

1. Einführung in die Untersuchung

als Hochrisikooanwendungen eingestuft werden, rückt ebenfalls eine risikobasierte Betrachtungsweise in den Vordergrund.

Die Forschungsarbeit lässt sich im weiteren Kontext dieses neuen, risikofokussierten Forschungszugangs verorten. In Abgrenzung zu bisher publizierten Beiträgen wird in diesem Buch eine systematisch entwickelte, dezidiert risikoethische Untersuchung vorgelegt, auf deren Basis sich sowohl Dilemma-Szenarien als auch alltägliche Fahrsituationen ganzheitlich ethisch wie auch gesellschaftlich-sozial interpretieren lassen. Das Fehlen entsprechender integrativer Beiträge, die dies im gegebenen Anwendungskontext leisten, kann als »blinder Fleck« eines Diskurses verstanden werden, der sich gerade neuformiert:

The discussion of risk and AVs is just beginning. We're at the stage where (a) a good case has been made for the importance of the discussion, and where (b) a smattering of different scenarios and questions about risk has been posed. (Evans, 2022, S. 8)

Der erarbeitete risikoethische Entwurf versteht sich dabei als Versuch, anhand einer risikoethischen Neuinterpretation des zugrundeliegenden Entscheidungsproblems zukünftige Auseinandersetzungen mit Unfallalgorithmen zu inspirieren und auf diese Weise den Diskurs voranzubringen. Die Ergebnisse dieser Forschungsarbeit bewegen sich auf einer Mesoebene, d. h. sie stellen ausdrücklich keine unmittelbar implementierbaren, konkreten Entscheidungsstrategien bzw. Normen dar, sondern schlagen vielmehr Rahmenbedingungen in Form risikoethischer Grenzkriterien vor, die einer rechtfertigbaren Risikopraxis als Grundlage dienen können.

1.3 Methodik und Struktur der Untersuchung

1.3.1 Methodischer Ansatz

Als praktisches Entwicklungsprojekt und zugleich gesellschaftliches Phänomen befindet sich das autonome Fahren an der Schnittstelle verschiedener Forschungsdisziplinen. Obwohl die Problematik moralischer Dilemma-Situationen primär eine ethische Fragestellung darstellt, muss sie aufgrund der spezifischen Merkmale ihres Kontextes als inter- und transdisziplinäres Problem aufgefasst werden. So bezieht sie an verschiedenen Stellen auch Methoden und Konzepte

anderer wissenschaftlicher Disziplinen, z. B. der Ingenieurwissenschaften, Soziologie und Rechtswissenschaften, mit ein. Als philosophische Untersuchung einer real-lebensweltlichen Fragestellung ist das Forschungsvorhaben im Bereich der Angewandten Ethik anzusiedeln. Ihr methodisches Vorgehen entspricht dem Selbstverständnis einer Angewandten Ethik, die sich selbst stets in interdisziplinärer Perspektive begreift, wobei sie anwendungsnah, aber zugleich auch theoretisch-wissenschaftlich reflektierend und begründend vorgeht. Damit beschreitet sie einen methodischen Pfad, den Filipović (2016, S. 46) als »Mittelweg zwischen ethischer Theoriebildung und erfahrungsbezogener Normfindung« beschreibt. In dieser Arbeit wird beabsichtigt, das Anwendungsproblem in seinen ethisch relevanten Facetten aufzugreifen und eine Antwort zu entwickeln, die auf der Basis einer kritischen Reflexion bestehende Entscheidungsstrategien zurückweist bzw. in begründeter Weise ergänzt.

Einer der originären Beiträge, welche dieses Buch zum Forschungsdiskurs leistet, besteht in der metaethischen Rekonstruktion der spezifischen Problemstruktur, die Unfalldilemmata kennzeichnet. Dabei bleiben metaethische Methoden jedoch auf den Rahmen des fünften Kapitels beschränkt; ihre Funktion ist es, die Problemstellung aus der Sichtweise des metaethischen Diskurses zu beleuchten und das zentrale Argument zu erweitern, welches sich aus der kritischen Analyse bisheriger Forschungszugänge ergibt.

