

自动驾驶车辆肇事刑事责任的认定与归责体系完善

李鑫颖

辽宁科技大学经济与法律学院, 辽宁 鞍山

收稿日期: 2025年8月13日; 录用日期: 2025年8月25日; 发布日期: 2025年9月16日

摘要

随着自动驾驶技术的发展, 其交通肇事的刑事责任认定成为重要法律问题。本文通过分析自动驾驶肇事刑事责任的认定困境, 从生产者、使用者和第三方等维度解析刑事责任划分, 并提出“动态归责理论”, 即责任权重并非固定不变, 应根据自动驾驶层级、事故场景中各主体的实际控制力表现来进行动态调整, 更精准地体现“权责随控制力变化而适配”的核心主张, 以在保障技术发展与维护公共安全之间寻求平衡。

关键词

自动驾驶车辆, 交通肇事, 刑事责任

The Determination of Criminal Responsibility for Self-Driving Vehicles and the Improvement of the Liability System

Xinying Li

School of Economics and Law, Liaoning University of Technology, Anshan Liaoning

Received: Aug. 13th, 2025; accepted: Aug. 25th, 2025; published: Sep. 16th, 2025

Abstract

With the development of autonomous driving technology, the determination of criminal responsibility

文章引用: 李鑫颖. 自动驾驶车辆肇事刑事责任的认定与归责体系完善[J]. 法学, 2025, 13(9): 2142-2147.

DOI: 10.12677/ojls.2025.139295

for traffic accidents has become an important legal issue. This article analyzes the dilemma of determining criminal responsibility for autonomous driving accidents, and analyzes the division of criminal responsibility from the dimensions of producers, users, and third parties. It proposes the “dynamic attribution theory”, which means that the weight of responsibility is not fixed and should be dynamically adjusted based on the actual control performance of each subject in the autonomous driving hierarchy and accident scenario, more accurately reflecting the core proposition of “rights and responsibilities adapt to changes in control”, in order to seek a balance between ensuring technological development and maintaining public safety.

Keywords

Autonomous Vehicles, Traffic Accidents, Criminal Liability

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

随着 L3 级自动驾驶技术的规模化落地与 L4 级系统的商业化试点，自动驾驶汽车交通肇事的刑事责任认定已成为法律理论与司法实践面临的核心挑战。传统刑事责任理论在应对这一技术变革时逐渐显现适配性缺陷，亟需理论创新回应技术发展需求。

现有法律理论在自动驾驶场景中暴露出明显局限。严格责任理论主张将责任直接归属于生产者，虽强化风险防控却可能抑制技术创新——在 L3 级人机共驾场景中，若系统未及时预警而单纯归咎于生产者，显然忽视了驾驶员的接管义务，美国优步测试事故中安全员被判过失杀人而车企仅负民事责任的判决，便凸显了这一理论的适用矛盾。过错责任理论聚焦主观过错认定，但自动驾驶系统的“算法黑箱”特性使过错举证陷入困境，特斯拉 Autopilot 乡村路口事故中，算法决策逻辑的不可解释性导致司法机关难以追溯制造商是否存在过失，传统过错标准在此类系统自主决策场景中逐渐失效。身份预设责任模式则以固定身份划分责任，默认生产者为主责方、使用者为次责方，这种静态划分忽视了技术层级差异带来的控制力变化：L2 级系统中用户需全程监控，而 L5 级系统中用户几乎无操控权，《江苏省道路交通安全条例》对 L3 级事故责任的一刀切规定，及【(2023)鄂 01 民终 25835 号】案例中因系统数据加密导致责任判定受阻的困境，均反映出静态模式与技术特性的脱节。

