

## 8. Fazit und Ausblick

### 8.1 Ergebnisse der philosophischen Untersuchung: Zusammenfassung

Das autonome Fahren stellt eines der faszinierendsten technologischen Entwicklungsprojekte unserer Zeit dar. Als disruptive Technologie ist es Schauplatz vielfältiger politischer, rechtlicher und nicht zuletzt ethischer Debatten, die miteinander in Beziehung gesetzt werden müssen, um ein umfängliches Verständnis seines Wesens als gesellschaftliches Phänomen zu erlangen.

In dieser Forschungsarbeit wurde das Problem der ethischen Gestaltung von Unfallalgorithmen als eines der zentralen Puzzleteile auf dem Weg zur anvisierten Mobilitätswende untersucht. Die vorliegende Auseinandersetzung fokussierte spezifische moralische Dilemma-Strukturen, die Entscheidungssituationen angesichts unabwendbarer Unfälle kennzeichnen. Sie lässt sich dabei im weiteren Kontext eines sich neu formierenden, risikofokussierten Forschungszugangs verorten. Die erarbeiteten Ergebnisse wurden entlang von Teilzielen, Arbeitshypothesen und Zwischenergebnissen entfaltet und strukturiert, was die Nachvollziehbarkeit der Argumentation erleichtern soll. An dieser Stelle werden abschließend die gedanklichen Schritte und zentralen Ergebnisse des in diesem Buch entwickelten argumentativen Narrativs in wertender Form zusammengetragen.

Kap. 1 widmete sich einer Einführung in die Untersuchung, in deren Rahmen zunächst das autonome Fahren als sozio-technisches Phänomen skizziert und hinsichtlich seiner ethischen Dimension problematisiert wurde. In diesem Zuge wurde die ethische Gestaltung von Unfallalgorithmen als das zentrale Forschungsanliegen des Buches motiviert, das entsprechende Erkenntnisinteresse prägnant formuliert und dessen Relevanz in theoretischer und praktischer Hinsicht aufgezeigt. Als zentraler Forschungsauftrag wurde die Entwicklung eines alternativen Problemzugangs herausgestellt, der relevante Dilemma-Szenarien als risikoethische Entscheidungs-

probleme interpretiert. Der methodische Ansatz der Forschungsarbeit wurde als interdisziplinär ausgerichtete, angewandte ethische Untersuchung charakterisiert, die Methoden und Konzepte aus der Maschinenethik, Digitalen Ethik und Ethik der Künstlichen Intelligenz einerseits sowie der Risikoethik andererseits integriert. Auf eine präzise Darstellung der forschungsleitenden Thesen und Ziele der philosophischen Untersuchung folgte schließlich eine Skizze des argumentativen Gedankengangs, welcher dem Narrativ dieses Buches zugrunde liegt.

Teil I begann mit einer einführenden Darstellung des Forschungsgegenstands. In Kap. 2 wurde ein Bild des autonomen Fahrens in seinem visionären Charakter und seinen vielfältigen Herausforderungen sowie der Verortung spezifischer ethischer Fragen im Forschungsdiskurs gezeichnet. Hierbei wurden zunächst die Grundlagen zum Verständnis moralischer Dilemma-Strukturen im gewählten real-lebensweltlichen Kontext gelegt. Die Senkung der sozialen Kosten motorisierter Individualmobilität wurde als übergeordnete Zielsetzung in der Entwicklung autonomer Fahrsysteme erläutert. Den primären Motivator stellt ein signifikant erhöhtes Sicherheitspotenzial dar, das im Zuge der Eliminierung menschlicher Fahrfehler zu erwarten ist. Anhand des Stufenmodells des autonomen Fahrens, das als zentrale Richtschnur für Forschung, Entwicklung und Regulierung gilt, konnten die Herausforderungen herausgearbeitet werden, mit denen sich das praktische Projekt der Verkehrssautomation konfrontiert sieht.

