeuge fungieren als die zentralen Treiber der anvisierten Mobilitätswende. Gemäß ihrer Vision entbinden sie Menschen von deren Fahraufgaben und schaffen Potenziale für mehr Sicherheit, Effizienz und Teilhabe. Auf diese Weise transformieren sie das Wesen der Mobilität, wie wir sie heute kennen. Die revolutionäre Dynamik der neuen Technologie liegt in einer ihrer zentralen Eigenschaften begründet, die ihr auch den Namen gibt: ihrer Autonomie. Der Begriff der Autonomie stammt aus dem Altgriechischen (*αὐτονομία* bzw. *autonomía*), wo er so viel wie ›Eigengesetzlichkeit‹ bedeutet. Im Zuge einer Auseinandersetzung mit komplexen praktischen Fragen wird allerdings schnell deutlich, dass der Autonomiebegriff im Kontext verschiedener Forschungsdisziplinen jeweils unterschiedliche, fachspezifische Bedeutungen aufweist. Es erscheint daher zweckmäßig, eine wissenschaftliche Abhandlung über autonome Fahrzeuge mit einer Klärung des Autonomieverständnisses zu beginnen, das hier zugrunde liegt.

Die Bezeichnung ›autonomes Fahrzeug‹ als Anwendungsbeispiel eines autonomen Systems bezieht sich auf den in der Informations-technik verwendeten Autonomiebegriff. Aus anwendungsbezogener und operationeller Sicht zeichnen sich autonome Systeme zunächst dadurch aus, dass sie unabhängig von direktem menschlichen Eingreifen operieren können. Eine solche triviale Auffassung von Autonomie ist jedoch nur für vollständig kontrollierbare Umgebungen, beispielsweise eine intelligente digitale Fabrik (*smart factory*), plausibel. Sobald das Umfeld komplexer wird, muss der Begriff erweitert werden. Autonomie ist dann als die Fähigkeit eines Systems zu interpretieren, das ohne menschliches Eingreifen eine Handlungsoption

wählt: »When such machines are called ›autonomous‹, it is meant that they are able to choose by themselves, without human intervention, the appropriate course of action in the manifold situations they encounter.« (Totschnig, 2020, S. 2474) Diese Verwendungsweise des Autonomiebegriffs setzt voraus, dass der Agent sich bei seiner Wahl an einem Ziel oder einer Nutzenfunktion orientiert, die extern festgelegt ist. In eine ähnliche Richtung weist auch die Definition von Hilgendorf (2017b, S. 47): »Unter einem autonomen technischen System soll ein System verstanden werden, das auf Probleme unabhängig von menschlichem Input situationsangemessen und somit ›intelligent‹ reagieren kann.« Bradshaw et al. (2013) betonen, dass das Design technischer Systeme in der Regel auf einen bestimmten Kompetenzbereich limitiert ist. Die Europäische Gruppe für Ethik der Naturwissenschaften und der neuen Technologien erläutert in ihrer »Erklärung zu künstlicher Intelligenz, Robotik und ›autonomen Systemen« (2018, S. 10–11):

Allerdings hat sich der Begriff des ›autonomen‹ Systems zur Bezeichnung eines Höchstmaßes an Automatisierung und maximaler Unabhängigkeit vom Menschen im Sinne einer operativen und entscheidungsbezogenen ›Autonomie‹ in der wissenschaftlichen Literatur und der öffentlichen Debatte sehr stark durchgesetzt.

Eine stärker technisch orientierte Konzeption definiert Autonomie als »die Fähigkeit eines rechnergestützten Systems, selbstständig aus Daten, die sowohl durch Programmierung vorgegeben als auch aus der Umwelt mittels Sensoren gewonnen sein können, zielgerichtete Pläne und Aktionen zu generieren.« (Deutscher Bundestag, 2020, S. 31) Eine solche Perspektive auf das Konzept der Autonomie setzt vor allem zwei zentrale Designelemente voraus. Dies ist einerseits die Implementierung einer Belohnungs- oder Nutzenfunktion, mittels derer das System seine Aktionen hinsichtlich gewünschter Zustände abgleichen kann. Andererseits sind Mechanismen des maschinellen Lernens unverzichtbar, die es dem System ermöglichen, sich an veränderte Umgebungsbedingungen anzupassen und sein Verhalten aufgrund von bisherigen, als unerwünscht bewerteten Aktionen zu optimieren.

Autonomie tritt grundsätzlich in verschiedenen Ausprägungen entlang eines Kontinuums auf (vgl. Etzioni & Etzioni, 2017, S. 408–409). Eine einflussreiche Auffassung, die sich an der Schnittstelle von informationstechnischer und (maschinen-)ethischer Perspekti-

ve bewegt, liefert der deskriptive Ansatz von Floridi und Sanders (2004, S. 357–364). Sie schreiben künstlichen Systemen genau dann eine Selbststurzprünglichkeit ihrer Handlungen zu, wenn diese drei relevanten Kriterien erfüllen. Neben der Fähigkeit zur Interaktivität (*interactivity*) und Adaption (*adaptability*) besteht eines dieser Kriterien in der Autonomie (*autonomy*):

Autonomy means that the agent is able to change state without direct response to interaction: it can perform internal transitions to change its state. So an agent must have at least two states. This property imbues an agent with a certain degree of complexity and independence from its environment. (Ebd., S. 357)

Die im Kontext künstlicher Systeme verwendeten Autonomiekonzeptionen unterscheiden sich deutlich von anspruchsvolleren philosophischen Auffassungen. Autonomie im Sinne Immanuel Kants, der sie als »Beschaffenheit des Willens, dadurch derselbe ihm selbst (unabhängig von aller Beschaffenheit der Gegenstände des Wollens) ein Gesetz ist« (1900ff, GMS, AA 04: 440.16-18)², interpretiert, wäre ausschließlich im Rahmen der Realisierung starker Künstlicher Intelligenz denkbar. Die Erschaffung einer menschenebenbürtigen Denkfähigkeit, die die Ziele ihres Handelns selbst definiert, ist zum gegenwärtigen Zeitpunkt allerdings lediglich eine Zukunftsvision; bei derzeitigen Systemen sind die relevanten Handlungsziele extern festgelegt.

Analog zur Vielschichtigkeit des ihnen zugrundeliegenden Autonomiebegriffs stellen autonome Systeme ein großes Forschungs- und Entwicklungsvorhaben an der Schnittstelle verschiedener Ingenieur-, Sozial- und Geisteswissenschaften dar. Die Konzeption autonomer Fahrzeuge ist dabei kein Selbstzweck, sondern durch spezifische praktische Erwartungen motiviert. Diese werden im nachfolgenden Unterkapitel skizziert.

2.1.2 Motivatoren des autonomen Fahrens

Die Vision einer automatisierten Mobilität ist verbunden mit der Erwartung, auf diese Weise die sozialen Kosten motorisierter Individualmobilität entscheidend zu senken. Die Idee des autonomen

² Kants *Grundlegung zur Metaphysik der Sitten* wurde erstmals 1785 veröffentlicht.

Fahrens begegnet den essenziellen Herausforderungen moderner Industriegesellschaften anhand dreier zentraler Kriterien, die die Mobilitätsqualität maßgeblich beeinflussen (vgl. Beiker, 2012, S. 1149–1152).

1. Sicherheitspotenzial

Primärer Motivator der Verkehrsautomatisierung ist ihr Potenzial, durch die Reduzierung bzw. Eliminierung menschlichen Fahrversagens mittel- bis langfristig die Zahl der Verkehrsunfälle signifikant zu reduzieren (vgl. Beiker, 2012, S. 1149–1150; Crew, 2015, o. S.; Fleetwood, 2017, S. 532). Wie ältere Studien belegen, gilt der ›Faktor Mensch‹ in der Forschung zur Verkehrssicherheit schon seit Langem als primäre Risikoquelle. Veränderungen im Fahrverhalten bieten dementsprechend eine vielversprechende Möglichkeit, um Schäden im Kontext von Verkehrsunfällen zu reduzieren: »Human factors are far more important than engineering factors. Among human factors, driver behavior (what the driver *chooses to do*) has much greater influence on safety than driver performance (what the driver *can do*).« (Evans, 1996, S. 784) Gemäß Datenerfassung des Statistischen Bundesamtes waren 88 % der Unfälle mit Personenschäden³ im Jahr 2021 auf menschliche Fahrfehler zurückzuführen. Die häufigsten Fehler geschahen dabei beim Abbiegen, Wenden, Rückwärtsfahren, Ein- bzw. Anfahren, bei Nichtbeachtung der Vorfahrtsregeln, aufgrund von ungenügendem Abstand oder nicht angepasster Geschwindigkeit (vgl. Statistisches Bundesamt, 2022, S. 50). Analog nennt eine in den USA durchgeführte Studie inkorrekt erkannte Fahrsituationen aufgrund von Unaufmerksamkeit oder Ablenkung als häufigste Unfallursache, gefolgt von schlechten Fahrentscheidungen wie zu hoher Geschwindigkeit, Fehleinschätzung anderer Verkehrsteilnehmer, aggressive Fahrweise und schließlich Fehlern im Fahrverhalten wie Überkompensation oder mangelnde Richtungskontrolle (vgl. National Highway Traffic Safety Administration, 2008, S. 24–25). Auch der Konsum berausender Mittel spielt eine

³ Bei Unfällen mit Personenschäden handelt es sich um solche Verkehrsergebnisse, bei denen Personen verletzt oder getötet werden, wobei die Höhe des Sachschadens irrelevant ist (vgl. Statistisches Bundesamt, 2023).

nicht unerhebliche Rolle.⁴ Die Ergebnisse einer neueren amerikanischen Studie decken sich mit diesen Erkenntnissen: Selbstfahrende Fahrzeuge sind insbesondere auf Schnellstraßen und bei Nebel sicherer als von Menschen gesteuerte Autos, nicht hingegen auf Landstraßen, bei Dämmerung und bei Abbiegemanövern (vgl. Abdel-Aty & Ding, 2024; Beck, 2024).