针对这些局限，本文提出的“动态归责理论”构建了全新的责任认定框架。其核心主张是摒弃身份预设的固定模式，以“事故场景中各主体的实际控制力”为核心变量，结合技术层级与注意义务履行程度动态调整责任权重，核心逻辑可凝练为“控制力越强，注意义务越高，对应责任权重越大”。这一逻辑包含三层递进内涵：控制力是责任划分的基础锚点，并非单纯物理操控权，而是对驾驶行为或系统运行的实质支配与风险防控能力的综合判定；注意义务与控制力需对称性匹配，强度随控制力大小成正比分布，这种匹配避免了权责失衡，如同 L5 级系统中用户注意义务仅限于不破坏系统；责任权重随控制力与义务履行状态动态调整，控制力强却未履行对应义务的主体责任权重显著提升，已尽义务者则相应降低，实现“控制力 - 义务 - 责任”的精准衔接。

相较于现有理论，动态归责理论实现了多维度突破。不同于严格责任的结果导向，它以实际控制力为核心，避免责任与技术支配状态脱节，在 L3 级系统未预警事故中，若制造商已履行测试义务则责任权重可降低，平衡风险与创新；区别于过错责任的主观要件中心主义，它通过“控制力 - 注意义务”的对

称匹配将抽象过错转化为客观标准,传感器失效事故中可直接评估供应商是否合规,破解举证难题;相比身份预设的静态划分,它结合技术层级与场景动态调整责任,与《深圳经济特区智能网联汽车管理条例》的算法透明化要求相契合,通过“缺陷主动报告”制度破解算法黑箱困境。

这一理论的实践价值与社会意义显著。在理论层面,它通过实质化控制力界定为多元主体责任划分提供可量化基础,以阶梯式逻辑破解权责失衡,在保障公共安全与技术创新间实现平衡,与科技部《驾驶自动化技术研发伦理指引》中责任随技术层级动态变化的要求形成呼应。在司法实践中,它依托“技术鉴定委员会”机制与典型案例指导提升裁判统一性,最高检发布的技术鉴定标准可作为控制力评估依据,美国佛罗里达州特斯拉诉讼案中陪审团按“设计缺陷”比例定责的逻辑,正与该理论的风险/效用失衡判定思路一致。在社会层面,它通过动态校准责任边界倒逼各主体履行义务,车企需加大极端场景测试投入,用户需提升接管敏感度,同时为L5级研发预留合理风险空间,避免责任泛化抑制创新,为构建适应智能社会的法律责任体系提供了关键理论支撑。

2. 自动驾驶肇事刑事责任的认定困境

自动驾驶技术的迭代重构了传统交通行为的责任逻辑,其肇事刑事责任认定面临多重挑战。从司法实践与学理研究来看,这一领域的核心困境集中于主体多元化、因果关系复杂化、过失标准模糊化等层面,需结合技术分级特性、注意义务动态划分及归责理论革新,构建适配技术发展的解决方案。

2.1. 主体认定困境

2.1.1. 责任主体的多元化与边界模糊

传统交通肇事责任主体多聚焦于驾驶人,但自动驾驶场景中,责任链条延伸至制造商、算法开发者、道路测试方、第三方服务提供商等多元主体,从单一主体到多元网络的责任弥散。根据工业和信息化部、公安部、交通运输部《智能网联汽车道路测试与示范应用管理规范(试行)》(以下简称管理规范)第三十四条,交通事故责任需依据当事人的违法行为等判定,但该规则未明确主体是直接操作的驾驶员,还是对系统安全负总责的制造商,抑或是提供算法支持的技术公司。司法实践已显现这一困境。【(2021)湘0725民初1299号】案例中,事故车辆涉及道路测试委托方、技术服务方及驾驶员三方主体,法院最终认为需结合三方合同约定与行为关联性综合认定责任比例。学理上更明确指出,自动驾驶系统因缺乏独立意志与刑事责任能力,无法成为单独责任主体,其决策失误的责任需追溯至背后控制力主体,即对系统设计、运行或监控具有实质影响力的生产者、使用者或管理者这种“主体弥散性”使得传统“谁行为,谁担责”的归责逻辑难以直接适用[1]。

