

schäden kommt? Inwiefern lassen sich Abwägungen zwischen individuellen Interessen rechtfertigen? Antworten auf diese und ähnliche Fragestellungen müssen sich einerseits in einen bestehenden Rechtsrahmen einfügen und zugleich technisch umsetzbar sein. Andererseits hängen sie in entscheidender Weise von gesellschaftlich geprägten, normativen Moral- und Wertvorstellungen ab, welche angesichts der fortschreitenden Digitalisierung neu zur Diskussion gestellt werden müssen.

Gibt es im moralischen Sinne keine eindeutige Lösung, ist unklar, wie selbstfahrende Fahrzeuge in entsprechenden Fällen agieren sollen. Aus ethischer Sicht lässt sich das Entscheidungsproblem, das unabwendbaren Unfallsituationen zugrunde liegt, aufgrund seiner komplexen Struktur als kontextspezifische Instanz moralischer Dilemmata interpretieren. Letztere repräsentieren einen spezifischen Typ moralischer Entscheidungsprobleme, bei dem sich miteinander inkompatible Handlungsalternativen gegenüberstehen, die alle aus moralischen Gründen jeweils richtig und falsch zugleich sind: Richtig in dem Sinne, dass ein moralischer Grund die jeweilige Handlung einfordert, und zugleich falsch in dem Sinne, dass durch die Wahl einer Option zwangsläufig diejenigen moralischen Gebote vernachlässigt werden, die mit den Alternativen assoziiert sind. Sowohl in der philosophischen Ethik als auch in der Moralphilosophie haben moralische Dilemmata eine lange Tradition. In Anlehnung an diese wurde bislang häufig versucht, die Problematik von Unfalldilemmata im Kontext des autonomen Fahrens mithilfe traditioneller ethischer Denkmuster zu adressieren. Diese stoßen aufgrund der spezifischen Vielschichtigkeit des Anwendungsproblems allerdings an ihre Grenzen. Trotz des seit nunmehr zehn Jahren andauernden Forschungsdiskurses ist die zentrale Frage, wie sich Entscheidungsstrategien in konkreten dilemmatischen Fällen ethisch begründen lassen, noch immer weitgehend ungeklärt.

1.2 Erkenntnisinteresse und Relevanz

Seit dem Beschluss über das »Gesetz zum autonomen Fahren« im Mai 2021 schreitet die Entwicklung autonomer Fahrzeuge auch in Deutschland rasant voran. Auch wenn öffentlich kommunizierte Zeitpläne meist ambitioniert sind und laufend korrigiert werden,

so ist dennoch klar: Der Tag, an dem von Fahrrobotern gesteuerte Fahrzeuge – in welchem Umfang auch immer – unsere Straßen im Regelbetrieb bevölkern werden, rückt immer näher. Derzeit befinden wir uns in einer Übergangsphase, in der erste Hersteller die Zulassung für Fahrzeuge mit automatisierten Teilfunktionen erhalten haben, so beispielsweise Mercedes-Benz mit seinem Autobahn-Staupiloten »Drive Pilot« (vgl. Rudschies & Kroher, 2024) oder die nächste Generation des BMW-Autobahnpiloten mit integrierter Überholfunktion per Blicksteuerung, verbaut im neuen Modell »i5« (vgl. Geiger et al., 2024). Treiber der Entwicklung des autonomen Fahrens sind einerseits der technische Innovationsdruck, der auf den Schultern der Automobilbauer lastet, und andererseits die Neuverfassung der Rechtslage durch die gesetzgebenden Institutionen.