Die Idee des autonomen Fahrens setzt an ebendieser Fehleranfälligkeit menschlichen Fahrverhaltens an, um diejenigen zu schützen, die selbst Quelle hoher Risiken sind:

It is the behavior of those whose lives are at stake in traffic that most influences risk in traffic. The least safe vehicle driven on the least safe road by some drivers poses far less risk than the safest vehicle driven on the safest road by other drivers. (Evans, 2008, S. 1)

Würden automatisierte bzw. autonome Systeme im Sinne eines defensiven Fahrstils programmiert, sodass sie die Fahrzeugumgebung ständig überwachen, im Bedarfsfall schnell reagieren können und dabei stets die Verkehrsregeln beachten, ließen sich fahrerbezogene Unfallursachen potenziell minimieren bzw. im Fall vollständig autonomer Fahrzeuge sogar gänzlich eliminieren (vgl. Goodall, 2020, S. 1). Auch wenn sich eine genaue Beizifferung des positiven Effekts⁵ derzeit noch nicht ausreichend mit harten Fakten untermauern lässt, so geben statistische Auswertungen jedoch Anlass zu der Annahme, dass ein solcher – zumindest in gewissem Maße – eintreten könnte. Die Zahlen polizeilich erfasster Unfälle aller Schweregrade sind in den letzten Jahren tendenziell rückläufig (vgl. Statistisches Bundesamt, 2022, S. 44), wobei ein nicht unerheblicher Anteil dieses Trends dem Einsatz automatisierter Teifunktionen im Fahrbetrieb zugeschrieben werden kann (vgl. Anderson et al., 2016, S. 14–16).⁶

-
- 4 Der Anteil unter Alkoholeinfluss begangener Fahrfehler an der Gesamtzahl der Unfälle mit Personenschäden lag im Jahr 2021 in Deutschland zwar nur bei 3,3 % (vgl. Statistisches Bundesamt, 2022, S. 50). Schlüsselt man allerdings die Ursachen von Unfällen mit tödlichem Ausgang weiter auf, so zeichnete Alkohol- oder Drogenkonsum beispielsweise im Jahr 2011 für mehr als 39 % der Fälle in den USA verantwortlich (vgl. Anderson et al., 2016, S. 16).
 - 5 Neben einer Reduzierung der Personenschäden hat eine erhöhte Verkehrssicherheit auch wirtschaftliche Auswirkungen, beispielsweise indem die Kosten für Versicherung, Reparaturen und Administration wegfallen bzw. sinken.
 - 6 In einer kürzlich veröffentlichten Studie dokumentieren Abdel-Aty und Ding (2024), dass mit Fahrerassistenzsystemen operierende Autos in vielen Szenarien

Die Ergebnisse einer vergleichenden Studie der Unfallraten von konventionellen Fahrzeugen und solchen, die im autonomen Modus operieren, legen auf Basis der aktuellen Datenlage nahe, dass autonome Fahrzeuge in weniger Unfälle aller Schweregrade involviert sein könnten (vgl. Blanco et al., 2016, S. i–iv). Während frühe Testfahrten diese These noch zu stützen schienen,⁷ sorgten in den letzten Jahren vermehrt tragische Unfälle, welche die bis dato technische Unreife der eingesetzten Fahrsysteme schonungslos aufzeigen, für öffentliches Aufsehen.⁸ Inwiefern autonome Fahrzeuge tatsächlich zu mehr Verkehrssicherheit beitragen, wird an anderer Stelle in diesem Buch kritisch diskutiert (siehe Kap. 2.2.3). Hier soll das (vermeintliche) Sicherheitspotenzial lediglich in seiner Funktion als primärer Motivator des autonomen Fahrens verstanden werden.

2. Effizienzsteigerung

Positive Wirkungen des autonomen Fahrens erhofft man sich auch in Sachen Zeit- und Verkehrseffizienz. Neben einer allgemeinen Steigerung der Lebensqualität für Personengruppen, die häufig im Verkehr unterwegs sind, hätte dies vor allem vorteilhafte ökonomische Auswirkungen. Zum einen würden die Beförderten durch die (vollständige) Automatisierung der Fahraufgabe von (aktiven) Fahrern zu (passiven) Passagieren, die ihre Zeit an Bord freier gestalten können. So würde es etwa Berufspendlern möglich, Transferzeiten produktiv als Arbeitszeit zu nutzen (vgl. Brändle & Grunwald, 2019, S. 282; Gurney, 2015, S. 192–193; KPMG LLP, Center for Automotive Research, 2012, S. 29). Zum anderen verspricht die Vernetzung und zentrale Koordination der Fahrzeuge eine effizientere Steuerung des Verkehrsflusses (vgl. Friedrich, 2015, S. 339–349). Von den verkürzten Transferzeiten durch weniger Staus würden nicht nur Menschen auf dem Weg zum Arbeitsplatz profitieren, sondern auch Lieferfahrzeuge des Gütertransports. Freigesetzte Optimierungspotenziale für

in weniger Unfälle verwickelt sind als menschengesteuerte Fahrzeuge, nicht jedoch in allen.

7 Googles selbstfahrender Prototyp legte zwischen Februar und Oktober 2015 knapp 3 Millionen Testkilometer mit lediglich geringfügigen Zwischenfällen zurück, welche er jedoch in keinem Fall selbst verursachte (vgl. Gurney, 2015, S. 188; Hulverscheidt, 2015).

8 Beispielhafte prominente Unfälle werden in Kap. 2.2.2 beschrieben.

logistische Prozesse könnten dazu beitragen, die gesamtwirtschaftliche Produktivität zu steigern.

Weiterhin birgt die neue Art der Beförderung das Potenzial, freie Kapazitäten in Fahrzeugen für innovative Geschäftsmodelle einzusetzen, ob als Shuttles, Taxis oder gemeinsam genutzte Fahrzeuge. Insbesondere im Bereich Carsharing ist ein aufsteigender Trend zu beobachten (vgl. Bagloee et al., 2016, S. 289; Fagnant & Kockelman, 2018, S. 147–156; Gogoll & Müller, 2017, S. 685; Gurney, 2015, S. 194; Lenz & Fraedrich, 2015, S. 184–189). Sollte sich dieser hinreichend etablieren, könnte er sogar eine Auflösung der Grenzen zwischen Individualmobilität und öffentlichem Verkehr bewirken und so das Tempo der Transformation bestehender Mobilitätsstrukturen weiter beschleunigen (vgl. Beiker, 2015, S. 204–206; Lenz & Fraedrich, 2015, S. 189–192).

Auch in städtebaulicher Hinsicht tun sich durch das automatisierte Fahren im Zuge einer Neugestaltung von Parkflächen (vgl. Bennett, 2022, S. 197) und einer zentralen Koordination von deren Auslastung neue Perspektiven auf. So könnten städtische Parkraumflächen künftig für alternative Nutzungszwecke zur Verfügung stehen, beispielsweise zur Schaffung neuen Wohnraums oder für Grünflächen (vgl. Anderson et al., 2016, S. 25–27; Sparrow & Howard, 2017, S. 212). Indem Emissionen, Energie- und Kraftstoffverbrauch durch eine zentrale, intelligente Routenplanung und vorausschauende, sparsame Fahrweise optimiert werden, sind nicht zuletzt auch positive Effekte für die Umweltbilanz zu erwarten (vgl. Anderson et al., 2016, S. 28–38; Bennett, 2022, S. 197–198; Brändle & Grunwald, 2019, S. 292; Gurney, 2015, S. 193–194; Lim & Taeihagh, 2018, S. 5; Lin, 2013a). Der Verkehrsautomatisierung wird in diesem Sinne gar eine führende Rolle zugeschrieben, um den Herausforderungen des Klimawandels zu begegnen (vgl. Hula et al., 2018, S. 91–93).

3. Veränderte Mobilitätsbedürfnisse

Nicht zuletzt thematisiert das automatisierte bzw. autonome Fahren die sich stetig im Wandel befindenden Mobilitätsgewohnheiten und -bedürfnisse heutiger und zukünftiger Generationen. Es treibt Veränderungen in intergenerationalen Einstellungen gegenüber Mobilität und Fahrzeughaltung weiter voran: Die Generation der Babyboomer empfand den Erwerb der Fahrerlaubnis und des ersten eigenen Autos noch als Inbegriff persönlicher Freiheit und eines gewissen

sozialen Status. Nun haben sich die Prioritäten der ständig vernetzten, jüngeren Generationen in Richtung völliger Flexibilität – sogenannter *mobility on demand* – verschoben (vgl. KPMG LLP, Center for Automotive Research, 2012, S. 7; Pavone, 2015, S. 400–402). Diese Bedürfnistransformation steht im Zeichen eines seit einigen Jahren zu beobachtenden Trends, der Ausdruck sich verändernder Konsumgewohnheiten ist: ›Nutzen statt besitzen‹ oder ›Zugang statt Besitz‹ lauten die Devisen, die sich als ökonomisches und soziales Phänomen unter dem Begriff ›Sharing-Economy‹ längst einen Namen gemacht haben (vgl. Sundararajan, 2017).

Das Konzept selbstfahrender Fahrzeuge stimuliert diese neuen Prioritäten, indem es die traditionelle Abhängigkeit zwischen ›Mobil-Sein‹ und Fahreignung bzw. Fahrfähigkeit auflöst. Höherstufig automatisierte Fahrzeuge operieren unabhängig von menschlichem Eingreifen und ermöglichen dadurch mobilitätseingeschränkten Personengruppen, z. B. Menschen mit körperlichen und geistigen Beeinträchtigungen,⁹ aber auch älteren Personen und Kindern, Zugang zu individueller Mobilität. Damit geht ein erheblicher Zuwachs an Lebensqualität durch mehr persönliche Unabhängigkeit und soziale Teilhabe einher (vgl. Beiker, 2012, S. 1151–1152; Brändle & Grunwald, 2019, S. 282–283; Gurney, 2015, S. 193; Hansson et al., 2021, S. 1398; Howard, 2013; Mladenovic & McPherson, 2016, S. 1137; Owens et al., 2019).

Hiermit sind die Ziele des autonomen Fahrens grob umrissen. Doch wie gestaltet sich deren Autonomie eigentlich aus technischer Sicht und vor allem im Zusammenspiel mit menschlichen Fahrern bzw. Insassen? Das folgende Unterkapitel illustriert, wie die Idee der Fahrzeugautomatisierung im Verlauf der letzten sechzig bis achtzig Jahre schrittweise durch verschiedene Systemkomponenten realisiert wurde. Zudem wird das sogenannte Stufenmodell erläutert, anhand dessen der rote Faden einer immer weniger an menschliche Fahrzeugsteuerung gebundenen Mobilität in den kommenden Jahren weitergesponnen werden wird.

9 Für eine Diskussion spezifischer ethischer und rechtlicher Fragen in diesem Kontext siehe Bradshaw-Martin und Easton (2014).

2.1.3 Von Informanten über Assistenten zu Automaten: Evolution und Stufenmodell der Fahrautomatisierung

Die Idee, die menschliche fahrzeugführende Person bei ihren Fahraufgaben zu unterstützen und von diesen zu entlasten, ist schon seit beinahe einem Jahrhundert Gegenstand technischen Innovationsbestrebens. Bereits in den 1940er-Jahren wurden erste grundlegende Bausteine der Fahrunterstützung konzipiert und im Laufe der folgenden Jahrzehnte mit dem Automatikgetriebe (1940), der Servolenkung (1952) und dem Bremskraftverstärker (1955) erprobt, bevor das Konzept der Fahrerassistenzsysteme in den 1960er-Jahren endgültig Einzug in die Forschungsagenden hielt (vgl. Beiker, 2012, S. 1146–1147). In der Folge wurden mechanische und elektronische Komponenten intensiv weiterentwickelt und zur Marktreife gebracht. Diese frühen Assistenzsysteme übernahmen primär Aufgaben in der Regelung der Fahrdynamik, insbesondere der Lenk- und Fahrstabilität, z. B. in Form des Antiblockiersystems ABS (1978) und des Elektronischen Stabilitätsprogramms (ESP) (*Electronic Stability Control*) (1995), mit dem sich bahnbrechende Erfolge in Bezug auf die Fahrsicherheit einstellten (vgl. Verband der Automobilindustrie (VDA) e.V., 2023).

