American Journal of Applied Science and Technology

VOLUME Vol.05 Issue 10 2025
PAGE NO. 291-296

Ethical Trajectories and Liability Architectures in Autonomous Vehicles: Bridging Moral Theory, Regulation, and Technical Design

Dr. Marcus Léonard

Department of Science, Technology, and Society, Université de Lyon

Received: 01 October 2025; **Accepted:** 14 October 2025; **Published:** 31 October 2025

Abstract

Background: The deployment of autonomous vehicles (AVs) has foregrounded a dense intersection of ethical decision-making, regulatory design, liability distribution, and algorithmic transparency. Existing scholarship highlights normative dilemmas (trolley-like trade-offs), heterogeneous societal preferences, and emergent legal ambiguities that together complicate responsible AV rollout (Bonnefon et al., 2016; Awad et al., 2018; Kriebitz et al., 2022).

Objective: This article synthesizes ethical theory, empirical findings on moral preferences, regulatory responses, and technical approaches to AV decision-making to propose an integrated framework—an ethical trajectory and liability architecture—that supports accountable AV behavior while remaining sensitive to plural moral perspectives and evolving technological constraints.

Methods: The study conducts an integrative conceptual analysis grounded strictly in the supplied literature. It combines normative ethical analysis, comparative interpretation of empirical moral-choice studies, and a technical review of ethical trajectory planning, localization and sensing, and simulation frameworks. The approach is methodical: (1) map moral problems identified in empirical and philosophical literature; (2) derive design principles for AV control systems and ethics modules; (3) analyze legal and liability frameworks; (4) propose a governance architecture aligning ethics-by-design with legal accountability. Each step is substantiated with cited findings and theoretical elaboration.

Results: The synthesis reveals three interdependent domains that must be aligned: (A) an ethics-of-risk orientation in system behavior rather than sole reliance on trolley-problem solutions (Geisslinger et al., 2021; Geisslinger et al., 2023); (B) a layered liability model distributing responsibility across manufacturers, software suppliers, and operators in accordance with foreseeability and control (Marchant & Lindor, 2012; Douma & Palodichuk, 2012; Xiao & Cao, 2017); (C) technical design patterns that operationalize fairness, transparency, and risk-sensitive planning through ethical trajectory planners, advanced localization, and sensor simulation-informed validation (Geisslinger et al., 2023; Kuutti et al., 2018; Elmquist et al., 2021).

Conclusions: Rather than seeking a single universal moral algorithm, effective governance of AVs requires integrated socio-technical architectures that explicitly trade off risks, embed normative pluralism into configurable policy layers, and align incentives via liability reforms. The proposed ethical trajectory and liability architecture offers a practical pathway for policymakers, designers, and legal actors to operationalize ethical values while safeguarding public trust and innovation. Future work must translate these conceptual prescriptions into empirical validation, juridical pilots, and participatory governance experiments. (Max ~400 words)

Keywords: autonomous vehicles, ethics of risk, liability, ethical trajectory planning, regulation, moral preferences

INTRODUCTION

The emergence of autonomous vehicles (AVs) constitutes a paradigmatic case of technology creating novel moral, legal, and social challenges. Philosophical debates and empirical investigations have repeatedly demonstrated that AVs do not simply port existing vehicles to automated control; they reconfigure agency, culpability, and societal expectations about road safety (Bonnefon et al., 2016; Awad et al., 2018). Two research strands dominate the current scholarship: the first interrogates public moral intuitions and their universals and variants (Awad et al., 2018; Awad et al., 2020), while the second scrutinizes legal liability and regulatory readiness (Marchant & Lindor, 2012; Kriebitz et al., 2022). Complementing these are technical contributions that propose concrete algorithmic mechanisms for ethically-informed planning (Geisslinger et al., 2023) and validation frameworks that assess sensor and localization components under realistic conditions (Kuutti et al., 2018; Elmquist et al., 2021). Together, these literatures paint a complex picture: neither purely philosophical prescriptions nor purely technical solutions suffice; regulatory responses must be informed by both empirical moral data and the practical limits of sensing and control.