Wie eine nähere Analyse zeigte, prägen Wechselwirkungen zwischen dem technischen Reifegrad der Fahrzeugsysteme bzw. deren Komponenten, ihrer Wirtschaftlichkeit und dem jeweils geltenden Rechtsrahmen das Voranschreiten der Mobilitätsrevolution maßgeblich. Einen Meilenstein in der Schaffung eines innovationsfördernden nationalen Rechtsrahmens markiert das 2021 beschlossene »Gesetz zum autonomen Fahren«. Nichtsdestotrotz sehen sich regulierende Institutionen der Notwendigkeit gegenüber, zwischen Sicherheits- und Entwicklungszielen zu vermitteln: Gelockerte Restriktionen dürfen nicht zulasten der Sicherheit gehen. Der kürzlich in Kraft getretene europäische *AI Act* stellt einen bedeutenden Schritt in diese Richtung dar, indem er Automobilunternehmen in die Pflicht nimmt, sich proaktiv mit möglichen negativen Auswirkungen der von ihnen entwickelten Hochrisiko-KI-Systeme auseinanderzu-

setzen. Vor diesem Hintergrund ist das autonome Fahren prinzipiell als Initiator ambivalenter Wirkungen anzuerkennen. Eine kritische Reflexion auf die assoziierten Erwartungen im Sinne eines ›digitalen Heilsversprechens‹ legte offen, dass assoziierte positive Effekte von ernstzunehmenden negativen Externalitäten begleitet werden.

In einem zweiten Schritt wurden in Kap. 3 ethische Problemfelder und die entsprechenden Diskurse überblicksartig skizziert. Dabei rückten angewandte ethische Fragen in den Vordergrund, die im Kontext des Sicherheitsversprechens autonomer Fahrzeuge stehen. Neben kurSORischen Einblicken in den relevanten Verantwortungsdiskurs konnte die Problematik unvermeidbarer Unfallsituationen sowohl in ihrer ethischen als auch informationstechnischen Dimension eruiert werden. Die Begrenztheit des Antizipationspotenzials autonomer Systeme, ihre generelle Fehleranfälligkeit sowie die Nicht-Eliminierbarkeit spezifischer Gefahrenpotenziale eines dynamischen Verkehrsgeschehens sind die entscheidenden Faktoren, aufgrund derer sich folgenreiche Unfälle nicht vollständig ausschließen lassen. Als Spezialfall beschränkt antizipierbarer Szenarien wurden Situationen evaluiert, in denen Personenschäden unabhängig von gewählter Trajektorie und Bremsverhalten des betreffenden Fahrzeugs unabwendbar sind. Derartige Szenarien weisen dilemmatische Strukturen auf, weshalb man sie im ethischen Diskurs als Instanzen moralischer Dilemmata auffasst: Ist eine Kollision unausweichlich, gibt es keine eindeutige ›richtige‹ Entscheidung.

Ferner wurde moralischen Dilemma-Szenarien sowohl in theoretischer als auch in praktischer Hinsicht eine essenzielle Rolle für die Realisierung einer automatisierten und vernetzten Mobilität attestiert. Eines der zentralen Argumente lautete dabei, dass die dilemmatischen Strukturen, die dem Entscheidungsproblem zugrunde liegen, sich tatsächlich in vielen alltäglichen Fahrsituationen stellen – oftmals ohne dass es uns bewusst ist. Das vor allem in der Tagespresse gezeichnete Bild tragischer Dilemma-Entscheidungen über Leben und Tod pointiert lediglich ein Problem, das omnipräsent ist: Überholvorgänge, Abstand zum Vordermann, gewählte Fahrgeschwindigkeit – in vielen Situationen sind ethisch relevante Abwägungen zwischen relevanten Handlungsgründen bzw. den Interessen Einzelner implizit enthalten. Die Komplexität, Vielfalt und Dynamik denkbarer Szenarien wurden als Gründe angeführt, weshalb sich Unfalldilemmata nicht anhand von Heuristiken normieren lassen;