2.1.2. 自动驾驶级别的差异对主体责任划分的影响

自动驾驶技术的分级特性进一步加剧了主体责任的认定难度。低级别自动驾驶,L0-L2以驾驶员主导为核心,系统仅提供辅助功能,车道保持、自适应巡航等,要求驾驶员全程保持监控与接管能力;而高级别自动驾驶L4-L5中,系统可在特定场景下完全主导决策,驾驶员的操作义务显著弱化。这种差异直接影响责任重心的分配。研究表明,低级别事故的责任多指向驾驶员的未及时纠正系统的辅助失误监督过失,而高级别事故若因系统故障引发,制造商往往成为责任核心[2]。但司法实践中其主体依旧存在争议,【(2023)鄂01民终25835号】案例中,事故车辆处于L3级自动驾驶模式,系统在复杂路口未识别行人导致碰撞,但因系统数据加密、算法黑箱特性,法院难以解析“是驾驶员未接管,还是系统未预警”,最终因证据不足导致责任判定受阻。这一案例凸显:技术黑箱不仅模糊了“行为与结果”的关联,更直接阻碍了“主体与责任”的对应,使得不同级别下的责任划分缺乏明确锚点。

2.2. 因果关系与过失认定困境

2.2.1. 因果关系链条的多因叠加与认定障碍

传统交通肇事的因果关系多为“一因一果”，而自动驾驶事故常呈现“多因一果”的复合形态，涉及系统算法缺陷、驾驶员操作失误、外部环境干扰(如暴雨导致传感器失效)、道路设施不匹配等多重因素。这种叠加效应使得因果关系的追溯与切割异常困难。程凡卿在《自动驾驶刑事风险研究——刑事追责的困境与对策》列举的典型案例直观体现了这一困境：某事故中，车道保持辅助系统突然失效(系统因素)，驾驶员因过度依赖系统未及时修正方向(人为因素)，同时对向车道车辆超速(外部因素)，三者共同导致碰撞。法院在审理中难以判定“哪个因素是事故的决定性原因”，最终只能依据各方过错比例酌情划分责任[3]。上述案例表明当下自动驾驶的技术复杂性已突破传统因果关系理论的解释范围，“条件说”“相当因果关系说”难以应对多因素交织的场景。

2.2.2. 过失标准的动态模糊与注意义务的冲突

过失认定的核心是“注意义务的违反”，但自动驾驶场景中，注意义务的主体、内容与边界均随技术介入程度动态变化，导致标准模糊化。人机共驾模式下需建立“动态注意义务分工”：系统主导决策时，驾驶员的注意义务应限于监控系统异常并在必要时接管，而非传统的全程主动驾驶；若系统发出接管提示，驾驶员的反应时间(如3秒内)应作为义务履行的合理限度。但这一标准在实践中难以落地——不同驾驶员的反应能力存在差异，不同场景(如城市道路与高速公路)对反应速度的要求也不同，导致“合理注意义务”的认定缺乏统一尺度[4]。同时，算法黑箱加剧了制造商过失的认定难度，由于算法决策过程不透明，司法机关难以判断制造商是否预见并避免了系统缺陷。

2.3. 产品缺陷归责的刑事法律适用空白

民事领域中，《最高人民法院关于审理道路交通事故损害赔偿案件适用法律若干问题的解释》第九条明确：因机动车产品存在缺陷导致交通事故，生产者需承担赔偿责任。但在刑事领域，现行刑法缺乏针对“自动驾驶系统缺陷”的直接罪名，导致制造商的刑事责任难以追究。实践中若因系统缺陷引发重大事故，只能通过“重大责任事故罪”“过失致人死亡罪”等传统罪名追责，但这些罪名的构成要件与产品缺陷导致的事故存在明显不匹配——例如，“重大责任事故罪”要求行为人“在生产、作业中违反安全管理规定”，但算法缺陷是否属于“安全管理规定的违反”，司法机关对此存在显著分歧。【(2021)湘1202民初7684号】舒某贵诉怀化春某汽车销售服务有限公司案例中，进一步暴露了这一矛盾。该案中，法院认定辅助驾驶系统未达到消费者“在常规路况下安全行驶”的合理期待，构成产品缺陷(民事认定)，但因刑法中无对应罪名，无法对制造商追究刑事责任。这种“民事可追责而刑事不可追责”的断层，使得产品缺陷的刑事威慑力严重不足，难以倒逼制造商履行更高标准的安全义务。此外，技术标准的缺失加剧了这一困境：目前自动驾驶系统的缺陷认定缺乏统一行业标准，司法机关只能依赖消费者合理期待或风险收益平衡原则酌情判断，导致同类案件的裁判结果差异较大[5]。