The dominant metaphor in early AV ethics discussions was the trolley problem: hypothetical, simplified dilemmas that force an agent to choose between two harms (Bonnefon et al., 2016). While useful for probing moral intuitions, the trolley framing has been criticized as insufficiently representative of real driving contexts, which are dynamic, probabilistic, and characterized by incompletely specifiable outcomes (Geisslinger et al., 2021). Empirical work across countries reveals both strong commonalities in basic moral patterns and significant variation that challenges the aspiration of a single, globally acceptable moral policy (Awad et al., 2018; Awad et al., 2020). From a legal standpoint, the introduction of AVs complicates fault-based liability systems that presuppose a human agent with volitional control and immediate foreseeability (Marchant & Lindor, 2012; Douma & Palodichuk, 2012). Consequently, policymakers have begun experimenting with regulatory instruments and explicit ethics clauses, as exemplified in jurisdictional developments such as the German Act on Autonomous Driving (Kriebitz et al., 2022), which acknowledges the continuing relevance of ethics in legislation.

This article aims to synthesize these disparate threads into a coherent, actionable framework that can guide both the engineering of AV ethics modules and the reform of liability regimes. The paper approaches the challenge by (1) critically reevaluating the trolley paradigm in light of the ethics-of-risk perspective, (2) mapping empirical moral preferences to governance-relevant design choices, (3) proposing an ethics-by-design architecture centered on ethical trajectory planning and rigorous simulation-informed validation, and (4) sketching a layered liability architecture that aligns with the technical realities of system design and public expectations. The methodological approach is integrative and strictly anchored in the references provided.

METHODOLOGY

The study employs an integrative conceptual-methodological approach designed to synthesize philosophical, empirical, legal, and technical sources into a practicable framework. Given the interdisciplinary nature of the problem and the instruction to work strictly with the supplied references, the methodology proceeds in four interlocking analytic stages.

1. Normative Landscape Mapping: We systematically extract normative claims, problem framings, and theoretical critiques from core philosophical and empirical texts that interrogate AV ethics. Sources such as Bonnefon et al. (2016), Awad et al. (2018, 2020), and Geisslinger et al. (2021, 2023) are read to identify recurring dilemmas (e.g., harm-minimization, distributive fairness, individual vs. collective preferences) and meta-ethical considerations (e.g., universality vs. cultural variation, procedural fairness).

2. Empirical-to-Design Translation: Empirical studies of moral preferences (Awad et al., 2018; Awad et al., 2020; Caro-Burnett & Kaneko, 2022) are used to infer design desiderata for AV decision-making systems. The translation emphasizes how statistical regularities and culturally contingent responses should influence configurable policy layers rather than prescribe a single algorithmic choice.

3. Technical Review and Synthesis: Technical literature on ethical trajectory planning (Geisslinger et al., 2023), localization and sensing potential (Kuutti et al., 2018), and simulation frameworks for validation (Elmquist et al., 2021) is analyzed to identify feasible implementation patterns, verification constraints, and residual uncertainties. The aim is to ground normative prescriptions in implementable techniques and to identify where trade-offs are unavoidable.

4. Legal Architecture Analysis: Legal scholarship on liability and criminal exposure for AVs (Marchant & Lindor, 2012; Douma & Palodichuk, 2012; Xiao & Cao, 2017) is synthesized to design a layered liability model. The analysis focuses on foreseeability, control, and the distributive consequences of assigning responsibility across OEMs, software providers, and vehicle owners/operators.

Throughout, the method adheres to strict citation: every major claim references the appropriate sources. The final product is an integrative framework—the ethical trajectory and liability architecture—whose components are justified by cross-referencing the literature and by elaborating theoretical implications, counterarguments, and normative trade-offs.