sie sind »nicht ethisch zweifelsfrei programmierbar« (Di Fabio et al., 2017, S. 11) und müssen als Einzelfälle ethisch gewürdigt werden. Wie ausgewählte Literatur zeigte, kommt Unfalldilemmata in gesellschaftlicher Hinsicht insofern eine hohe Bedeutung zu, als sie eine große psychologische Rolle für potenzielle Nutzer spielen und auf diese Weise zu den erfolgskritischen Determinanten der Marktdurchdringung autonomer Fahrzeuge zählen. Schlussendlich erfordert die technische Robustheit der Entscheidungslogik von Fahrsystemen, dass für alle denkbaren Fälle jederzeit definierte Zustände und entsprechende Handlungsvorgaben abrufbar sind.

In Teil II wurde das zentrale Argument entwickelt, welches die erste forschungsleitende These begründet: Ein Zugang zur Gestaltung von Unfallalgorithmen, wie ihn der einschlägige Forschungsdiskurs bisher diskutiert, lässt (zu) viele Fragen offen. In Kap. 4 erfolgte eine kritische Auseinandersetzung mit bis dato dominanten Forschungszugängen, unter denen moralische Dilemma-Situationen als Problematik moralischer Designentscheidungen interpretiert werden. Auf der Basis maschinenethischer Grundlagen wurde argumentiert, dass durch Algorithmen getroffene Entscheidungen sich in moralisch relevanter Weise von menschlichen Entscheidungen unterscheiden. Im Zuge ihrer moralphilosophischen Problematisierung konnte eine systematisierte Zusammenstellung repräsentativer Dilemma-Szenarien erarbeitet werden, auf welche die Argumentation im weiteren Verlauf der Forschungsarbeit immer wieder zurückgriff.

Ausgehend von einer kritischen Darstellung bisheriger Forschungsliteratur wurde dem im Diskurs weit verbreiteten Narrativ einer möglichen Analogie zum bekannten Trolley-Problem nachgespürt, wobei essenzielle Diskrepanzen offengelegt werden konnten. Eine eingehende Untersuchung des praktischen Kontextes, in den real-lebensweltliche Unfalldilemmata eingebettet sind, lieferte wesentliche Anhaltspunkte für eine mangelnde Eignung bisher vorgeschlagener Ansätze. Dilemma-Szenarien sind – anders als durch die vermeintliche Trolley-Analogie suggeriert – keine individuellen Entscheidungsprobleme, sondern vielmehr soziale Problemstellungen, die in einem Netzwerk gesellschaftlicher Verflechtungen stehen. Es wurde dargelegt, dass Unfallalgorithmen als Gegenstand politischer Regulierung in einem Spannungsverhältnis zwischen individuellen Präferenzen einerseits und pluralistischen Wertvorstellungen ande-

rerseits verhaftet sind. Aus konzeptionellen und praktischen Gründen ließ sich die Option personalisierter ethischer Einstellungen für Unfallalgorithmen zurückweisen. Weiterhin wurde das Verhältnis von Ethik und Politik im Hinblick auf die Gestaltung von Unfallalgorithmen ausgelotet. Die dargestellte Argumentation bekräftigte, dass ethische Begründbarkeit und gesellschaftliche Akzeptanz im Sinne einer öffentlichen Vernunft nach Rawls'scher Lesart Hand in Hand gehen.