3. “动态归责理论”的提出与核心内涵

针对自动驾驶肇事刑事责任认定中主体多元化、因果关系复杂化及责任边界模糊化等困境，本文提出“动态归责理论”，旨在突破传统静态归责框架的局限，实现刑事责任的精准化与适配性划分。

3.1. 理论核心内涵的深化阐释

“动态归责理论”的核心主张是：自动驾驶肇事的刑事责任认定应摒弃身份预设责任的固定模式，不是生产者必然主责、使用者一律担责，转而以“事故场景中各主体的实际控制力”为核心变量，结合

自动驾驶技术层级与注意义务履行程度，动态调整责任权重。其核心逻辑可凝练为“控制力越强，注意义务越高，对应责任权重越大”，具体包含三层递进内涵。

控制力是责任划分的基础锚点。控制力并非单纯指物理操控权，而是对驾驶行为或系统运行的实质支配能力与风险防控能力的综合判定。生产者的控制力体现在对车辆的物理控制、系统提示的及时反馈、接管请求的快速处理，以及对系统功能的合理运用；传感器供应商、道路管理者等第三方的控制力，则体现在组件质量保障与环境维护效能上，比如激光雷达测距精度的稳定性、道路标线的清晰程度等。这种对控制力的实质化界定，为责任划分提供了可量化的事实基础。

注意义务与控制力要对称性匹配。即注意义务强度应与主体控制力的大小成正比，高控制力主体，如 L4 级车企需履行全场景风险预见与防控义务，包括符合 ISO 21448 预期功能安全标准、对极端场景进行测试验证、建立系统缺陷快速响应机制；中控制力主体，如 L3 级用户需履行监督与及时接管义务，包括保持对路况的持续关注、在系统发出接管提示后 3 秒内响应、不因分心丧失接管能力；低控制力主体，如 L2 级第三方供应商需履行组件质量合规义务，如确保传感器在宣称的环境条件下正常工作。这种匹配性避免了控制力弱却承担重义务的权责失衡，例如 L5 级系统中，用户因控制力极弱，其注意义务仅限于不故意破坏系统，而非全程监控[6]。

责任权重随控制力与义务履行状态动态调整。责任权重的分配并非固定不变，而是基于控制力大小与注意义务履行程度综合判定。控制力越强的主体，若未充分履行对应强度的注意义务，其责任权重将显著提升；若主体已尽到与自身控制力匹配的注意义务，责任权重则会相应降低。这种动态调整机制让责任认定摆脱了身份预设的局限，实现了“控制力-义务-责任”的精准衔接。

3.2. 理论适用的具体路径与挑战应对

将“动态归责理论”落地于司法实践，需构建标准化的适用流程，同时直面技术特性带来的现实制约。其核心路径需围绕技术事实认定、责任权重量化、争议解决机制展开，并针对性回应“算法黑箱”“数据安全”“边缘场景”等技术挑战。

实践中，技术特性带来的制约需通过配套制度破解。“算法黑箱”导致的决策过程不透明是核心障碍，需建立独立强制性数据记录标准，要求自动驾驶车辆安装符合国家标准“黑匣子”，强制记录算法输入数据、决策逻辑节点、系统状态变化等关键信息，且数据存储格式需满足司法机关直接调取与核验需求，避免车企以技术保密为由隐瞒核心证据。针对算法安全性的追溯难题，应引入第三方安全审计制度，要求车企在产品上市前提交核心算法(脱敏后)至指定第三方机构进行安全测试，审计内容需覆盖极端天气、突发障碍物等边缘场景的决策逻辑，审计结果作为责任认定中“注意义务履行程度”的重要依据。对于暴雨、隧道出入口等“边缘场景”的责任模糊问题，需建立场景测试认证机制，由行业主管部门联合科研机构制定边缘场景测试清单，车企需证明其系统对清单内场景的覆盖能力，未通过认证场景发生的事故，可直接推定生产者未履行风险防控义务。