RESULTS

The synthesis yields a multi-part result: (A) reframing ethics for AVs around risk-managed behavior; (B) specifying design principles for ethics-aware AV systems; (C) formulating a layered liability model; and (D) detailing governance implications and practical deployment strategies. Each is described below with careful attention to supporting literature.

A. Ethics of Risk: Moving Beyond the Trolley Paradigm

Philosophical and empirical literature converge on a core insight: realistic driving decisions are rarely pure, discrete choice dilemmas of the trolley type; instead, they unfold as probabilistic assessments of an evolving risk landscape (Geisslinger et al., 2021). The trolley problem remains heuristically valuable for highlighting moral intuitions but misrepresents the decision space of AVs by abstracting away uncertainty, sensor noise, temporal dynamics, and the layered causes behind outcomes. Empirical experiments with large, culturally diverse samples demonstrate both universals (e.g., preference for minimizing fatalities) and cross-national differences in the weighting of particular victim characteristics, which suggests that any ethical module must be sensitive to pluralistic inputs rather than enforcing a single moral calculus (Awad et al., 2018; Awad et al., 2020).

The ethics-of-risk orientation reframes ethical behavior as the calibration of operational policies to minimize expected harm while respecting procedural fairness and legal constraints. Practically, this means incorporating probabilistic risk assessments, continuous trajectory optimization to reduce exposure to high-consequence events, and explicit policy layers that constrain the search space of permissible maneuvers. Geisslinger et al. (2021) articulate an argument for ethics that centers on anticipated risk reduction and proportionality rather than deterministic moral absolutes; this orientation is better aligned with the realities of sensing, localization, and control.

B. Design Principles for Ethics-Aware AV Systems

From the technical literature, several design principles emerge as necessary to instantiate an ethics-of-risk orientation in AV systems:

1. Layered Policy Architecture: Ethical decision-making should be implemented as multiple layers: a normative policy layer (high-level constraints that instantiate societal/legal values), a tactical planner (ethical trajectory planner that operationalizes those constraints in short-term motion decisions), and a low-level controller respecting physical vehicle dynamics (Geisslinger et al., 2023). This layering preserves clear boundaries between value-laden choices and control-level safety mechanisms.

2. Configurable Ethical Profiles: Given documented cross-cultural and intra-societal variation in moral preferences, AV systems should support configurable ethical profiles that reflect jurisdictional regulation and local social norms (Awad et al., 2018; Caro-Burnett & Kaneko, 2022). Configuration should be transparent, legislatively bounded, and subject to oversight.

3. Risk-Sensitive Trajectory Planning: Ethical trajectory planning must explicitly model probabilistic outcomes and expected harms, integrate uncertainty from localization and sensing modules, and prioritize maneuvers that minimize aggregated expected harm over time rather than focusing on instantaneous binary choices (Geisslinger et al., 2023; Kuutti et al., 2018).

4. Explainability and Record-Keeping: Transparent logging of sensor inputs, decision rationales, and trajectory choices is essential for post-incident analysis and public accountability. The logs must be structured to support both technical forensic analysis and legal evaluation (Elmquist et al., 2021).

5. Continuous Simulation-Based Validation: Given the gap between controlled testing and real-world complexity, AV ethics modules must be validated through extensive sensor and system simulation frameworks that reproduce edge-case scenarios and long-tailed risk distributions (Elmquist et al., 2021; Kuutti et al., 2018).

These principles operationalize normative aims into design constraints and technical mechanisms while acknowledging engineering limitations and legal expectations.

C. Layered Liability Model

Legal scholars anticipate tensions between existing fault-based liability systems and AVs, particularly when control is transferred from a human to an automated system (Marchant & Lindor, 2012; Douma & Palodichuk, 2012; Xiao & Cao, 2017). A layered liability architecture reconciles technical realities with legal norms by distributing responsibility according to control, foreseeability, and the allocation of decision-making authority.

1. Design/Manufacturing Liability: Original equipment manufacturers (OEMs) and suppliers retain primary responsibility for defects in hardware and system-level design flaws that unreasonably increase risk. This mirrors traditional product liability where foreseeable misuse and design defects attract manufacturer liability (Marchant & Lindor, 2012).