Auf der Grundlage dieser Ergebnisse demonstrierte eine umfassende Rekonstruktion der relevanten Forschungsliteratur, dass bisher dominante Forschungszugänge sowohl in methodischer als auch inhaltlicher und struktureller Hinsicht bedeutende Schwächen offenbaren. Anhand einer systematischen Analyse konnte erörtert werden, dass es weder deskriptiven noch normativen Herangehensweisen bisher gelungen ist, überzeugende Entscheidungsstrategien für Unfalldilemmata bereitzustellen. Utilitaristische, deontologische, tu-gendethische, vertragstheoretische, rechtsphilosophische, meta-normative und pluralistische Ansätze stoßen ebenso wie Konzepte und Methoden experimenteller Ethik diesbezüglich an ihre Grenzen. Als eine der essenziellsten Fehlinterpretationen, die dominanten Forschungszugängen aufgrund ihrer Fokussierung auf das Trolley-Problem zugrunde liegt, wurde die Annahme sicherer Handlungskonsequenzen bzw. des Eintretens bestimmter Umweltzustände akzentuiert. Es wurde betont, dass Entscheidungen in Bezug auf Unfalldilemmata tatsächlich mit Unsicherheiten hinsichtlich der zu erwartenden Folgen behaftet sind. Die Verfügbarkeit probabilistischer Bewertungsmethoden und stochastischer Entscheidungsmodelle erlaubt es, Eintrittswahrscheinlichkeiten grob zu bestimmen, weshalb man von Entscheidungen unter Risiko ausgeht. Dies motiviert die Entwicklung eines alternativen, spezifisch risikobasierten Zugangs, wie er in Teil III entwickelt wurde.

In Kap. 5 wurde das zentrale Argument, welches die erste These stützt, um eine metaethische Perspektive erweitert. Eine metaethische Rekonstruktion der Wertekonflikte, welche die relevanten moralischen Dilemma-Strukturen begründen, ermöglichte es, eine ganzheitliche Perspektive auf die Problemstellung zu eröffnen. Als zentrales Argument in der Begründung der Existenz unlösbarer Dilemmata wurde die Idee der Inkommensurabilität spezifischer moralischer Werte bzw. Gebote erörtert und anhand zweier Konzeptio-

nen veranschaulicht. Der Ansatz des unausweichlichen moralischen Fehlverhaltens nach Christopher Gowans fußt auf der Inkonvertibilität bzw. Nicht-Einlösbarkeit von Werten, die aufgrund einer einzigartigen Verantwortung gegenüber Personen besteht. Im Fall einer Verletzung derselben kommt es zu unersetzbaren Verlusten, die, so folgert Lisa Tessman im Zuge ihrer Weiterentwicklung des Konzepts von Gowans unter Bezugnahme auf Martha Nussbaums *Capability Approach*, nur dann moralisch relevant sind, wenn sie Kosten auferlegen würden, die nicht kompensiert werden können. Können Werte weder substituiert noch kompensiert werden, muss der Akteur zwangsläufig moralisch scheitern, die entsprechenden moralischen Gebote sind nicht verhandelbar. Es folgte eine Diskussion der Ergebnisse der metaethischen Analyse im Hinblick auf ihre Relevanz für den Anwendungskontext moralischer Unfalldilemmata. Die Argumentation machte deutlich, dass das zentrale Problem, welches die Komplexität moralischer Dilemma-Situationen im Kontext des autonomen Fahrens theoretisch-formal begründet, in der inkompatiblen Verschränkung moralisch gleichrangiger legitimer Interessen von Individuen und der gleichzeitigen Nicht-Verrechenbarkeit deontologischer Pflichten zum Schutz des Lebens besteht. Das entscheidende Argument bildet dabei die Inkommensurabilität unersetzbarer und einzigartiger Werte, auf die das Fehlen systematischer Herangehensweisen an derartige Dilemmata zurückzuführen ist.

Sodann folgte eine Untersuchung konkreter Entscheidungsstrategien, die auf metaethischer Basis argumentieren. Mit dem Verweis auf verantwortungsethische Gründe sowie eine fragwürdige Haltung der Indifferenz gegenüber Leben und Rechten der Betroffenen erfolgte eine Zurückweisung des Vorschlags, Dilemma-Szenarien per Zufallsprinzip zu entscheiden. Im Anschluss an Thomas Nagel wurde ein alternativer Ansatz dargelegt, der einen vielversprechenden Ausweg skizziert und dabei auf der Grundannahme basiert, dass sich, wenn formale Lösungen nicht möglich sind, auf praktischer Ebene dennoch Perspektiven eröffnen können. Daraufhin wurden Entscheidungsstrategien für echte moralische Dilemmata, die sich an der Maßgabe einer pragmatischen Ethik orientieren, zum Desiderat erklärt. Diese abstrahieren von theoretischen Analysen und fokussieren praktisch zielführende Abwägungen auf der Basis legitimierter Werte. Auf diese Weise ließ sich das in Kap. 4 entwickelte Argument dahingehend vervollständigen, dass sich aufgrund der