Die interdisziplinäre Dimension des Anwendungsproblems bringt direkte Implikationen für die wissenschaftliche Standortbestimmung dieser philosophischen Abhandlung mit sich. Sie steht im weiteren Sinne im Kontext jenes technikethischen Diskurses, der sich mit solchen algorithmischen Entscheidungen befasst, die eine moralische Dimension aufweisen. Im engeren Sinne ist sie an der Schnittstelle von Maschinenethik, Digitaler Ethik und Ethik der Künstlichen Intelligenz zu verorten. Durch die Ausarbeitung einer risikoethischen Perspektive auf das Anwendungsproblem ist sie vor allem auch in der Risikoethik anschlussfähig. Diese widmet sich als Teilgebiet der Ethik denjenigen Problemstellungen, die im Zusammenhang mit der moralischen Bewertung von unsicheren und risikobehafteten Handlungen im Kontext gesellschaftlicher Risiken anzusiedeln sind. Dies schließt technikinduzierte Risiken ausdrücklich ein. Während beispielsweise die Maschinenethik danach fragt, inwiefern Entscheidungen über Leben und Tod von Maschinen getroffen werden kön-

nen bzw. sollen, verfolgt die Risikoethik das Ziel, rechtfertigbare risikopraktische Kriterien und Rahmenbedingungen für den Einsatz künstlicher Systeme zu definieren. Die forcierte Integration verschiedener disziplinspezifischer ethischer Zugänge ist ein Charakteristikum dieses Buches, das Elemente der Digitalen Ethik, der Metaethik und der Risikoethik zusammenführt.

1.3.2 Ziele und Hypothesen

Ein *erstes Teilziel* der vorliegenden Untersuchung ist es, die Problemstellung innerhalb des Forschungsdiskurses zu verorten und dabei die Bedeutung und Relevanz von Unfallszenarien sowohl in praktischer als auch theoretischer Hinsicht zu evaluieren (*Zwischenergebnis*). Das *zweite Teilziel* besteht in einer kritischen Reflexion moralphilosophischer Problemzugänge, die Unfallalgorithmen als moralisches Designproblem interpretieren und den Forschungsdiskurs seit seinen Anfängen dominieren. Eine differenzierte Auseinandersetzung legt offen, dass unter dieser Problemperspektive zentrale Fragen ungeklärt bleiben; diese stehen der finalen Begründbarkeit eines moralphilosophisch fundierten Designs von Unfallalgorithmen entgegen und beinhalten zugleich die Forderung nach alternativen Zugängen (*zweites Zwischenergebnis*). Anhand des *dritten Teilziels* wird die Absicht verfolgt, mithilfe einer metaethischen Rekonstruktion der spezifischen Wertekonflikte, die den relevanten moralischen Dilemma-Strukturen zugrunde liegen, eine ganzheitliche Perspektive auf die Problemstellung zu eröffnen. Auf diese Weise wird die Notwendigkeit einer pragmatischen, der praktischen Sache dienlichen Ausrichtung möglicher Strategien vor dem Anwendungshintergrund von Unfalldilemmata begründet (*drittes Zwischenergebnis*). Schließlich wird im Rahmen des *vierten Teilziels* ein alternativer Problemzugang entworfen, der eine risikoethische Perspektive auf die spezifische Problematik von Unfallalgorithmen einnimmt und durch eine pragmatische Herangehensweise an zum jetzigen Zeitpunkt ungeklärte ethische Fragen den Forschungsdiskurs bereichert.

Die Ziele der Abhandlung werden in Form zweier forschungsleitender Arbeitshypothesen konkretisiert. Die Bearbeitung des zweiten Teilziels erfolgt im Hinblick auf die (*erste*) These, dass der bisher dominante moralphilosophische Zugang zur Problemstellung aufgrund methodischer sowie inhaltlicher und problemstruktureller Schwä-

chen zentrale Fragen offenlässt. Er ist daher inadäquat, um glaubwürdige und begründbare Entscheidungsstrategien hinsichtlich des Anwendungsproblems zu etablieren. Im Zuge des dritten Teilziels wird die Plausibilität dieser These gestützt. Anhand einer Analyse der metaethischen Grundlagen dilemmatischer Wertekonflikte wird das Fehlen systematischer Strategien zur Entscheidung von Unfalldilemmata aufgezeigt und gleichzeitig das vielversprechende Potenzial pragmatischer Entscheidungsstrategien akzentuiert. Im Anschluss an die erste These wird im Rahmen des vierten Teilziels eine risikoethische Auseinandersetzung vorgelegt. Dieser liegt die (*zweite*) These zugrunde, dass sich unter einem risikoethischen Zugang zentrale Fragen des Anwendungsproblems klären lassen und normative Implikationen freigelegt werden können, die neuen, vielversprechenden Entscheidungsperspektiven den Weg bereiten.

1.3.3 Gedankengang

Der Argumentations- und Gedankengang dieses Buches folgt einem Narrativ, das sich in drei Teile gliedert und dabei an den forschungsleitenden Thesen orientiert.