2. Algorithmic and Software Liability: Providers of critical software modules, including ethical trajectory planners and perception stacks, bear liability proportionate to the control they exert and the ability to foresee harmful outputs. Where software enumerates normative policy constraints, the provider may be liable if those constraints are inadequately specified, improperly tested, or misaligned with regulatory requirements (Xiao & Cao, 2017).

3. Configuration and Operator Liability: Vehicle owners and operators who intentionally modify ethical profiles or disable safety features should retain liability to the extent their actions create foreseeable risks. However, where operators have no meaningful control over software updates or configuration (e.g., locked profiles), shifting liability away from owners may be appropriate.

4. Shared/Strict Liability for High-Risk Failures: For high-consequence failures arising from systemic complexity—where pinpointing a single culpable actor is infeasible—the architecture recommends conditional strict liability regimes (e.g., mandatory compensation funds or insurer frameworks) balanced by robust recourse and audit rights to avoid moral hazard.

This layered approach aligns with the technical realities that AV outcomes result from the interaction of hardware, software, configuration, and environmental contingencies. It also incentivizes rigorous design and testing practices by tying liability to foreseeability and control.

D. Governance and Deployment Strategies

Operationalizing the above requires regulatory and governance practices that balance innovation incentives with public accountability. The German Act on Autonomous Driving demonstrates a legislative recognition that ethics still matters and that the law must address the socio-technical complexity of AVs (Kriebitz et al., 2022). Building on this precedent, the article proposes governance strategies:

1. **Mandated Ethical Impact Assessments:** Before market deployment, AV systems should undergo ethical impact assessments that evaluate how configurable profiles, safety constraints, and decision heuristics interact with local norms and legal obligations.

2. **Regulatory Sandboxes and Pilots:** Jurisdictions should permit controlled pilots with clear data-collection mandates, enabling iterative refinement of both technical systems and liability frameworks.

3. **Public Deliberation and Participatory Standard-Setting:** Because moral preferences vary and affect public trust, policymakers should institutionalize participatory mechanisms for translating societal values into regulatory constraints (Awad et al., 2018; Caro-Burnett & Kaneko, 2022).

4. Transparent Certification and Audit Regimes: Certification should require documentation of simulation validation, sensor performance, and ethical module behavior, supported by independent auditing bodies.

These strategies create a feedback loop between design, legal accountability, and societal acceptance, promoting safe and legitimate AV adoption.

DISCUSSION

The synthesis above presents an integrated way forward, but it also raises significant conceptual and practical challenges that merit deep analysis. This section delves into theoretical implications, counter-arguments, limitations, and directions for future work.

Theoretical Implications and Normative Trade-offs

The shift from trolley-problem salience to an ethics-of-risk orientation has important normative consequences. It reframes morality in operational terms—focusing on harm minimization under uncertainty and procedural fairness—rather than privileging particular outcomes in stylized dilemmas. This orientation aligns with consequentialist intuitions about harm reduction, yet it can be reconciled with deontological commitments by embedding hard constraints (e.g., prohibitions on discriminatory targeting) in the normative policy layer. Geisslinger et al. (2021) argue persuasively for such a trajectory-based ethics that is sensitive to risk; importantly, this avoids the unrealistic demand that AVs solve philosophical thought experiments in full generality.

However, several tensions remain. First, risk-based optimization can obscure distributive concerns: minimization of expected harm may systematically disadvantage certain groups if risk models are biased or if the social value assigned to certain outcomes reflects discriminatory priors (Feuerriegel et al., 2020). Thus, fairness must be an explicit constraint—not an emergent property—of planning systems (Feuerriegel et al., 2020). Second, preference heterogeneity complicates democratic legitimacy: if moral preferences differ across cultures and within populations (Awad et al., 2018; Awad et al., 2020), whose values should be encoded? The suggested solution—configurable ethical profiles within regulatory bounds—creates a pragmatic compromise but invites questions about cross-border interoperability and potential forum shopping by manufacturers seeking permissive jurisdictions.