Konflikte incommensurabler Werte eine pragmatische Ethik als vielversprechende Bewältigungsstrategie anbietet. Die Sachbezogenheit auf die Spezifika von Dilemma-Szenarien ist der entscheidende Blickwinkel, von dem ausgehend offene Fragen im Sinne eines pragmatischen Ansatzes geklärt werden können. Damit liefern die Ergebnisse dieses Kapitels kein eigenes Argument im engeren Sinne, sondern erlauben als argumentative Stütze einen Brückenschlag zwischen der ersten und zweiten These des Buches: Im Hinblick auf eine explizite ethische Auseinandersetzung mit dem Unsicherheitsaspekt, der der Gestaltung von Unfallalgorithmen anhaftet und in dieser Forschungsarbeit im Fokus steht, erscheint eine risikoethische Perspektive vielversprechend.

In Teil III wurde schließlich eine entsprechende risikoethische Interpretation der Problematik von Unfallalgorithmen präsentiert. Diese orientierte sich an der zweiten These, welche besagt, dass sich unter einem risikoethischen Zugang zentrale Fragen des Anwendungsproblems klären lassen und normative Implikationen freigelegt werden können, die neuen Entscheidungsperspektiven den Weg bereiten. In Kap. 6 war es zunächst das Anliegen, begriffliche und konzeptionelle Grundlagen in Bezug auf wissenschaftliche Verortung, Historie, Gegenstandsbereich und Paradigmen der Risikoethik zu legen. Es folgte eine Veranschaulichung möglicher Begründungsansätze hinsichtlich einer generellen moralischen Zulässigkeit von Risikoübertragungen anhand dreier risikopraktischer Paradigmen (konsequentialistisch-objektivistisch, postmodern-subjektivistisch, partizipatorisch). Im Kontext einer Kritik rationaler Risikopraxis wurden gängige entscheidungstheoretische Kriterien für fragwürdig befunden, da diese aufgrund ihrer konsequentialistischen Grundlage weitgehend unvereinbar sind mit dem Selbstverständnis einer Risikoethik, welche die Legitimität eines risikobehafteten Systems stets an deren Wirkungen auf Einzelinteressen bemisst. Das Kapitel mündete in der Forderung nach einem Paradigmenwechsel in risikopraktischen Fragen, der mit der Abkehr von einer konsequentialistischen Grundorientierung einhergeht. Angesichts der konfliktierenden incommensurablen Werte, die Unfalldilemmata kennzeichnen, lässt sich im Hinblick auf die Gestaltung von Unfallalgorithmen für die Hinwendung zu einer deontologischen Risikoethik werben.

## 8. Fazit und Ausblick

Grundzüge einer solchen konnten in Kap. 7 systematisch entwickelt werden. Übergeordnetes Ziel war es dabei, die spezifischen Umstände zu bestimmen, unter denen Risiken übertragen werden dürfen. In einem ersten Schritt ging es darum, eine Interpretation vorzulegen, welche die Gestaltung von Unfallalgorithmen als risikoethisches Verteilungsproblem betrachtet, das sich im Spannungsfeld zwischen soziologischer Risikoakzeptanz und ethischer Risikoakzeptabilität bewegt. Es fand eine Abgrenzung der Forschungsarbeit von bisheriger risikoethischer Forschung zur Thematik des autonomen Fahrens statt, die sich auf eine Analyse von Effekten distributiver Gerechtigkeit sowie Entwürfe implementierungsnaher, formaler Ansätze konzentriert. Schließlich wurde eine dezidiert ethische Perspektive systematisch mit dem Ziel erarbeitet, den ›blindten Fleck‹ des Diskurses auszumerzen. Um die entsprechende Forschungslücke zu schließen und den Diskurs voranzubringen, fand zunächst eine Neuinterpretation der bisherigen Leitfrage nach der moralphilosophischen Begründung von Entscheidungsprinzipien für Unfallalgorithmen statt. In diesem Zuge rückten die ethische Rechtfertigbarkeit und Verteilung risikoinduzierter Vor- und Nachteile in den Vordergrund. Neben konsequentialistischen wurden auch kontraktualistische Kriterien der Risikoakzeptabilität evaluiert und für lediglich begrenzt adäquat befunden.