Durch technologischen Fortschritt vor allem in der Perzeption des verkehrlichen Umfelds entwickelten sich aus reinen Assistenzfunktionen bald komplexere Systeme, die spezifische Fahrfunktionen teilautomatisiert ausführen. Als erste das Fahrzeugumfeld erfassende Technologie kamen Ultraschallsensoren in der ersten Hälfte der 1990er-Jahre in der aktiven Parkassistenz zum Einsatz. Im Zusammenwirken mit erstmals eingesetzten Radarsensoren ermöglichten sie die Realisierung des Abstandsregeltempomats (*Adaptive Cruise Control (ACC)*), durch den sich das Fahren im gebundenen Verkehr teilautomatisieren ließ und eine Vielzahl von potenziell gefährlichen Situationen bereits in der Entstehung vermieden werden konnte. Dies markierte einen Meilenstein in der Geschichte der Fahrerassistenz. Die zunehmende technologische Reife von Ultraschall- und Radarsensorik sowie ein Durchbruch in der Kameratechnologie machten seit der Jahrtausendwende die Entwicklung komplexerer Systeme möglich, die verschiedene Technologien integrieren. Mithilfe der über Sensorfusion verknüpften Daten unterschiedlicher Sensorsorten gelang die Erzeugung von Fahrzeug-Umfeldmodellen mit

stark verbesserten Präzisionsgraden, womit eine wichtige Voraussetzung für das automatisierte Fahren geschaffen war (vgl. Bengler et al., 2014, S. 6–7).¹⁰

Während traditionelle Fahrerassistenzsysteme als automatisierte Teilkünste der fahrzeugführenden Person lediglich vorübergehend in spezifischen Anwendungsfällen assistieren, übernimmt beim höherstufig automatisierten Fahren ein Fahrroboter (nahezu) alle Fahraufgaben. Die Basisarchitektur gegenwärtig entwickelter Systeme der Fahrzeugautomatisierung umfasst im Wesentlichen drei Komponenten: Perzeption, Prädiktion und Aktion. Die Hardware besteht aus einer Kombination fortschrittlicher Sensoren (Stereokameras, Radar unterschiedlicher Reichweite, Laser, GPS, Ultraschall), die es ermöglichen, Strukturen und Objekte in der Umgebung des Fahrzeugs sowie dessen Position zu erkennen. Diese Sensoren wirken in Verbindung mit Aktuatoren, Steuergeräten und integrierten Software-Algorithmen, welche die produzierten Daten verarbeiten und durch Ansteuerung der Aktuatoren die Aktionen des Fahrzeugs determinieren (vgl. Beiker, 2012, S. 1147; KPMG LLP, Center for Automotive Research, 2012, S. 10). Hierbei spielt die Integration von Anwendungen, welche auf Künstlicher Intelligenz basieren, eine große Rolle; zentrale Komponenten bei der Entwicklung autonomer Fahrsysteme sind vornehmlich neuronale Netze und datengestützte *Deep-Learning*-Techniken des maschinellen Lernens.¹¹

-
- 10 Diese Fusionstechnik kommt beispielsweise beim Stauassistenten zum Einsatz, wo Radar- und Kamerasensoren zusammenwirken, um die Funktionalitäten des ACC und des Spurhalteassistenten für den Einsatz im gebundenen Verkehr zu integrieren.
 - 11 Das in diesem Zusammenhang angewandte Konzept Künstlicher Intelligenz begreift diese im Sinne der schwachen KI-Hypothese als Systeme, die in der Lage sind, menschliche Intelligenz in Bezug auf spezifische, genau definierte Aufgaben zu simulieren, zu erweitern oder mit ihr zu konkurrieren (vgl. Vallor & Bekey, 2017, S. 339–340). Ein zentraler Aspekt bei der Entwicklung derartiger Systeme sind Techniken maschinellen Lernens, deren Realisierung über neuronale Netze (*neural networks*) erfolgt, welche dem Aufbau des menschlichen Gehirns nachempfunden sind und Eingabedaten in mehreren komplexen Schichten (*layers*) via *Deep-Learning*-Techniken (vgl. Reed et al., 2021, S. 784) zu kontrollierten Aktionen verarbeiten: »[...] the network gradually ›learns‹ from repeated ›experience‹ (multiple training runs with input datasets) how to optimize the machine's ›behavior‹ (outputs) for a given kind of task.« (Vallor & Bekey, 2017, S. 340).

Das autonome Fahren stellt das Endziel einer technologischen Entwicklungsagenda dar, die sicherheitskritische Fahrfunktionen schrittweise automatisiert. Die Evolution der Fahrautomatisierung kann verstanden werden als eine Agenda zunehmender Systemintegration; sie führt von passiven Warn- und Informationssystemen über assistiertes hin zu automatisiertem Fahren und schließlich zum autonomen, selbstfahrenden Fahrzeug (vgl. Beiker, 2012, S. 1147–1148). Experten in Forschung und Entwicklung unterscheiden daher verschiedene Level der Automatisierung, anhand derer sich bestehende und zukünftige Fahrzeugsysteme einordnen lassen. Grundlage der Kategorisierung ist eine deskriptive Taxonomie, die Teil der erstmals 2014 von der SAE International (ehemals »Society of Automotive Engineers«) veröffentlichten Norm J3016 ist. Anhand funktionaler Mindestanforderungen an das jeweilige System werden dabei sechs Stufen beschrieben, die sich an den spezifischen Rollen orientieren, die menschlichen Nutzern einerseits und dem automatisierten Fahrsystem andererseits im Hinblick auf die dynamische Fahraufgabe¹² zukommen. Bei der Einstufung des Automatisierungsgrades werden dabei nur solche Systeme berücksichtigt, die sich dauerhaft auf die dynamische Fahraufgabe oder Teile davon auswirken und nicht nur kurzzeitig in potenziell gefährlichen Situationen aktiviert werden, wie beispielsweise ein Notbremsassistent (vgl. SAE On-Road Automated Vehicle Standards Committee, 2014, S. 1–2).¹³

-
- 12 Der Begriff ›dynamische Fahraufgabe‹ (*dynamic driving task*) beschreibt die Kontrolle des Fahrzeugs. Er umfasst die operativen (Lenken, Bremsen, Beschleunigen etc.) und taktischen (Reagieren auf Ereignisse) Aspekte der Fahraufgabe, wohingegen er den strategischen Aspekt (Bestimmen von Zielen und Wegpunkten) außer Acht lässt (vgl. SAE On-Road Automated Vehicle Standards Committee, 2014, S. 2).
- 13 Hinsichtlich der Fokussierung auf spezifische Aspekte existieren alternative Klassifizierungen der Automatisierungsgrade des automatisierten Fahrens. So liegt der Schwerpunkt bei der vom VDA konzipierten Version, die inhaltlich deckungsgleich ist mit der Taxonomie der SAE, auf der Beschreibung der Aufgabenteilung zwischen menschlichen Fahrern und dem System in laienverständlicher Sprache. Dabei werden bereits Systeme der Stufe 3 als hochautomatisiert, 4 entsprechend als vollautomatisiert und 5 als fahrerlos bezeichnet. Auch die amerikanische National Highway Traffic Safety Administration bietet ein alternatives Schema an, das die Level 4 und 5 zusammenfasst. Der Fokus liegt dabei auf dem Anteil, den automatisierte Komponenten im Hinblick auf sicherheitskritische Funktionen des Fahrzeugs innehaben (vgl. National Highway Traffic Safety Administration, 2013, S. 4–5).

Die einzelnen Level des Stufenmodells implizieren keine zwangsläufige Reihenfolge der Markteinführung; jedoch lässt der inkrementelle Entwicklungsprozess automatisierter Fahrfunktionen die verwendeten Technologien reifen und legt somit den Grundstein für das Erreichen der jeweils nächsten Stufe (vgl. Bengler et al., 2014, S. 12). Die Automatisierungslevel sind wie folgt charakterisiert:

Level 0 (No automation): Der Fahrer hat jederzeit die Kontrolle über das Fahrzeug, i. e. er führt dauerhaft alle dynamischen Fahraufgaben aus, wobei ihm Warn- und Informationssysteme wie Spurhalte- und Notbremsassistenten Hilfestellung geben.

Level 1 (Assistiertes Fahren/driver assistance): Eine oder mehrere spezifische Kontrollfunktionen des Fahrzeugs, die unabhängig voneinander operieren, sind automatisiert. Sie unterstützen den Fahrer in Form von Fahrerassistenzsystemen (*driver assistance systems*), die entweder das Steuern (z. B. *lane centering*) oder Beschleunigen bzw. Bremsen in einem bestimmten Fahrszenario (*driving mode*) übernehmen. Ein Beispiel für ein solches System ist der Einparkassistent.

Level 2 (Teilautomatisiertes Fahren/partial automation): Im Wesentlichen wie Level 1, nur wirken hier mindestens zwei automatisierte Kontrollfunktionen zusammen; ein oder mehrere Fahrerassistenzsysteme übernehmen zeitgleich sowohl das Steuern als auch Beschleunigen bzw. Bremsen in bestimmten Situationen. Ein Beispiel für ein entsprechendes System ist hier der Stauassistent.

Im Allgemeinen gilt für die Level 0 bis 2, dass der menschliche Fahrer ›fährt‹, d. h. er muss das System und die Fahrumgebung jederzeit überwachen, um gegebenenfalls eingreifen zu können. Er ist vollständig verantwortlich für den sicheren Betrieb des Fahrzeugs. Im Übergang von Level 2 zu 3 findet gewissermaßen ein Bruch statt, es kommt zu einer grundsätzlichen Veränderung in der Aufgabenteilung und Verantwortlichkeit. Der Fahrer gibt schrittweise weite Teile der Fahraufgabe an das System ab, das fortan die Fahrzeugkontrolle und Überwachung der Fahrumgebung übernimmt und ›fährt‹.

Level 3 (Bedingt automatisiertes Fahren/conditional automation): Sicherheitskritische Funktionen sind unter bestimmten Fahrbedingungen komplett automatisiert; das System übernimmt in spezifischen Szenarien die Fahrzeugkontrolle vollständig. Der

2.2 Herausforderungen im Kontext der Entwicklungsagenda

Fahrer muss das System nicht mehr jederzeit überwachen, aber verfügbar sein, wenn es im konkreten Anwendungsfall an seine Grenzen stößt und den Menschen zur Übernahme des Steuers auffordert. Der Fahrer fungiert als Absicherung bzw. Rückfallebene für das System. Tritt ein solcher Fall ein, ›fährt‹ fortan der Fahrer. Ein beispielhaftes System für diese Stufe ist das Staufolgefahren.

Level 4 (Hochautomatisiertes Fahren/high automation): Ein Eingreifen des Fahrers ist nicht mehr erforderlich. Das System übernimmt die Kontrolle in bestimmten Situationen vollständig. Wenn der Fahrer nicht oder nicht angemessen auf eine Aufforderung zur Übernahme reagiert, geht das Fahrzeug in einen sicheren Zustand über, indem es z. B. auf dem Seitenstreifen anhält.

Level 5 (Vollautomatisiertes Fahren/full automation): Der Fahrmodus wird als ›autonom‹ bzw. das Fahrzeug als ›selbstfahrend‹ bezeichnet. Alle Fahrfunktionen sind vollständig automatisiert; das System übernimmt zu jeder Zeit und in allen Situationen sowie unter allen Umwelt- und Fahrbahnbedingungen die Kontrolle über das Fahrzeug und überwacht die Fahrumgebung ständig. Der Fahrer greift ins Fahrgeschehen nicht mehr ein.

An dieser Stelle kann das Phänomen des autonomen Fahrens als ausreichend eingeführt gelten. In den folgenden Unterkapiteln wird sich nun der ethischen Dimension zugewandt, die selbstfahrende Fahrzeuge entfalten. Der Einstieg in diese komplexe Thematik erfolgt zunächst über eine Betrachtung der Zusammenhänge zweier Perspektiven: autonome Fahrzeuge als technische Innovation einerseits und als ökonomisches (Konsum-)Gut andererseits.

2.2 Herausforderungen im Kontext der Entwicklungsagenda

2.2.1 Die Wechselbeziehung zwischen technischer Reife und Wirtschaftlichkeit

Die Einführung des automatisierten Fahrens in den öffentlichen Straßenverkehr wird maßgeblich durch drei Aspekte stimuliert: die technische Reife der Fahrzeugsysteme und deren Komponenten, die Wirtschaftlichkeit des Gutes ›automatisiertes Fahrzeug‹ und den

jeweils geltenden Rechtsrahmen. Zwischen diesen bestehen Abhängigkeiten und Wechselwirkungen unterschiedlichen Ausmaßes.