Vor diesem Hintergrund drängt angesichts der raschen Entwicklung der vernetzten und automatisierten Mobilität auch die Klärung ethischer Fragen. Moralische Dilemma-Szenarien zählen dabei zu den theoretisch und praktisch bedeutungsvollsten ethischen Herausforderungen. Eine vollumfängliche Automatisierung der Individualmobilität muss den Steuerungsalgorithmen autonomer Fahrsysteme idealerweise für jegliche denkbare Situation eine Handlungsempfehlung an die Hand geben. Besonders brisant aus ethischer Sicht ist diese Thematik nicht zuletzt deshalb, weil in einem automatisierten Verkehrsgeschehen zukünftig Algorithmen moralische Entscheidungen »treffen« müssen, die bisher Menschen getroffen haben; aus instinktiven Handlungen werden systematische (vgl. Baker et al., 2018). In Extremfällen erfordern Unfalldilemmata Entscheidungen über Leben und Tod; ethisch relevant sind sie jedoch auch ohne die tragische Komponente, die Szenarien aus Tages- und Wochenzeiten oft anhaftet. Dilemma-Szenarien pointieren ein Problem, das sich auch in alltäglichen Fahrsituationen stellt: Entscheidungen über Abstände, Geschwindigkeiten oder Trajektorien implizieren allesamt ein Abwägen der Interessen derjenigen, die von möglichen Handlungen in einer spezifischen Situation betroffen sind.

Die ethische Gestaltung von Unfalgorithmen wurde im bisherigen Forschungsdiskurs mehrheitlich unter der zentralen Fragestellung diskutiert, an welchen ethischen Prinzipien sich die betreffenden Systemalgorithmen bei der Entscheidungsfindung orientieren sollen. Von Expertengremien entwickelte Richtlinien erwiesen sich bisher als zu unkonkret; sie sind lediglich als Empfehlungen zu

verstehen, denen es an Durchsetzungskraft mangelt. So wurden im Bericht der vom Bundesministerium für Verkehr und Digitale Infrastruktur (BMVI) eigens zur Reflexion ethischer Fragen in Bezug auf das autonome Fahren eingesetzten Kommission aus dem Jahr 2017 die Bandbreite ethischer Herausforderungen aufgezeigt sowie zentrale ethische Prinzipien festgelegt, z. B. ein Diskriminierungsverbot, der Schutz des Lebens vor anderen Erwägungen wie Mobilitätschancen oder eine Transparenz in Verantwortungsfragen (vgl. Lütge et al., 2020).¹ Was verbleibende offene Fragen anbelangt, wurde hingegen lediglich auf weiteren Forschungsbedarf verwiesen. Auch der nunmehr zehn Jahre währende internationale ethische Forschungsdiskurs wurde zwar – analog zur fortschreitenden technischen Entwicklung – mit der Zeit komplexer und vielschichtiger, konnte aber noch keine durchschlagskräftigen Strategien präsentieren. Dies ist vor allem darauf zurückzuführen, dass bisherige Forschungszugänge zu viele Fragen offenlassen, um als praktische Entscheidungshilfe in Erwägung gezogen zu werden.

In den letzten zwei bis drei Jahren kristallisierte sich eine neue, bis dato noch unterrepräsentierte Forschungsströmung heraus, die eine alternative Sichtweise auf die Problemstellung einnimmt: Dilemma-Szenarien werden dabei nicht länger als ein moralisches Designproblem betrachtet, das mittels traditioneller Ansätze der Moralphilosophie zu problematisieren ist, sondern vielmehr als risikoethisches Entscheidungsproblem im Hinblick auf eine rechtfertigbare Risiko-praxis. Eine Auffassung als risikoethische Fragestellung erfreut sich zunehmender Aufmerksamkeit innerhalb des Forschungsfelds. Diese spiegelt sich nicht zuletzt in einer Vielzahl laufender oder kürzlich abgeschlossener Forschungsprojekte wider, die auf Bundesebene gefördert und zumeist in enger Verzahnung von technologischer Entwicklung durch Unternehmen oder unternehmensnahe Forschungsinstitutionen einerseits sowie Theoriebildung in universitären oder HAW-Instituten andererseits bearbeitet werden. Das kürzlich in Kraft getretene Europäische Gesetz zur Künstlichen Intelligenz (*EU Artificial Intelligence Act/AI Act*), demzufolge autonome Fahrsysteme

1 Auf den Bericht der Ethik-Kommission wird im Verlauf dieser Forschungsarbeit an verschiedenen Stellen zurückgegriffen. Auch wenn es sich dabei nicht um starke Argumente, sondern vielmehr um Richtlinien mit pragmatischem Fokus handelt, werden sie jedoch nicht ganz unbegründet in den Raum gestellt.