Als Antwort auf deren festgestellte Schwächen erfolgte die Skizze einer deontologischen Risikoethik, die das Ziel verfolgt, einerseits Werte wie Freiheit, Autonomie und individuelle Rechte zu schützen, die durch Risikoübertragungen bedroht sind, aber andererseits auch nicht zu restriktiv zu sein, um eine vollständige Paralyse sozialen Miteinanders zu vermeiden. Hier wurde im Anschluss an den schwach deontologischen, vor dem Hintergrund politischer Fragen kontraktualistisch verfassten Ansatz von Julian Nida-Rümelin, Johann Schulenburg und Benjamin Rath die Zielvorgabe einer kohärenten Risikopraxis ausgerufen. Diese integriert konsequentialistische und deontologische Elemente entsprechend ihrer jeweiligen Rechtfertigungsgrundlage und besteht darauf, dass deontologische Kriterien bei der Implementierung von Algorithmen harte Grenzen (*hard constraints*) darstellen, die vor quantitativen Optimierungszielen Vorrang haben müssen.

Schließlich wurde ein Entwurf einer kohärenten Risikopraxis skizziert, der auf zwei zentralen Prinzipien beruht, die als Grenz-

kriterien nicht-verhandelbare Rahmenbedingungen konstituieren, innerhalb derer sich die Verteilung von Vor- und Nachteilen aus Risikoübertragungen als Optimierungsproblem bewegt. Als erstes und absolutes Prinzip wurde die Zumutbarkeit von Risiken begründet. Diese fokussiert das, was rational betrachtet akzeptiert werden sollte, und nicht, was tatsächlich sozial akzeptiert ist. Die entscheidenden Aspekte, aufgrund derer sich subjektive Risikopräferenzen in objektivierte Risikobewertungen transferieren lassen, bilden die Reziprozität in der Übertragung zumutbarer Risiken und der daraus resultierende Vorteil, keine absoluten Einschränkungen individueller Handlungsfreiheit und in der Verfolgung persönlicher Ziele hinnnehmen zu müssen. Zur Bestimmung des zumutbaren Risikos in Unfalldilemmata wurden moralische Sorgfaltspflichten als Kernkriterium einer entsprechenden Schwellenwert-Konzeption vorgeschlagen. Das zentrale moralische Argument, das dabei zur Anwendung kommt, besteht wiederum in deren Reziprozität, durch die soziale bzw. sozio-technische Gefüge wie Mobilität erst tragfähig werden.

Als zweites und relatives Prinzip wurde das Postulat einer fairen Verteilung von Vor- und Nachteilen im Kontext von Risikoübertragungen formuliert. Bezugnehmend auf die Rawls'sche Idee der ›Gerechtigkeit als Fairness‹ ließen sich risikopraktische Strategien genau dann als fair bezeichnen, wenn die sich daraus ergebende Verteilung von Vor- und Nachteilen als fair gelten kann. Ferner wurde argumentiert, dass sich die spezifische Problematik von Unfallalgorithmen nur anhand von egalitaristischen Gerechtigkeitskonzeptionen adäquat fassen lässt. Auch eine Ungleichverteilung kann fair sein – nämlich genau dann, wenn berechtigte Ansprüche auf bestimmte Vorteile anderen gegenüber bestehen. Die Überprüfung dieser Hypothese identifizierte das Ausmaß, in dem eine Person Risiken für andere verursacht, als ethisch relevanten Faktor für die Fairness von Risikoübertragungen. Hierbei wurde erläutert, inwiefern sich auf der Basis einer Kategorisierung von Verkehrsbeteiligten entlang ihrer jeweiligen Vulnerabilität die Gewährung von Sicherheitsvorteilen für besonders gefährdete Gruppen rechtfertigen lässt. Erklärtes Ziel ist es dabei, bestehende kontingente Ungleichheiten im Sinne einer ergebnisegalitaristischen Position fair zu berücksichtigen.