Fortschritte in der technologischen Entwicklung entlang des Stufenmodells sind gleichbedeutend mit einem höheren Automatisierungsgrad; sie lassen sich nur erzielen, wenn automatisierte Fahrzeuge zunehmend unabhängiger von menschlichem Eingreifen operieren, ohne dabei die anvisierten Sicherheitsziele aus den Augen zu verlieren. Die zentralen Aufgaben für Design und Entwicklung autonomer Fahrsysteme (Level 5 und teilweise 4) stellen sich hinsichtlich Perzeption, Kognition, Verhaltensentscheidung und -ausführung (vgl. Wachenfeld & Winner, 2015, S. 466). Durch die intensive Forschungsarbeit und die hohen Entwicklungsinvestitionen des letzten Jahrzehnts sind die wesentlichen technischen Voraussetzungen für das automatisierte Fahren zum heutigen Tag bereits in Ansätzen geschaffen, jedoch noch nicht zur vollständigen technischen Reife gelangt.

Die mitunter größten Herausforderungen bei der technischen Realisierung höherstufiger Automatisierung bestehen in einer kontinuierlichen Verbesserung der maschinellen Wahrnehmungsleistung (vgl. Beiker, 2012, S. 1148–1149; Bengler et al., 2014, S. 6; KPMG LLP, Center for Automotive Research, 2012, S. 12). Maschinenlesbare Informationen sind die Basis für algorithmische Verhaltensentscheidungen. Ein zentrales Element derselben stellt die Situationsprädiktion¹⁴ dar, die mögliche Entwicklungen einer wahrgenommenen Szene aus dem verkehrlichen Umfeld des Fahrzeugs innerhalb eines festgelegten Zeithorizonts von wenigen Sekunden vorausberechnet. Auf Basis dieser sogenannten Episoden werden sodann Aktionen geplant und schließlich korrespondierende Bewegungspfade – sogenannte Trajektorien – anhand spezifischer Kriterien errechnet (vgl. Dietmayer, 2015, S. 421). Problematisch ist hierbei nun, dass die Prädiktionsfähigkeit gegenwärtiger Wahrnehmungssysteme im Notfall nicht ausreicht: Im Fall von auftretenden Funktionseinschränkun-

14 Die Situationsprädiktion baut auf der Situationserkennung auf, bei der einzelne Komponenten des dynamischen Fahrzeug-Umfeldmodells zueinander in Beziehung gesetzt werden. Dietmayer (2015, S. 420–432) nennt Unsicherheiten hinsichtlich dreier Aspekte, durch welche die maschinelle Wahrnehmung beeinträchtigt wird: Zustandsunsicherheit, Existenzunsicherheit, Klassenunsicherheit.

gen bei Systemen auf Level 3 und 4 beträgt der kritische Zeithorizont, den das Fahrzeug für die Rückgabe an den Fahrer selbst überbrücken muss, ca. fünf bis zehn Sekunden; ebenso lange benötigt ein Fahrzeug auf Level 5 für die Erreichung eines eigensicheren Zustands.¹⁵

Trotz eingebauter sensorischer Redundanz können momentane Systeme eine zuverlässige Prozesskette bis hin zur sicheren Fahrzeugführung in Notsituationen nur unzureichend gewährleisten. Erforderliche Fortschritte im Bereich der maschinellen Perzeption und Kognition umfassen daher sowohl die Detektion von Objekten und ihre physikalische Vermessung als auch die Zuordnung korrekter semantischer Bedeutungen (vgl. Dietmayer, 2015, S. 420–426). Es reicht nicht aus, ein Verkehrsschild nur als solches zu erkennen; das Fahrzeug muss auch dessen Bedeutung, z. B. ›Stop‹ oder ›Vorfahrt gewähren‹, erfassen können. Gleches gilt für die Gestenerkennung menschlicher Verkehrsteilnehmer, die einen wesentlichen Teil der Kommunikation im Verkehrsgeschehen ausmacht (vgl. Holzbock et al., 2023a). Hierfür ist die qualitative Weiterentwicklung von Kamera- und Sensortechnologien auf Hardwareseite ein kritischer Erfolgsfaktor.

Um Perzeption und Situationsverständen auch in komplexen dynamischen Umgebungen zuverlässig sicherzustellen, spielt Künstliche Intelligenz mit zunehmendem Automatisierungsgrad, v. a. ab Level 3, eine bedeutende Rolle (vgl. Bengler et al., 2014, S. 12). Vor allem der Einsatz von *Deep-Learning*-Techniken ist kontinuierlich zu optimieren, mittels derer das Fahrzeugsystem zunächst in Simulationen und später in kontrollierten Feldversuchen ›trainiert‹ wird, um in realen Situationen zuverlässig den korrekten Output generieren zu können (vgl. Vallor & Bekey, 2017, S. 341). Zum Beispiel präsentiert ein neuerer Ansatz von Holzbock et al. (2023b) ein Fahrzeug-Umgebungsmodell, das die Körperhaltung von Fußgängern als Indiz für deren Bewegungsabsichten mithilfe neuronaler Netzwerke interpretiert. Auf Level 5 wirken intelligente, sensorgestützte Sicherheitssysteme schließlich mit vernetzter Fahrzeugtechnologie zusammen, um mit der verkehrlichen Infrastruktur zur Laufzeit zu kommunizie-

15 Eigensicherheit (*intrinsic safety*) wird einem System dann zugesprochen, wenn es über spezifische Konstruktionsmechanismen bzw. -prinzipien verfügt, die sicherstellen, dass auch im Fall einer Störung oder eines Systemversagens keine grundsätzliche Gefährdung vom System selbst ausgeht.

ren. Dies ermöglicht eine stetige Optimierung des Fahrverhaltens beispielsweise hinsichtlich Geschwindigkeit, Routenplanung oder Antizipation potenzieller Gefahren, wodurch Verkehrseffizienz und Fahrsicherheit erhöht werden. So schlagen Schumann et al. (2023) beispielsweise einen erweiterten Pfadplanungsalgorithmus für große Umgebungen vor, der eine neuartige Methode zur Erkennung signifikanter Umgebungsveränderungen beinhaltet und eine effiziente Nutzung der Manövriertfähigkeit zur Pfadplanung in engen Umgebungen ermöglicht.

Spezifische Herausforderungen auf dem Weg zur Marktreife des hoch- und vollautomatisierten Fahrens stellen sich im Hinblick auf die Konvergenz von Kommunikations- und Sensor technologien sowie das Entwicklungspotenzial einzelner Komponenten. Konkrete Aufgabenstellungen sind z. B. die Optimierung der GPS-Technologie zur Positionsbestimmung, hochauflösende Kartierungsmethoden zur Verbesserung der Umfeldwahrnehmung, Bildverarbeitungsverfahren sowie zuverlässige und intuitive Mensch-Maschine-Schnittstellen für die Interaktion zwischen Fahrer und System (vgl. Bengler et al., 2014, S. 7). Neben der technischen Ausstattung der Fahrzeuge muss auch der Ausbau einer digitalisierten Verkehrsinfrastruktur vorangetrieben werden; dringlich ist hier vor allem die flächendeckende Ausstattung des Verkehrsnetzes mit dem neuen Mobilfunkstandard 5G (vgl. Verband der Automobilindustrie (VDA) e.V., 2022b).

Abseits von technologischen Durchbrüchen rückt auch die Entwicklung von Test- und Bewertungskonzepten künftig verstärkt in den Fokus; der ›kritische Pfad zum autonomen Fahren‹ (Bengler et al., 2014, S. 16) führt über die Bereitstellung innovativer Metriken, mittels derer sich die Leistungsfähigkeit von Menschen und Fahrrobotern messen und vergleichen lässt, mittelfristig Testverfahren verbessert werden können und schließlich die Zuverlässigkeit der zur Reife gebrachten Systeme steigt. Letztere muss sich an den Anforderungen der funktionalen Sicherheit¹⁶ orientieren; einschlägige

16 Die funktionale Sicherheit (*functional safety*) ist ein Teilbereich der allgemeinen Sicherheit, der sich auf die korrekte Funktionsweise eines Systems bezüglich seiner Eingabeverarbeitung bezieht. Sie ist dann gegeben, wenn jede spezifizierte Sicherheitsfunktion ausgeführt und die jeweilige Anforderungsstufe erreicht wird. Konkret bedeutet das, dass zur Erreichung der funktionalen Sicherheit alle

Referenzen für den Automobilbereich sind hierbei die Norm ISO 26262 (»Road vehicles – Functional Safety«)¹⁷ und insbesondere deren Schlüsselkomponente, das Risiko-Klassifizierungsschema *ASIL* (*Automotive Safety Integrity Level*)¹⁸, das für automatisierte Fahrzeuge die höchste Sicherheitsstufe (*ASIL D*) fordert.

Zweifellos ist die Entwicklung fortschrittlicher Technologien hin zur technischen Reife eine erfolgskritische Voraussetzung für die anvisierte Einführung des automatisierten Fahrens. Jedoch ist deren Gelingen auch abhängig von ökonomischen Faktoren. So sind automatisierte Fahrzeuge neben allen strategischen Zielen zur Stärkung des Technologie- und Automobilstandorts Deutschland schließlich auch ein Wirtschaftsgut, das den Gesetzmäßigkeiten der Marktwirtschaft unterliegt. Gemäß den Mechanismen von Angebot und Nachfrage ist eine großflächige Serienproduktion automatisierter Fahrfunktionen nur realistisch, wenn die entsprechende Nachfrage vorhanden ist. Auf diese Weise ließen sich zum einen die über einen langen Zeitraum notwendigen hohen Investitionen der Hersteller über den laufenden Markt refinanzieren. Zum anderen sind mittelfristig Netzwerkeffekte zu erwarten: Je mehr Menschen automatisierte Fahrzeuge nutzen, desto stärker wirken Skaleneffekte auf die Produktionskosten und damit auch den Preis, was wiederum die Attraktivität für weitere Nutzer erhöht (vgl. KPMG LLP, Center for Automotive Research, 2012, S. 20). Die Akzeptanz am Markt hat einen entscheidenden Einfluss auf die Durchdringung des Fahrzeug-

Maßnahmen gehören, die zur Fehlervermeidung dienen, ebenso wie Vorgänge, mit deren Hilfe während des Betriebs auftretende Fehler beherrscht werden können.

- 17 Die ISO 26262 stellt eine Abwandlung der IEC 61508 dar, die an die spezifischen Anforderungen im Automobilbereich angepasst ist: Lebenszyklus eines Fahrzeugs, Schnittstellen bei verschiedenen Zulieferaufträgen, konfigurierbare Software usw. Anwendbar ist die Norm auf alle Fahrzeugklassen bis 3500 kg (vgl. International Organisation for Standardisation, 2018).
- 18 Die Klassifizierung gemäß *ASIL* bezeichnet das Maß, in dem eine Fehlfunktion eines Systems relevant für dessen Sicherheit ist. Sie orientiert sich an drei Parametern: Häufigkeit der Situation mit Relevanz der jeweiligen Fehlfunktion (*exposure*), Kontrollierbarkeit der Fehlfunktion (*controllability*) und Schwere der Auswirkung bei geringer Kontrollierbarkeit (*severity*). Je nach Belegung der Parameter kann der *ASIL* auf einer Skala von A bis D bestimmt werden und die sich daraus ergebenden speziellen Anforderungen sind zusätzlich zur ISO 26262 zu erfüllen (vgl. International Organisation for Standardisation, 2018, Teil 9).

bestands mit automatisierten Fahrzeugen.¹⁹ Diese wiederum stimuliert einerseits Umfang und Geschwindigkeit des technologischen Fortschritts und andererseits auch das Maß, in dem das autonome Fahren positive Wirkung entfaltet. Letztere ist skalierbar sowohl mit der Anzahl zugelassener Fahrzeuge als auch mit dem Grad der Automatisierung: Je mehr Fahraufgaben das System übernimmt, desto weniger verbleiben bei der fahrzeugführenden Person und desto seltener führen menschliche Fahrfehler zu Unfällen. Je mehr Fahraufgaben automatisiert ablaufen, desto mehr Szenarien können effizient koordiniert werden, z. B. zeit- und kraftstoffeffizientes Fahren im gebundenen Verkehr, im Stau oder bei der Parkplatzsuche. Für mobilitätseingeschränkte Personengruppen, die zum Führen konventioneller Autos nicht ermächtigt bzw. in der Lage sind, werden automatisierte Fahrzeuge zudem erst dann nutzbar, wenn kein menschliches Eingreifen mehr erforderlich ist, d. h. ab Level 4.