Teil I: Autonomes Fahren und Unfalldilemmata: Ethischer Problemhorizont und Relevanz

Die Untersuchung beginnt mit einer einführenden Darstellung des autonomen Fahrens sowohl als technologisches Entwicklungsprojekt als auch gesellschaftliches Phänomen. Hinsichtlich der Funktion der Verkehrsautomatisierung als eines zentralen Elementes der anvisierten Mobilitätswende westlicher Gesellschaften werden in Kap. 2 Evolution, Agenda, Ziele und Herausforderungen der Entwicklung autonomer Fahrzeuge knapp skizziert und einer kritischen Betrachtung unterzogen. Ziel ist es hierbei, den praktischen Forschungsgegenstand in der Bandbreite seiner vielschichtigen Facetten darzustellen und die Problematiken anzudeuten, die sich für verschiedene Forschungsdisziplinen ergeben. In Kap. 3 werden sodann ethische Herausforderungen fokussiert, die sich im Kontext von Entwicklung und Einsatz autonomer Fahrsysteme stellen. Neben einer überblicksartigen Darstellung verschiedener ethischer Diskurse wird das Hauptaugenmerk auf jene Unfallalgorithmen gelenkt, die im Fall unvermeidbarer Kollisionen aktiviert werden und in dabei entste-

henden moralischen Entscheidungsdilemmata die Aktionen autonomer Fahrzeuge steuern. Diese Problemstellung, die im Zentrum der Forschungsarbeit steht, wird sowohl in ihrer ethischen als auch informationstechnischen Dimension thematisiert. Schließlich werden Bedeutung und Relevanz moralischer Dilemma-Szenarien für das autonome Fahren erörtert.

Teil II: Problemzugänge in zwei Diskursen: Darstellung und Kritik

Im zweiten Teil wird das zentrale Argument entwickelt, welches die erste These des Vorhabens begründet: Ein Zugang zur Gestaltung von Unfallalgorithmen, wie er bisher im einschlägigen Forschungsdiskurs diskutiert wurde, lässt (zu) viele Fragen offen. Hier erfolgt eine kritische Auseinandersetzung mit dem Fokus bisheriger Forschung, unter dem moralische Dilemma-Situationen im Kontext autonomer Fahrsysteme beinahe ausschließlich als Problematik moralischer Designentscheidungen betrachtet werden. Anhand einer umfassenden Rekonstruktion der relevanten Forschungsliteratur wird in Kap. 4 der Nachweis geführt, dass der bisher dominante Forschungszugang sowohl in methodischer als auch inhaltlicher und struktureller Hinsicht erhebliche Schwächen an den Tag legt. Im Rahmen von Kap. 5 wird dieses Argument um eine metaethische Perspektive ergänzt, mittels derer die zugrundeliegende Problemstellung moralischer Dilemmata aus metaethischer Sicht erörtert und hinsichtlich des Anwendungskontextes von Unfalldilemmata spezifiziert wird. Es wird substantiiert, weshalb pragmatische Strategien zur Bewältigung echter moralischer Dilemmata besonders vielversprechend sind.

Teil III: Risikoethische Auseinandersetzung: Entwurf eines alternativen Problemzugangs

Im dritten Teil wird auf Basis der bisherigen Ergebnisse schließlich ein alternativer Zugang zum Anwendungsproblem entworfen. Dabei wird das Feld der Risikoethik beschritten, die – wie die zweite These der Forschungsarbeit postuliert – adäquate Mittel bereitstellt, um sowohl die Spezifika des praktischen Problemkontextes als auch deren theoretische Dimension im Sinne einer pragmatischen Herangehensweise zu integrieren. In Kap. 6 werden im Rahmen einer theoretischen Grundlegung und Einführung in wissenschaftliche Verortung, Historie, Gegenstandsbereich und Paradigmen der Risikoethik

zunächst begriffliche und konzeptionelle Grundlagen gelegt. Es wird gezeigt, dass eine rationale Risikopraxis angesichts der spezifischen Problematik von Unfallalgorithmen nicht glaubwürdig ist. In Kap. 7 werden sodann die Grundzüge einer deontologischen Risikoethik systematisch entwickelt. Im Zuge einer eingehenden risikoethischen Analyse wird die im Diskurs bisher dominante Frage nach der moralphilosophischen Begründung von Entscheidungsprinzipien für Unfallalgorithmen zur Frage der ethischen Rechtfertigbarkeit von Dilemma-Risiken sowie der fairen Verteilung daraus resultierender Vor- und Nachteile transformiert. Es wird für eine kohärente Risikopraxis plädiert, welche die Gestaltung von Unfallalgorithmen als Optimierungsproblem begreift, dem durch deontologische Grenzkriterien unverhandelbare Beschränkungen auferlegt sind.

Zum Abschluss werden die zentralen Ergebnisse der Untersuchung in Kap. 8 nochmals zusammengefasst und im Hinblick auf ihre Bedeutung für weiterführende Forschung kritisch evaluiert.