Um den Rückgriff auf egalitäre Gerechtigkeitskonzeptionen im Hinblick auf das gewählte Anwendungsproblem zu rechtfertigen, wurden zwei prominente Konzeptionen rekonstruiert und kritisch

beleuchtet. Zum einen wurde unter Anwendung von John Rawls' *Differenzprinzip* ausgeführt, dass die Orientierung an denjenigen Personen, die am schlechtesten gestellt sind, höchste Priorität genießen sollte. Eine kritische Diskussion deutete an, dass eine heuristisch interpretierte Instanz des Rawls'schen *Schleier des Nichtwissens*, wie sie im Diskurs häufig verwendet wird, berechtigte Ansprüche in Bezug auf Unfalldilemmata nicht final begründen kann. Zum anderen wurde ein Ansatz von Sven Ove Hansson vorgestellt, der Risikoübertragungen genau dann als zulässig betrachtet, wenn sie Teil eines gerechten Systems von Risiken sind, welches dem Einzelnen Vorteile bringt. Die Idee einer distributiven Gerechtigkeit wird hier durch die Reziprozität hergestellt, die jedem Einzelnen zum Vorteil gereicht. Abschließend blieb festzuhalten, dass beide Ansätze hinsichtlich ihrer Eignung, finale Begründungen fairer risikoethischer Verteilungsstrategien für Unfallalgorithmen zu liefern, unvollständig sind. Aufgrund ihrer Erweiterungsbedürftigkeit begründen sie weiterführende Forschungsdesiderate.

## **8.2 Kritische Reflexion und Ausblick: Wissenschaftliche Relevanz, Forschungsdesiderate und Limitationen**

Im Rahmen der Forschungsarbeit wurde das anspruchsvolle Vorhaben verfolgt, Perspektiven der Digitalen Ethik, Metaethik und Risikoethik in einer pragmatisch orientierten Interpretation des Anwendungsproblems zusammenzuführen. Ziel war es dabei, den Problemhorizont bisheriger dominanter Forschungszugänge über eine moralphilosophische Sichtweise hinaus zu erweitern und auf diese Weise Raum für die Berücksichtigung theoretisch-formaler Grundlagen einerseits und anwendungsspezifischer Determinanten andererseits zu schaffen. Der Auftrag war es, eine dezidiert ethische Betrachtungsweise zu erarbeiten, die insbesondere dem Risikoaspekt moralischer Unfalldilemmata Rechnung trägt.

Der Beitrag, den dieses Buch zur Weiterentwicklung des Forschungsdiskurses leistet, besteht zu wesentlichen Teilen in der Begründung der Relevanz und Eignung einer risikoethischen Neuinterpretation des Anwendungsproblems. So wurde nicht nur erläutert, über welche Potenziale die Risikoethik bei der Gestaltung von Unfallalgorithmen verfügt, sondern es wurden auch konzept-

tionelle Grundlagen gelegt, auf die weiterführende Forschung aufbauen kann. Dabei ging es ausdrücklich nicht darum, bisherige Forschungsansätze grundsätzlich zurückzuweisen, sondern lediglich zu pointieren, inwiefern diese durch die Integration einer risikoethischen Perspektive profitieren können. Zudem bleiben die erarbeiteten Ergebnisse nicht auf Dilemma-Situationen beschränkt, sondern lassen sich ebenfalls auf Routinefahrsituationen anwenden, in denen – von uns oftmals unbemerkt – implizite Abwägungen zwischen persönlichen Zielen und Werten stattfinden. Wenn etwa Fahrentscheidungen über eingehaltene Abstände bei alltäglichen Überholvorgängen getroffen werden, bilden sich situationsspezifische Risikokonstellationen, die von nicht zu vernachlässigender ethischer Relevanz sind.