Wann können wir mit der Marktreife derartiger Systeme rechnen? Im nachfolgenden Unterkapitel wird der aktuelle Entwicklungsstand autonomer Fahrzeuge resümiert und sein Verhältnis zu entsprechenden gesetzlichen Bestimmungen ausgelotet.

2.2.2 Wo stehen wir heute? Aktueller technischer Stand und regulative Verordnungen

Im gegenwärtigen Realbetrieb auf öffentlichen Straßen sind Fahrzeuge mit teilautomatisierten Fahrfunktionen des Levels 2 weitgehend etabliert. In bestimmten Fahrsituationen, wie bei Stau- oder Autobahnfahrten, kann das System bereits die Fahrzeugkontrolle übernehmen, wobei die fahrzeugführende Person stets die Fahrumgebung überwachen und jederzeit zum Eingreifen bereit sein muss. Systeme, die sich am Übergang zwischen Level 2 und 3 befinden, stehen zur Serienproduktion bereit bzw. werden als Prototypen im (beschränkten) Realverkehr zur Weiterentwicklung der KI-gestützten Software erprobt. Automobilbauer setzen dabei verstärkt auf Ko-

¹⁹ Becker und Axhausen (2017) erarbeiten einen systematischen Überblick über Studien, die sich mit der Akzeptanz von automatisierten Fahrzeugen beschäftigen. Als relevante Faktoren identifizieren sie u. a. personenbezogene Charakteristika (z. B. Geschlecht, Alter), Präferenzen für Einsatzszenarien (z. B. Stadt, Autobahn) und eine grundsätzliche Affinität zu Systemen der Fahrerassistenz.

operationen mit Unternehmen außerhalb des Automobilbereichs; so entwickelt Daimler seine Mercedes-Benz-Modelle in enger Zusammenarbeit mit Grafikprozessor- und Chipsätze-Entwickler Nvidia, um KI-Komponenten zur Reife zu bringen. Auch Volkswagen testet die Komponenten seines Entwicklungspartners Argo AI auf Teststrecken in typischen Verkehrssituationen. Ein Durchbruch gelang kürzlich Mercedes-Benz, dem für den Staupiloten »Drive Pilot« als erstem deutschen Automobil die Zulassung für ein Level-3-System erteilt wurde; dieses fährt im Stau mit Geschwindigkeiten bis zu 60 km/h selbstständig, ohne dass der Fahrer das System ständig überwachen muss (vgl. Rudschies & Kroher, 2024). Konkurrent BMW hingegen erhielt kürzlich vom Kraftfahrtbundesamt die Genehmigung für seinen Autobahnassistenten, der seit Oktober 2023 im neuen 5er-Modell erhältlich ist. Dabei handelt es sich um ein sogenanntes *Level-2-Hands-off-System*, d. h. das System fährt selbstständig bis zu 130 km/h, der Fahrer muss allerdings jederzeit eingreifen können.

Kontinuierlich arbeitet die Automobilbranche in Deutschland an der Neuentwicklung höherstufiger Systeme. Wegbereiter ist hier hauptsächlich das automatisierte *Valet Parking*, mit der das Level 4 des hochautomatisierten Fahrens erstmals beschriften wird. Es stellt eine der ersten in Serie realisierten autonomen Fahrfunktionen dar, die bereits eine Zulassung für den Alltagsbetrieb erhalten haben. Als erster Automobilhersteller hat Mercedes-Benz die entsprechende Funktion in einem Serienfahrzeug verbaut, gemeinsam mit Partner Bosch im Parkhaus des Mercedes-Benz Museums in Stuttgart erprobt und darf diese nun in fest definierten Bereichen im Regelbetrieb einsetzen (vgl. Mercedes-Benz Group AG, 2020). Ende 2022 erteilte das Kraftfahrt-Bundesamt den beiden Stuttgarter Unternehmen die gemeinsame Zulassung für das weltweit erste zertifizierte *Automated-Valet-Parking-System*, den fahrerlosen »Intelligent Park Pilot«, der seitdem in einem speziell ausgerüsteten Parkhaus am Stuttgarter Flughafen operiert. Gesteuert durch das Fahrsystem und die von Bosch entwickelte intelligente Infrastruktur des Parkhauses erfolgt der Parkvorgang vollständig automatisiert, nachdem der Fahrer das Fahrzeug in einer definierten Übergabezone verlassen hat (vgl. Bosch Mobility Solutions, 2021; Mercedes-Benz Group AG, 2022). Weitere Meilensteine auf dem Weg zum autonomen Fahren aus Sicht deutscher Hersteller konnten im August 2024 vermeldet werden: In Kooperation mit dem chinesischen Unternehmen

WeRide schickte Mercedes-Benz erste Level-4-Testfahrzeuge auf die Straßen Pekings, zunächst nur auf Autobahnen und in spezifisch ausgewiesenen, stark frequentierten Bereichen in der Innenstadt. Erprobungsziel ist dabei u. a. die Umfelderkennung mittels verschiedener Sensoren (vgl. Köllner, 2024c). Zudem erhielt eine von der deutschen Firma Vector entwickelte AUTOSAR²⁰-Basissoftware die Zertifizierung gemäß ASIL D und ist ab sofort für automatisierte Funktionen in der Fahrzeugentwicklung einsetzbar (vgl. Köllner, 2024b).

In den USA sind höherstufige Systeme bereits fester Bestandteil alltäglicher Mobilität. Allen voran Automobilriese General Motors investiert an verschiedenen Standorten weltweit hohe Summen in die Entwicklung automatisierter Fahrsysteme. Als Pionier im Bereich der Autopiloten agierte lange Zeit der von US-Milliardär Elon Musk geführte amerikanische Konzern Tesla, der bereits 2015 ein System präsentierte, das an der Grenze zu Level 3 auf Autobahnen operiert. Trotz seiner Innovationskraft – oder vielleicht gerade deswegen – wird Tesla regelmäßig mit negativen Schlagzeilen bedacht, die auf eine Vielzahl aufsehenerregender Unfälle und diverse laufende Klagen zurückzuführen sind. Unlängst entschied ein kalifornisches Gericht, dass das Autopilotensystem des Herstellers nicht für einen Unfall im Jahr 2019 verantwortlich zu machen ist, als ein Tesla-Fahrzeug von der Fahrbahn abgekommen und in Flammen aufgegangen war (vgl. ARD-Tagesschau, 2023). Dies könnte ein richtungsweisender Erfolg nicht nur für den Konzern, sondern auch für die US-Justiz sein. Diese wird sich in den kommenden Jahren mit zahlreichen ähnlichen Klagen befassen müssen, denn spektakuläre Unfälle im Zusammenhang mit automatisierten Fahrzeugen sind in den USA nicht selten. So stand auch die General-Motors-Tochter Cruise zeitweise in der Kritik, nachdem sie zuvor eine Vorreiterrolle eingenommen hatte. Als Folge eines fatalen Unfalls im vergangenen Jahr, an dem ein Robotaxi der Firma beteiligt war, wurde dieser vorerst die Lizenz entzogen. Inzwischen sind die Cruise-Fahrzeuge mit mittelfristigen Expansionsplänen auf die Straßen zurückgekehrt,

20 AUTOSAR (*AUTomotive Open System ARchitecture*) ist die Bezeichnung einer globalen Partnerschaft führender Unternehmen in der Automobil- und Softwareindustrie mit dem Ziel, standardisierte Software-Frameworks und -architekturen für zukünftige intelligente und sichere Mobilitätslösungen zu entwickeln und zu etablieren (vgl. AUTOSAR GbR, 2024).

werden zunächst allerdings von Menschen gesteuert, um Kartenmaterial zu aktualisieren und Vertrauen zurückzugewinnen (vgl. Kennemer, 2024).

Erfolgsgeschichte schreibt hingegen die Google-Schwestergesellschaft Waymo, deren Flotten von Robotaxis bereits seit Längerem in Phoenix, Arizona, San Francisco und Los Angeles unterwegs sind, teilweise ohne Überwachung durch einen Sicherheitsfahrer. Als weltweit erstes Unternehmen, das Robotaxis zur Marktreife gebracht hat, droht Waymo sogar Tesla bei möglichen Investoren den Rang abzulaufen (vgl. Göpfert, 2024). Bei allem ökonomischen Erfolg verbleiben allerdings auch hier Sicherheitsbedenken, die verstärkt ins Blickfeld der US-Verkehrsbehörde National Highway Traffic Safety Administration (NHTSA) geraten, wenngleich abgesehen von Situationen mit Blechschäden bislang keine folgenschweren Zwischenfälle zu vermelden waren (vgl. Ohnsman, 2024).

Es ist naheliegend, dass die jeweils gültige Rechtslage eine entscheidende Rolle dahingehend spielt, wo und in welchem Umfang selbstfahrende Fahrzeuge bereits praktisch zum Einsatz kommen. Einen Rechtsrahmen für das autonome Fahren und seine spezifischen ethischen und juristischen Herausforderungen zu schaffen, ist dabei eine rechtspolitische Aufgabe (vgl. Hilgendorf, 2019, S. 357–358). Regulatorische Instrumente haben sich in der Vergangenheit als effektive Stimuli erwiesen, um die Marktdurchdringung neuer Technologien im Automobilbereich zu erhöhen und auf diese Weise die Weiterentwicklung der Automatisierungsfunktionen entlang des Stufenmodells zu unterstützen. Umgekehrt können fehlende oder unzureichende gesetzliche Bestimmungen auch hemmend auf den Innovationsfortschritt und eine großflächige Markteinführung wirken (vgl. Hilgendorf, 2018a, S. 681). In den vergangenen Jahren stand die Gesetzeslage in Deutschland der Verkehrsautomatisierung noch eher konservativ gegenüber. So trat im Juni 2017 zunächst eine Änderung des Straßenverkehrsgesetzes in Kraft, wonach Systeme auf Level 3 zeitweise die Fahraufgabe übernehmen dürfen, wenn weiterhin eine fahrzeugführende Person im Notfall die Kontrolle übernehmen kann. Die Erprobung von Prototypen höherstufig automatisierter Fahrzeuge im Realverkehr wurde bis zu diesem Zeitpunkt noch durch gesetzliche Hürden ausgebremst; sie durften nur

mit Sondergenehmigungen auf fest definierten Betriebsflächen oder Teststrecken fahren.²¹

Diese über weite Strecken zögerliche Anpassung des Rechtsrahmens für den Regelbetrieb von autonom agierenden Fahrzeugen im öffentlichen Verkehr ist ursächlich dafür, dass Branchenexperten der deutschen Automobilindustrie trotz aller Investitionskraft einen Entwicklungsrückstand auf die USA attestieren. Dort sind die rechtlichen Rahmenbedingungen aufgrund des föderalen *Laissez-Faire*-Ansatzes günstiger; die Entscheidungsbefugnis über das Verkehrsrecht liegt bei den Bundesstaaten, sodass sich die jeweiligen Regelungen in verschiedenen Staaten unterscheiden. Einige erlauben den Regelbetrieb selbstfahrender Fahrzeuge bereits seit geraumer Zeit; die liberalsten Regularien bestehen in Arizona, Texas, Nevada und Michigan. Im Januar 2021 haben die USA zudem die Zulassungsanforderungen in Bezug auf die zu erfüllenden Sicherheitsnormen für Roboterautos gelockert, um die Einführung in den Realverkehr und damit mittelfristig auch die Markteinführung zu beschleunigen. Diese Regelung gilt allerdings zunächst nur für vollautomatisierte Fahrzeuge im Gütertransport, da sich hier keine schutzwürdigen fahrzeugführenden Personen mehr an Bord befinden (vgl. Schmidt, 2021). Im Herbst 2023 hat die US-amerikanische NHTSA neue regulative Entwürfe unter der Bezeichnung »AV STEP« (*ADS-equipped Vehicle Safety, Transparency, and Evaluation Program*) vorgelegt. Diese sollen die Einführung selbstfahrender Autos weiter beschleunigen, indem sie die Höchstgrenzen für die maximal zulässige Anzahl an Fahrzeugen mit Assistenz- und Automatisierungssystemen aufheben und zugleich eine vermehrte Datenpreisgabe von Seiten der Unternehmen fordern (vgl. Chasins, 2024; McElligott, 2023).