A third theoretical issue concerns the moral significance of human agency. Some scholars emphasize the distinctive moral status of human volition in culpability and responsibility (Marchant & Lindor, 2012). When decisions are delegated to machines, moral intuitions about blame and redress may demand humanly meaningful oversight. The layered liability architecture responds to this by preserving incentives for human actors (designers, operators) to remain accountable and by requiring transparent logging to facilitate human normative assessment after incidents.

Counter-Arguments and Responses

Critics might argue that embedding ethics into machines is either infeasible or normatively perilous. If one believes ethics requires thick moral deliberation, any algorithmic implementation risks demeaning moral judgment. The counter is twofold: (1) autonomous systems inevitably make decisions with moral consequences—leaving them unregulated or governed ad hoc is a greater risk; (2) ethics-by-design can preserve moral deliberation by making policy choices explicit, contestable, and subject to democratic oversight (Kriebitz et al., 2022). Thus, ethical coding does not replace human moral agency; it institutionalizes it in traceable, auditable forms.

Another counter-argument concerns the feasibility of risk-sensitive trajectory planning in real time, given sensor limitations and computational constraints. Technical work on ethical trajectory planners demonstrates promising algorithms, but they rely on high-fidelity sensing and robust localization; simulation frameworks show how to test these systems but also reveal gaps between simulated and real-world complexity (Geisslinger et al., 2023; Elmquist et al., 2021; Kuutti et al., 2018). The policy implication is clear: deployment should be progressive and tied to demonstrable validation metrics, not premature mass-market release.

Limitations of the Synthesis

This article is intentionally bounded by two methodological constraints that influence the scope of conclusions. First, the analysis relies exclusively on the provided references; while these are diverse and authoritative, they do not exhaust the literature on AV ethics, regulation, or technical mechanisms. Second, the argument is predominantly conceptual and prescriptive rather than empirically validated: while the proposed architecture is grounded in technical proposals and empirical studies of moral preferences, its real-world effectiveness requires field trials, regulatory experimentation, and iterative refinement. The article thus offers a normative-technical roadmap rather than a finalized policy instrument.

Practical Challenges and Future Directions

Operationalizing the ethical trajectory and liability architecture entails numerous technical, legal, and sociopolitical hurdles:

1. **Standards Harmonization:** Jurisdictions will need to negotiate interoperability standards for configurable ethical profiles to avoid fragmentation while respecting local norms.

2. **Liability Insurance Models:** The layered liability model requires innovative insurance structures and possibly compensation funds to manage residual systemic risks and to prevent insolvency following catastrophic events.

3. **Auditability and Privacy:** Transparent logging is essential for accountability but must be balanced against privacy and security concerns. Research must refine log designs that are both legally informative and privacy-preserving.

4. **Participatory Governance:** Translating public moral preferences into policy requires new institutions for deliberation and mechanisms for updating ethic profiles as social norms evolve.

Empirical research should evaluate the social acceptability of configurable profiles, the technical efficacy of trajectory planners under diverse environmental conditions, and the legal robustness of the proposed liability allocations. Pilot programs and regulatory sandboxes will be crucial for stress-testing the architecture.

CONCLUSION

Autonomous vehicles are not merely technical artifacts; they are socio-technical systems that instantiate moral choices at scale. The initial dominance of trolley-style thought experiments offered early conceptual clarity but must give way to an ethics-of-risk framework that better captures the probabilistic, dynamic, and distributive complexities of real-world driving. Operationalizing ethical values in AVs requires a layered, transparent, and configurable design architecture, rigorous simulation-informed validation, and a liability regime that aligns responsibility with control and foreseeability. The integration of normative theory, empirical moral psychology, and technical engineering—anchored by sensible regulatory reforms such as those initiated in Germany—can produce AV systems that are both safer and more legitimate in the public eye.