为保障理论落地的统一性与权威性，还需完善支撑机制。应设立跨领域的“自动驾驶事故技术鉴定委员会”，吸纳法律、计算机、交通工程等领域专家，统一把控控制力评估、算法缺陷认定等专业问题，其出具的鉴定意见可作为司法裁判的重要参考。同时需强化算法透明化与数据合规的平衡，借鉴《深圳经济特区智能网联汽车管理条例》的“缺陷主动报告”制度，要求车企在事故后及时提交脱敏算法框架与测试记录，既保障技术秘密又满足责任追溯需求。

3.3. “动态归责理论”的价值

其一，理论破解传统归责的静态性局限。通过构建动态关联机制，该理论有效消解了多主体责任划分的模糊性，避免了传统框架下一刀切式的责任分配。其二，强化责任预防功能。该理论通过明确“控

制力与义务对等”的逻辑，倒逼各主体履行与其控制力匹配的注意义务：车企需加大极端场景测试投入以覆盖潜在风险，用户需提升对系统提示的敏感度以保障及时接管，从而从源头降低事故发生概率。其三，平衡技术发展与安全保障。一方面，理论为技术探索预留合理空间，例如 L5 级自动驾驶研发中的合理风险可适当降低责任权重，避免因责任泛化抑制创新活力；另一方面，通过动态校准责任边界，防止相关主体以技术复杂为由逃避应承担的责任，为自动驾驶产业的健康发展提供法治保障。

4. 结论

随着自动驾驶技术的发展，其交通肇事的刑事责任认定成为重要法律问题。主体多元化、因果关系复杂化、法律适配性不足等困境，使得依赖“单一主体归责”或“静态责任划分”的传统模式难以应对技术革新带来的复杂场景。

本文通过剖析自动驾驶肇事刑事责任认定的核心难题，提出“动态归责理论”作为解决方案。该理论以“控制力 - 注意义务 - 责任权重”的联动关系为核心，主张摒弃固定化责任分配模式，以事故场景中各主体的实际控制力为核心变量，结合自动驾驶技术层级与注意义务履行程度动态调整责任权重。其核心逻辑“控制力越强，注意义务越高，责任权重越大”，既破解了传统归责的静态性局限，清晰区分不同场景下的责任差异，又通过强化各主体的义务履行，从源头降低事故概率，更在保障技术创新与维护公共安全之间实现了平衡。未来，随着自动驾驶技术的进一步发展，相关的刑事责任认定规则还需不断完善和细化，以更好地适应社会发展的需求。

基金项目

本论文由辽宁科技大学大学生创新创业训练计划项目经费支持，编号：X202510146274。

参考文献

- [1] 廖兴存. 无人驾驶汽车交通肇事过失刑事责任论纲[J]. 湖南社会科学, 2021(3): 122-130.
- [2] 蔡仙. 自动驾驶中过失犯归责体系的展开[J]. 比较法研究, 2023(4): 65-81.
- [3] 程凡卿. 自动驾驶刑事风险研究——刑事追责的困境与对策[J]. 比较法研究, 2022(5): 118-130.
- [4] 袁彬, 徐永伟. “人机共驾”模式下交通肇事罪的适用困境及突围[J]. 广西大学学报(哲学社会科学版), 2019, 41(5): 76-84.
- [5] 冯文杰. 人工智能刑法研究的法益保护观及其适用[J]. 理论月刊, 2025(6): 129-141+160.
- [6] 彭文华. 自动驾驶车辆犯罪的注意义务[J]. 政治与法律, 2018(5): 86-99.