Andererseits können gesetzliche Vorschriften auch als »Mittel einer reflektierten Innovationsförderung« (Hilgendorf, 2018b, S. 93) wirken, beispielsweise indem sie die Serienausstattung von Neufahrzeugen mit bestimmten Fahrfunktionen verpflichtend vorschrei-

21 Dies erfolgte beispielsweise als innerbetriebliche Fahrten auf dem Firmengelände oder im Rahmen von Forschungsprojekten wie dem »EVA-Shuttle« für den Personentransport zwischen Haustür und Haltestelle des öffentlichen Nahverkehrs im Karlsruher Stadtgebiet oder als *Mobility-on-Demand*-System zur Reduzierung des Individualverkehrs am Touristenziel Hambacher Schloss im Projekt »Hambach-Shuttle« (vgl. Bundesministerium für Digitales und Verkehr, 2022a).

ben.²² Ein zentrales Themenfeld, das kritisch für Investitionsrisiken und damit auch Investitionsanreize der Hersteller ist,²³ betrifft die Klärung haftungsrechtlicher Fragen (vgl. Borenstein et al., 2017, S. 67–68; Marchant & Lindor, 2012, S. 1337). Sollen Hersteller in vollem Umfang haften, wenn ein automatisiertes Fahrzeug in einen Unfall verwickelt ist? Die einschlägige juristische Fachliteratur wächst kontinuierlich; bisher wurden vor allem Ansätze einer angepassten und erweiterten Produkthaftung diskutiert, die Defizite traditioneller Regelungen ausgleichen und die spezifischen Anforderungen der neuen Technologien berücksichtigen sollen (vgl. Douma & Palodichuk, 2012; Gasser, 2015; Gurney, 2013, 2017; Hilgendorf, 2018a; Koch, 2022; Wu, 2015, 2020).²⁴

Um die Konkurrenzfähigkeit der europäischen Automobilwirtschaft zu erhalten, haben die EU-Staaten eine Harmonisierung der Vorschriften auf internationaler (EU-)Ebene zum Ziel erklärt. Da dieses jedoch in seiner Ausformulierung bis auf Weiteres auf sich warten lässt, schuf die damalige Bundesregierung schließlich auf eigene Initiative einen Rechtsrahmen, um das automatisierte Fahren weiter voranzubringen. Im Mai 2021 beschloss der Deutsche Bundestag mit Zustimmung des Bundesrates schließlich das für das Bundesgebiet gültige sogenannte »Gesetz zum autonomen Fahren«. Damit war Deutschland das erste Land weltweit, das einen nationalen Rechtsrahmen für den Regelbetrieb hochautomatisierter Fahrzeuge (Stufe 4) in festgelegten Betriebsbereichen im öffentlichen Straßenverkehr bereits ab 2022 setzte. Ziel dieser Regelung ist es, Innovationen in der Fahrautomatisierung schnell zu etablieren und die Weiterentwicklung durch Erprobung unter Realbedingungen zu

22 So gehört etwa das Elektronische Stabilitätsprogramm seit 2014 zur Pflichtausstattung aller in der Europäischen Union zugelassenen Pkw (vgl. Bengler et al., 2014, S. 2–4).

23 Garza (2011, S. 616) erläutert, dass die Entwicklung autonomer Fahrzeuge sich trotz anfänglich hoher Haftungsrisiken mittelfristig für die Hersteller lohnt, da zu erwarten ist, dass sich langfristig Häufigkeit und Schwere negativer Vorfälle mit steigender Nutzerakzeptanz reduzieren.

24 Zu beachten ist, dass Lösungsansätze zu spezifischen rechtlichen Fragen wie der Unfallhaftung sich nur eingeschränkt verallgemeinern lassen, da sie sich zumeist auf spezifische staatliche Rechtsrahmen beziehen, wie beispielsweise die englischsprachige Literatur auf die Rechtsordnung der USA. In Bezug auf die deutsche Rechtsordnung sind hier vor allem die Ansätze von Hilgendorf (2018a, 2018b, 2019) einschlägig.

fördern (vgl. Verband der Automobilindustrie (VDA) e.V., 2022a). Als erste Einsatzszenarien sind u. a. ein Shuttlebetrieb, z. B. zur Aufwertung des ÖPNV in städtischen Randbereichen oder im ländlichen Raum, Betriebsfahrten (*Hub2Hub*), Logistik oder *Dual-Mode*-Fahrzeuge wie beim *Automated Valet Parking* vorgesehen (vgl. Bundesministerium für Digitales und Verkehr, 2021). Neben Regelungen für technische Anforderungen an das Fahrzeug, Kriterien für die Erteilung der Betriebserlaubnis und den Umgang mit Daten sowie Pflichten von Haltern und herstellenden Unternehmen führt das neue Gesetz die sogenannte »Technische Aufsicht« als menschliche Kontrollinstanz ein (Art. 1, § 1d, Abs. 3). Dabei handelt es sich um eine natürliche Person, die das Fahrzeug deaktivieren oder Fahrma-növer freigeben kann und durch eine Haftpflichtversicherung abgesichert ist. Experten aus Politik und Praxis bewerten das Gesetz zwar als wichtigen Schritt in die richtige Richtung, kritisieren aber u. a. die auf den ÖPNV und die gewerbliche Nutzung beschränkte Anwendbarkeit, die fehlende Regelungsklarheit hinsichtlich Datensicherheit und Datenschutz sowie Bestimmungen bezüglich der Haftungsproblematik, die zu stark zu Gunsten der Hersteller ausfällt (vgl. Deutscher Bundestag, 2021; Fahrenholz, 2021).²⁵ Zudem werden viele ethische Fragen nur vage thematisiert (vgl. Kriebitz et al., 2022).

Ein Jahr später folgte die »Verordnung zur Regelung des Betriebs von Kraftfahrzeugen mit automatisierter und autonomer Fahrfunktion und zur Änderung straßenverkehrsrechtlicher Vorschriften« (Bundesministerium für Digitales und Verkehr, 2022b), die den nationalen Rechtsrahmen zum autonomen Fahren vorerst vervollständigte.²⁶ Gemäß ihrer Digitalstrategie ist es das proklamierte Ziel der amtierenden Bundesregierung, das autonome Fahren bis 2025 vom Pilotprojekt zu einem festen Bestandteil der Praxis alltäglicher Mobilität voranzubringen (vgl. Bundesministerium für Digitales und Verkehr, 2023, S. 20). Zu diesem Zweck sollen zukünftig weitere

25 Für Anmerkungen und eine Übersicht über offene juristische Fragen das neue Gesetz betreffend siehe Hilgendorf (2017a, S. 226–229, 2019, S. 359–361).

26 Diese befasst sich mit den Forderungen von Verbänden und Unternehmen nach konkreten Vorgaben für Genehmigungsverfahren, Normung, Zertifizierung und Standards (vgl. Verband der Automobilindustrie (VDA) e.V., 2022a), der Sicherstellung von Kompatibilität (vgl. KPMG LLP, Center for Automotive Research, 2012, S. 15) und der Überwindung zulassungs- und haftungsrechtlicher Hürden.

Fördergelder in die ganzheitliche Entwicklung investiert werden, um die Produktionskosten angesichts des zeitversetzt zu erwartenden Refinanzierungseffekts über den Markt zu kompensieren.

Eine zentrale Bedeutung bei der Regulierung der Verkehrsautomatisierung kommt dem im Mai 2024 von den 27 Mitgliedsstaaten der EU verabschiedeten und wenige Monate später durch die Veröffentlichung im Amtsblatt der EU in Kraft getretenen prominenten Rechtsakt zur umfassenden Regulierung Künstlicher Intelligenz zu. Der sogenannte *AI Act* stellt einen horizontalen Regulierungsrahmen bereit, der eine Risikoklassifizierung von KI-Systemen vornimmt und spezifische Transparenzpflichten für vier definierte Risikokategorien vorgibt, die anhand von Kriterien wie Fairness, Transparenz, Robustheit und Zuverlässigkeit festgelegt werden. Während Systeme mit inakzeptablen Risiken²⁷ fortan verboten sind, werden Hochrisikoanwendungen mit weitreichenden Auflagen bedacht. Unter diese Risikokategorie fallen Systeme, die »eine erhebliche Bedrohung für die Gesundheit, Sicherheit oder die Grundrechte der Einzelnen darstellen« (Hengl, 2024) können; dies gilt neben beispielsweise Systembauteilen in der kritischen Infrastruktur oder medizinischen Geräten auch für autonome Fahrzeuge.

Nun ist der *AI Act* keine explizite Regelung für die Automobilindustrie, betrifft diese aufgrund der umfangreichen Risikoauflagen aber in besonderem Maße. So wird zukünftig vorgeschrieben, dass KI-Systeme selbstfahrender Fahrzeuge strenge Test- und Validierungsverfahren durchlaufen müssen, um zugelassen zu werden. Weiterhin ist auch die KI, die beim Test und der Validierung der Fahrzeugsoftware zum Einsatz kommt, Gegenstand neuer Regularien. Als besonders herausfordernd gestaltet sich dabei der Umstand, dass die Validierung im Hinblick auf dynamische und unvorhersehbare Umgebungen erfolgen muss, um als sicher und zuverlässig gelten zu können.²⁸

27 Zu diesen zählen etwa kognitive Verhaltensmanipulation, Emotionserkennung oder Sozialkreditsysteme (vgl. Köllner, 2024a).

28 Darüber hinaus sieht die neue Verordnung die Einrichtung einer zweistufigen Governance-Architektur vor. Dabei obliegt nationalen Behörden die Durchsetzung der Vorschriften für KI-Systeme und deren Beobachtung, während KI-Modelle mit allgemeinem Verwendungszweck (*General Purpose AI, GPAI*) auf EU-Ebene beaufsichtigt werden. Diese Modelle »können für vielfältige Aufgaben eingesetzt werden und bilden die Grundlage für viele KI-Systeme in

Ist der *AI Act* nun eher als Innovationshemmer oder Erfolgsfaktor für die Entwicklung selbstfahrender Fahrzeuge zu bewerten? Der VDA befürchtet zunächst ein Ausbremsen von Innovation aufgrund des neuen KI-Gesetzes, denn die Regelungen sollen mittelfristig in die EU-Verordnung für die Genehmigung von Kraftfahrzeugen übernommen werden. Insbesondere die Regularien zur Typgenehmigung und Marktüberwachung von Kraftfahrzeugen sollen gezielt ergänzt und somit sektoruell reguliert werden, was zunächst mit hohen bürokratischen Belastungen verbunden sein wird. Grundsätzlich sei dieser delegierte Rechtsakt jedoch zu befürworten, um den spezifischen Herausforderungen der Automobilbranche gerecht zu werden (vgl. Köllner, 2024a). Zunächst ist allerdings ein verlangsamter Markteintritt aufgrund von hohen Compliance-Anforderungen zu erwarten. Mittelfristig bringen diese jedoch die Branche voran: Verbindliche und klare Regelungen schaffen Vertrauen und Akzeptanz, Rechtssicherheit minimiert das unternehmerische Risiko und befähigt die Innovationsbereitschaft, die Förderung ethischer Standards führt zu nachhaltigen und verantwortungsvollen Innovationen, die europäischen Unternehmen einen Wettbewerbsvorteil verschaffen könnten.