The proposed ethical trajectory and liability architecture is a practical contribution toward that integration. It articulates design-level principles, legal allocations of responsibility, and governance strategies that together support accountable AV deployment. Realizing this vision requires multi-stakeholder collaboration, ongoing empirical evaluation, and the political will to align innovation incentives with public safety and justice.

REFERENCES

1. Kriebitz, A.; Max, R.; Lütge, C. The German Act on Autonomous Driving: Why ethics still matters. Philos. Technol. 2022, 35, 29.
2. Caro-Burnett, J.; Kaneko, S. Is society ready for AI ethical decision making? Lessons from a study on autonomous cars. J. Behav. Exp. Econ. 2022, 98, 101881.
3. Awad, E.; Dsouza, S.; Kim, R.; Schulz, J.; Henrich, J.; Shariff, A.; Bonnefon, J.-F.; Rahwan, I. The moral machine experiment. Nature 2018, 563, 59–64.
4. Awad, E.; Dsouza, S.; Shariff, A.; Rahwan, I.; Bonnefon, J.F. Universals and variations in moral decisions made in 42 countries by 70,000 participants. Proc. Natl. Acad. Sci. USA 2020, 117, 2332–2337.
5. Jean-François, B.; Azim, S.; Iyad, R. The social dilemma of autonomous vehicles. Science 2016, 352, 1573–1576.
6. Azim, S.; Jean-François, B.; Iyad, R. Psychological roadblocks to the adoption of self-driving vehicles. Nat. Hum. Behav. 2017, 1, 694–696.
7. Marchant, G.E.; Lindor, R.A. The coming collision between autonomous vehicles and the liability system. Santa Clara L. Rev. 2012, 52, 1321.
8. Douma, F.; Palodichuk, S.A. Criminal liability issues created by autonomous vehicles. Santa Clara L. Rev. 2012, 52, 1157.
9. Xiao, S.; Cao, J. On the Civil Liability of Artificial Intelligence. Sci. Law 2017, 35, 166–173.
10. Geisslinger, M.; Poszler, F.; Lienkamp, M. An ethical trajectory planning algorithm for autonomous vehicles. Nat. Mach. Intell. 2023, 5, 137–144.
11. Kuutti, S.; Fallah, S.; Katsaros, K.; Dianati, M.; Mccullough, F.; Mouzakitis, A. A Survey of the State-of-the-Art Localization Techniques and Their Potentials for Autonomous Vehicle Applications. IEEE Internet Things J. 2018, 5(2), 829–846.
12. Elmquist, A.; Serban, R.; Negrut, D. A Sensor Simulation Framework for Training and Testing Robots and Autonomous Vehicles. J. Auton. Veh. Syst. 2021, 1(2).
13. Patil, A. A.; Patel, N.; Deshpande, S. Ethical Decision-Making In Sustainable Autonomous Transportation: A Comparative Study Of Rule-Based And Learning-Based Systems. International Journal of Environmental Sciences 2025, 11(12s), 390–399.
14. Geisslinger, M.; Poszler, F.; Betz, J.; Lütge, C.; Lienkamp, M. Autonomous Driving Ethics: from Trolley Problem to Ethics of Risk. Philos. Technol. 2021, 34(4), 1033–1055.
15. Keeling, G.; Evans, K.; Thornton, S. M.; Mecacci, G.; Santoni de Sio, F. Four Perspectives on What Matters for the Ethics of Automated Vehicles. Lect. Notes Mobil. 2019, 49–60.
16. Poszler, F.; Geißlinger, M. AI and Autonomous Driving: Key ethical considerations. Inst. Ethics Artif. Intell. 2021.
17. Andersson, P.; Ivehammar, P. Benefits and Costs of Autonomous Trucks and Cars. J. Transp. Technol. 2019, 09(02), 121–145.
18. Feuerriegel, S.; Dolata, M.; Schwabe, G. Fair AI: Challenges and Opportunities. Bus. Inf. Syst. Eng. 2020, 62(4), 379–384.