Als kritische Reflexion auf die erreichten Ergebnisse bleibt festzuhalten, dass die Limitationen der vorgelegten Untersuchung zugleich als Forschungsdesiderate zu verstehen sind. Diese bestehen zum einen hinsichtlich inhärenter Schwächen traditioneller ethischer Denkschulen, die im Hinblick auf das Anwendungsproblem zutage treten. Es wurde argumentiert, dass ein deontologisches Framework grundsätzlich offen für die Integration diverser ethischer Prinzipien ist. Die Forschungsarbeit verzichtet auf eine konkrete Ausarbeitung dessen, wie sich das neu definierte risikoethische Optimierungsproblem jenseits der skizzierten deontologischen Grenzkriterien von Zumutbarkeit und Fairness inhaltlich gestalten lässt. In welchem Umfang können konsequentialistische Elemente Beachtung finden? Inwieweit ist eine Schadensminimierung aus kontraktualistischer Sicht über Ex-Ante-Kompensation legitimierbar? An relevanten Stellen wurden diese und ähnliche offene Fragen explizit als solche benannt, um zukünftiger Forschung Anknüpfungspunkte zu bieten und kreative Ansätze zu inspirieren. Spezifische Probleme ethischer Theorien existieren grundsätzlich unabhängig von einer risikoethischen Sichtweise; im Zuge ihrer Bewältigung im Kontext des Anwendungsproblems müssen jedoch Risiken und Unsicherheiten als entscheidungstheoretische Komponenten Berücksichtigung finden.

Überdies ist die Sphäre moralischer Verantwortung für Risikoübertragungen im Rahmen dieser Forschungsarbeit weitgehend unbearbeitet geblieben. Eine fundierte und explizite Auseinandersetzung mit diesem Themenfeld bedarf mehr als einer Randnotiz, was dieses Buch aufgrund seiner Fokussierung auf den konkreten

## 8. Fazit und Ausblick

Anwendungsfall von Unfallalgorithmen sowie risikoethische Fragen der Zulässigkeit und Fairness nur unzureichend leisten kann. Hierbei sollten auch Perspektiven Erwähnung finden, die auf technikinduzierte Risiken im Speziellen rekurrieren. So wird es mit zunehmender Automatisierung technischer Innovationen immer schwieriger, mögliche Schadensfolgen zuverlässig abzuschätzen und verantwortlichen Instanzen zuzurechnen. Die Grenzen der Verantwortlichkeit verschwimmen angesichts kollektiver Entscheidungslagen sowohl im vorwärts- als auch im rückwärtsgerichteten Sinne oder müssen hinsichtlich der Verantwortungsunfähigkeit autonomer Maschinen neu ausgelotet werden.

Zum anderen stellt der vorgelegte Entwurf keine abgeschlossene risikoethische Untersuchung dar, sondern ist vielmehr als skizzenhafter Impuls für weitere risikoethische, inter- und transdisziplinäre Forschung zu verstehen. Forschungsdesiderate bestehen z. B. hinsichtlich der Risikobewertung bzw. Risikoabschätzung, die notwendig ist, um zu implementierbaren Konzeptionen zu gelangen. Die Definition konkreter handlungsrelevanter Prinzipien und Algorithmen für spezifische Dilemma-Szenarien liegt ebenfalls jenseits des definierten thematischen Rahmens dieser Forschungsarbeit. Diese schafft jedoch eine argumentative Grundlage, auf der Konzepte einer möglichen Implementierung aus ethischer Sicht gerechtfertigt oder zurückgewiesen werden können. Technische Umsetzbarkeit allein ist nicht ausreichend, um die Wahl bestimmter Algorithmen zu begründen. Es ist zu hoffen, dass die vorgelegte risikoethische Auseinandersetzung als Hilfestellung dienen und auf diese Weise zu einer rechtfertigbaren Entscheidungsfindung beitragen kann.