Als allgemeiner Orientierungssatz für die Evaluierung bestehender und zukünftiger Rechtsrahmen gilt: Diese müssen so gestaltet sein, dass die Entwicklung des hoch- und vollautomatisierten Fahrens nicht behindert, sondern vielmehr gefördert und begleitet wird, wobei jedoch keine Kompromisse bei der Sicherheit zugelassener Systeme gemacht werden dürfen. Zielkonflikte scheinen hierbei vorprogrammiert. Inwiefern dies Abstriche hinsichtlich der Erwartungen an autonome Fahrzeuge bedeutet, wird im folgenden Unterkapitel diskutiert.

der EU. Einige davon könnten systemische Risiken bergen, wenn sie sich als besonders leistungsfähig erweisen oder eine weite Verbreitung finden.« (Europäische Kommission, 2024, o. S.) Das Europäische Gremium für Künstliche Intelligenz (KI-Gremium) soll die EU-weite Zusammenarbeit koordinieren. Verstöße gegen die Regularien werden künftig anhand von prozentualen Geldbußen sanktioniert. Das Testen von KI-Systemen unter realen Bedingungen soll ohne größere Hürden möglich werden, sofern gewisse Schutzvorkehrungen gewährleistet sind (vgl. ebd.).

2.2.3 Zwischen Utopie und Dystopie: Die Ambivalenz des autonomen Fahrens

Die erwarteten positiven Wirkungen erheben die Automatisierung der Mobilität zu einem ›digitalen Heilsversprechen‹, das Lösungen für viele unserer gegenwärtigen gesellschaftlichen Problematiken suggeriert. Sie gilt als zentrales Puzzleteil auf dem Weg zur Verwirklichung einer sicheren, effizienten und inklusiven Mobilität. Doch wie realistisch sind diese Erwartungen eigentlich? Wie wird die Zukunft der Mobilität tatsächlich aussehen, in der (voll-)automatisierte Fahrzeuge unsere Straßen bevölkern?²⁹

Seit einiger Zeit mahnen kritische Stimmen vermehrt an, dass das weitgehend positive Bild vom autonomen Fahren wesentliche Aspekte vernachlässigt und sich mittelfristig ambivalente Folgeeffekte einstellen könnten. In der Forschungsliteratur finden mögliche sekundäre und tertiäre Effekte bisher noch wenig Beachtung. Eine weniger optimistische Haltung gegenüber der Einführung autonomer Fahrsysteme gründet sich im Wesentlichen auf Erfahrungen aus früheren Reduktionsstrategien. So waren Versuche, die Umweltbelastung durch technische Innovationen in der Energieeffizienz zu reduzieren, zu Beginn stets vielversprechend, haben sich jedoch in der Realität nur marginal ausgewirkt bzw. letztlich sogar eine weitere Verschlechterung der Umweltbilanz herbeigeführt:

We have made these mistakes before, even for automobiles. When the automobile was first introduced, two of its primary purposes were to increase safety and protect the environment. Horses were dangerous and pollutive. The automobile was supposed to address these two problems and make roadways safer and cleaner. However, driver error replaced horse error. Carbon dioxide replaced horse manure, and its nonvisible nature and lack of smell make it a more challenging source of pollution. (Gurney, 2022, S. 147)

In der Vergangenheit waren sogenannte Reboundeffekte – neben unbeabsichtigten Externalitäten – u. a. auf erhöhte Konsumanreize (vgl. Jackson, 2009, S. 62–63) zurückzuführen, die durch neue Technologien geschaffen wurden. Eine ähnlich kontraintuitive Dynamik

29 Siehe hierzu auch den Beitrag von Ryan (2020), der anhand einer Szenarienanalyse mögliche Treiber und Hemmnisse der Entwicklung selbstfahrender Fahrzeuge bis zum Jahr 2025 untersucht.

wird auch für das autonome Fahren befürchtet: Durch die komfortable Form der Mobilität entstehen Kauf- und Konsumanreize für Privathaushalte und andere potenzielle Nutzer, deren Zahl sich um mobilitätseingeschränkte Personen erweitert. Es muss daher realistischerweise davon ausgegangen werden, dass sich die Zahl der gefahrenen Fahrzeugkilometer oder gar der Gesamtfahrzeugbestand mit der Einführung des autonomen Fahrens erhöht (vgl. Gogoll & Müller, 2017, S. 685). Damit sind weitere negative Externalitäten assoziiert, wie erhöhte Staugefahr oder die Untergrabung der Komfortvorteile des öffentlichen Nahverkehrs, die Emissionen weiter in die Höhe treiben. Aus ökonomischer Sicht drohen finanzielle Einbußen für die Kommunen, wenn gebührenpflichtige Parkräume wegfallen, und weniger Schadensfälle lassen die Gehaltsbudgets von Versicherungen und Werkstätten schrumpfen (vgl. Anderson et al., 2016, S. 38–40).

Auch die gesellschaftlichen Effekte lassen sich, zumindest in mittelfristiger Perspektive, nicht ausschließlich positiv bewerten. So ist anzunehmen, dass zunächst vor allem finanzielle Individuen von den Vorteilen autonomer Fahrzeuge profitieren werden, wodurch soziale Ungleichheiten nicht reduziert, sondern verstärkt werden. Gurney (2022, S. 148–153) erläutert weitere Beispiele möglicher unbeabsichtigter sekundärer Folgen: Es ist zu befürchten, dass im Fall reduzierter Unfallzahlen die Verfügbarkeit von Spenderorganen zurückgeht, wodurch sich die Zahl der an Organversagen Verstorbenen erhöhen wird. Negative Effekte sind auch für Industrien alternativer Mobilitätsformen zu erwarten, beispielsweise für die Luftfahrt oder den Schienenverkehr, denen autonome Fahrzeuge durch ihre Attraktivität den Rang abzulaufen drohen.

Weitere Fragezeichen im Hinblick auf eine allzu rosige Vorstellung von der automatisierten Mobilität bestehen im Hinblick auf deren primären Motivator, das Sicherheitspotenzial. Gemäß DIN 31000 beschreibt der Begriff der Sicherheit eine Sachlage, in der das Risiko das Grenzrisiko nicht übersteigt. Entscheidend für die Sicherheit in einer Fahrsituation ist die Gesamtleistungsfähigkeit des Systems, welches das Fahrzeug im jeweiligen Fall steuert. Dieses besteht bis einschließlich Level 2 aus Fahrer und Assistenz- bzw. Teilautomatisierungsfunktionen. Analysen von Unfalldaten belegen, dass Letztere menschliche Fahrfehler gut kompensieren können, wodurch ihnen zumindest in Routinesituationen ein allgemeiner Sicherheitszuwachs

gegenüber Fahrzeugen ohne Automatisierungsfunktionen zugesprochen werden kann. Fahrsysteme der Hoch- und Vollautomatisierung müssen sich in Bezug auf ihre Leistungsfähigkeit allerdings dem Vergleich mit der Fahrfähigkeit eines menschlichen Fahrers stellen, dessen »Bestleistung« sie nicht nur erreichen, sondern für ein erhöhtes Sicherheitspotenzial sogar übertreffen müssen (vgl. Winkle, 2015, S. 373).³⁰ Kahn (2022) stellt die berechtigte Frage, ab wann autonome Fahrsysteme denn als »besser« als menschliche Fahrer gelten können – und welcher Bewertungsmaßstab dabei angelegt werden soll. Wenn immer mehr Fahraufgaben automatisiert ablaufen, droht langfristig sogar der Verlust der menschlichen Fahrkompetenz (vgl. Sparrow & Howard, 2017, S. 208). Dieses »Paradoxon der Automatisierung« offenbart einen Zielkonflikt zwischen Sicherheit und Effizienz bzw. Fahrkomfort, der Systeme auf Level 2 und 3 wiederum als besonders kritisch im Hinblick auf ihre Fahrsicherheit erscheinen lässt.³¹

Beim teilautomatisierten Fahren (Level 2), das eine Überwachung und ein potenzielles Eingreifen durch die fahrzeugführende Person noch erforderlich macht, kommt zusätzlich ein weiterer, spezifischer Aspekt hinzu: Je geringer der Umfang, in dem Personen in die Fahraufgaben involviert sind, desto eher tendieren sie zur Unaufmerksamkeit, was dazu führt, dass sie in Notsituationen zu langsam, gar nicht oder nicht angemessen reagieren (können) (vgl. Köllner, 2021; Sparrow & Howard, 2017, S. 207–208).³² Ein reales Beispiel in diesem Kontext ist der tödliche Unfall mit einem automatisierten

30 Systeme der bedingten Automatisierung (Level 3) stellen hier einen Grenzfall dar: Fahren müssen zwar in Notfällen eingreifen können, sind also theoretisch Teil des Fahrsystems, aber in allen anderen Fällen irrelevant für dessen Gesamtleistungsfähigkeit.

31 Das Phänomen der übermäßigen Abhängigkeit vor allem von Autopiloten ist in der Luft- und Schifffahrt ein etabliertes Risiko, das auftritt, wenn Kontrollsysteme automatisiert werden und menschliches Eingreifen nur noch in Notfällen notwendig ist. Problematisch ist dabei insbesondere, dass die Aufgabenteilung mit dem System häufig unklar ist, was die Aufgaben des Piloten nicht einfacher, sondern komplexer macht (vgl. Wolmar, 2018, S. 42–47).

32 Dies genügt insbesondere nicht dem Konzept der *Meaningful Human Control*, das seit Kurzem in der Forschungsliteratur zur Kontrollierbarkeit und Verantwortungsfähigkeit autonomer Systeme diskutiert wird: »The concept of MHC appeals to the intuition that when autonomous systems are deployed in unstructured, dynamic and potentially unpredictable environments, simply having a human agent involved at some point in the decisional chain [...] may not

UBER-Fahrzeug, der sich 2018 im amerikanischen Arizona ereignete und infolge des großen Medieninteresses traurige Berühmtheit erlangte (vgl. Lee, 2019; Levin & Wong, 2018): Abgelenkt durch ihr Smartphone nahm die Sicherheitsfahrerin eine Fußgängerin, die die Straße überqueren wollte, zu spät wahr. Wie Untersuchungen bestätigten, war in diesem Fall ein Softwarefehler bzw. eine Voreinstellung ursächlich dafür, dass das automatisierte Fahrzeug die Fußgängerin fälschlicherweise als irrelevantes Objekt einstufte (vgl. Beutnagel, 2018). Dennoch wäre die Sicherheitsfahrerin trotz des aktivierten Autopiloten dazu verpflichtet gewesen, die Fahrumgebung stetig zu überwachen, und hätte so womöglich die Fehleinschätzung des Systems überstimmen können.

Weiterhin ist auch die zeitliche Perspektive des erwarteten Sicherheitszuwachses fraglich. Einer der wesentlichen Skalierungsfaktoren für die Größenordnung der Sicherheitswirkung, die autonome Fahrzeuge entfalten, ist der Grad der Marktdurchdringung.³³ So wird geschätzt, dass bei einem Marktanteil von 10 % bereits 1100 Leben pro Jahr gerettet werden könnten, während diese Zahl bei 90 % Anteil auf 21.700 Leben steigt. Nun bedeutet ein Fahrzeugkauf jedoch eine große Investition für private Haushalte, weshalb die Erneuerungszyklen in der Regel recht lang sind (vgl. Altenburg et al., 2018, S. 2). Es ist daher anzunehmen, dass sich neue Automatisierungstechnologien nicht disruptiv, sondern nur langsam im Bestand durchsetzen werden. Frühestens ab 2050 ist eine mengenmäßige Durchdringung im Gesamtfahrzeugbestand zu erwarten, die nennenswerte Sicher-

be sufficient to prevent unwanted mistakes and so-called accountability gaps; human persons must maintain a role that is as prominent as possible.« (Mecacci & Santoni de Sio, 2020, S. 104) Bisher wurde es primär im Kontext autonomer Waffensysteme betrachtet, lässt sich aber auch auf das automatisierte Fahren anwenden (vgl. Santoni de Sio & van den Hoven, 2018; Santoni de Sio, 2021, S. 720).

33 Dieser Zusammenhang gilt analog für die beiden Aspekte Effizienz und Mobilitätsbedürfnisse: Je höher der Anteil automatisierter bzw. autonomer Fahrzeuge am Verkehrsaufkommen ist, desto größer ist auch der Anteil des effizient steuerbaren Verkehrsflusses und des umweltfreundlichen Fahrverhaltens am gesamten Verkehrsaufkommen. Die gesamte Zeitersparnis durch effizienten Verkehrsfluss, insbesondere Stauvermeidung, wird bei einem Marktanteil von 10 % auf 756 Millionen Stunden pro Jahr geschätzt, bei 90 % Marktanteil auf 2772 Millionen Stunden pro Jahr (vgl. Fagnant & Kockelman, 2015, S. 172–175). Je mehr automatisierte Fahrzeuge zugelassen werden, desto mehr Personen können von neuen bzw. wiedergewonnenen Mobilitätschancen profitieren.

heitseffekte mit sich bringt. Diese werden sich erwartungsgemäß zunächst in einer Reduzierung der Sachschäden äußern. Auf signifikante Auswirkungen auf die Zahl der Personenschäden wird man noch länger warten müssen, denn schwere Unfälle mit Todesfolge passieren besonders häufig auf Landstraßen, wo auf absehbare Zeit noch kaum Automatisierungsfunktionen greifen werden. Demnach kann realistischerweise erst langfristig – im Zeithorizont von frühestens dreißig Jahren – mit einem relevanten Sicherheitszuwachs gerechnet werden (vgl. ebd., S. 40–47). Ebenfalls fragwürdig ist, inwiefern die Vision einer vernetzten und zentral gesteuerten Mobilität ein realistisches Ziel darstellt, solange Sicherheitsbedenken in Bezug auf Datenschutz und Cyberattacken fortbestehen:

Ob in Zukunft eine dem Bahn- und Luftverkehr entsprechende vollständige Vernetzung und zentrale Steuerung sämtlicher Kraftfahrzeuge im Kontext einer digitalen Verkehrsinfrastruktur möglich und sinnvoll sein wird, lässt sich heute nicht abschätzen. Eine vollständige Vernetzung und zentrale Steuerung sämtlicher Fahrzeuge im Kontext einer digitalen Verkehrsinfrastruktur ist ethisch bedenklich, wenn und soweit sie Risiken einer totalen Überwachung der Verkehrsteilnehmer und der Manipulation der Fahrzeugsteuerung nicht sicher auszuschließen vermag. (Di Fabio et al., 2017, Regel Nr. 13)

Für höherstufig automatisierte Systeme gilt ferner, dass es aufgrund ihrer eingeschränkten maschinellen Wahrnehmungsfähigkeit und daraus resultierender Fehleinschätzung von Situationen zu neuen, bislang unbekannten Unfallkonstellationen kommen kann, die in einem Verkehrsgeschehen ohne Automatisierung nicht auftreten würden (vgl. Brändle & Grunwald, 2019, S. 289). Dies wäre beispielsweise dann der Fall, wenn eine Programmierung auf Schadensminimierung dazu führt, dass andere Verkehrsteilnehmer die Risikoaversion autonomer Fahrsysteme ausnutzen und sich bewusst unachtsam und rücksichtslos verhalten (vgl. Millard-Ball, 2018, S. 10–11). Dieser Aspekt wirft die Frage auf, ob autonome Fahrsysteme letztendlich überhaupt einen Mehrwert liefern. Sofern sie tatsächlich ein signifikantes Sicherheitspotenzial besitzen, ist ihre flächendeckende Einführung streng genommen ein moralischer und gesellschaftlicher Imperativ (vgl. Hevelke & Nida-Rümelin, 2015b, S. 621; Nyholm,

2018c, S. 6).³⁴ Konsequent zu Ende gedacht würde ein solcher auch das Verbot konventioneller Fahrzeuge implizieren, die dem durch Fahrautomatisierung gesetzten Sicherheitsstandard nicht mehr gewachsen sind: »Once they are safer than human drivers when it comes to risks to 3rd parties, then it should be illegal to drive them: at that point human drivers will be the moral equivalent of drunk robots.« (Sparrow & Howard, 2017, S. 206) In diesem Sinne bekräftigen auch Müller und Gogoll (2020, S. 1550–1561), dass konventionelle Fahrzeuge ein ungerechtfertigtes Risiko für andere Verkehrsteilnehmer in sich bergen und daher aus moralischen Gründen verboten werden müssten: »To put it more drastically: In a nutshell, manual driving in a scenario in which autonomous cars are affordable is the moral equivalent of running around with a hand grenade for pleasure.«³⁵

Tatsächlich stellt gerade der Mischverkehr eine große Herausforderung in puncto Sicherheit dar. Dies ist u. a. darauf zurückzuführen, dass sich der menschliche Fahrstil und die Funktionsweise eines Fahrroboters grundlegend unterscheiden:

Self-driving cars have optimizing driving styles and are very strict rule followers. Human drivers, in contrast, are typically satisficers: They drive just well enough to satisfy their driving goals. And humans are more flexible in their attitudes to traffic rules. (Nyholm, 2018c, S. 7)

In der Folge kann es zu Problemen in der Abstimmung und Erwartungsbildung zwischen Mensch und Maschine kommen, aus denen eine erhöhte Unfallgefahr resultiert. So können autonome Fahrzeuge menschliches Fahrverhalten nur sehr schwer antizipieren, während ihre mangelnde Fähigkeit zu nonverbaler und informeller Kommunikation beispielsweise durch Gestik es wiederum menschlichen Fahrern erschwert, das Verhalten autonomer Fahrzeuge korrekt zu deuten (vgl. Färber, 2015, S. 137–143; Sparrow & Howard, 2017, S. 2011).

34 Ein ähnliches Argument verwendet John Harris in seiner berühmten *Survival Lottery* (1975).

35 Müller und Gogoll beziehen sich in ihrer Begründung auf ein Argument von Sven Ove Hansson (2003), der Risikoübertragungen genau dann als gerechtfertigt betrachtet, wenn diese ihrerseits Vorteile für alle Beteiligten mit sich bringen. Dieses Argument wird in Kap. 7.3.3.4 näher betrachtet.

Ein mögliches rigoroses Verbot konventioneller Fahrzeuge tangiert jedoch nicht zu vernachlässigende soziale und kulturelle Aspekte. So wird das eigenhändige Führen eines Fahrzeugs häufig als »Ausdruck eines besonderen Lebensgefühls [...], in dem sich die Dokumentation des eigenen Status, Sehnsucht nach Freiheit, Freude an Sport und gelegentlich auch Lust auf Abenteuer verbinden« (Hilgendorf, 2017a, S. 225), angesehen. Entsprechende Restriktionen würden die individuelle Selbstverwirklichung und Autonomie empfindlich beschneiden (vgl. Borenstein et al., 2019, S. 392; Hansson et al., 2021, S. 1402; Moor, 2016). Wie die Ethik-Kommission (Di Fabio et al., 2017, S. 11, Regel 6) konstatiert, wäre ein solcher Eingriff in die Selbstbestimmung ethisch nicht vertretbar:

Umgekehrt ist eine gesetzlich auferlegte Pflicht zur Nutzung vollautomatisierter Verkehrssysteme oder die Herbeiführung einer praktischen Unentzerrbarkeit ethisch bedenklich, wenn damit die Unterwerfung unter technische Imperative verbunden ist (Verbot der Degradierung des Subjekts zum bloßen Netzwerkelement).

Bisher sind einige alternative Ansätze vorgeschlagen worden, um ein striktes Verbot konventioneller Fahrzeuge zu umgehen und gleichzeitig negative Effekte des Mischverkehrs abzuschwächen. Müller und Gogoll (2020, S. 1563–1564) entwerfen das visionäre Szenario eines *AI-Supervised Human Driving*, bei dem der Mensch die dynamische Fahraufgabe ausführt, dabei jedoch von einem intelligenten künstlichen System überwacht wird, das im Notfall den Fahrer überstimmen kann. Hingegen plädieren Nyholm und Smids (2020, S. 339–340) dafür, bestimmte Aspekte des menschlichen Fahrens an eine robotergestützte Fahrweise anzupassen, um auf diese Weise einerseits die Koordinationsprobleme zwischen Mensch und Maschine zu überwinden und andererseits potenzielle Risikoquellen für menschliche Fahrfehler einzudämmen.³⁶ Praktisch umsetzbar wäre dies beispielsweise über restriktive Anpassungen der Verkehrs vorschriften oder mithilfe bestimmter Technologien, z. B. durch geschwindigkeitsregulierende Komponenten, Frühwarnsysteme bei erhöhter Kollisionsgefahr oder Alkohol-Interlocks. Roy (2016, o. S.) argumentiert in eine ähnliche Richtung: »[...] if autonomous cars

36 Nyholm und Smids (2020, S. 338–339) liefern verschiedene Argumente, weshalb eine umgekehrte Anpassung automatisierten Fahrverhaltens an das menschliche wenig zielführend ist.

can set a certain, provable safety standard, it might make sense that the licensing requirements would therefore require a human driver to prove a similar competency—which raises the safety bar for human-driven cars, as well.«

Schlussendlich verbleiben mögliche Effekte, die sich heute noch gar nicht absehen lassen. Eine gewisse Ambivalenz ihrer Wirkungen ist innovativen Technologien inhärent; sie können sowohl unbeabsichtigte positive als auch negative Konsequenzen nach sich ziehen und langfristig in Weisen wirken, die nicht vorhersehbar sind (vgl. Lin, 2013a). Dieser unter dem Begriff ›Schmetterlingseffekt‹ (*butterfly effect*) bekannte Mechanismus erschwert es, zum aktuellen Entwicklungszeitpunkt eine umfassende Bewertung des automatisierten Fahrens in langfristiger Perspektive vorzunehmen. Der Schwerpunkt aktueller Bemühungen sollte daher primär auf den generellen Wirkungen der Fahrautomatisierung liegen, die kurz- und mittelfristig zu erwarten sind – und auf Maßnahmen, die dazu dienen, mögliche negative Effekte so gut wie möglich zu antizipieren und abzuschwächen. Um dieses Ziel zu erreichen, ist eine Auseinandersetzung mit ethischen Fragen rund um autonome Fahrzeuge unerlässlich. Im folgenden dritten Kapitel wird nun zunächst der bestehende ethische Diskurs rekonstruiert, wobei insbesondere moralische Dilemma-Szenarien in den Blick genommen werden.