The Ethical Implications of Artificial Intelligence for Road Safety

Bahare Khajehpour[a]*, Kritika Maheshwari[a], Neelke Doorn[a], Niek Prins[b], Jan-Pieter Paardekooper[c,d], Oscar Oviedo-Trespalacios[a]

* Corresponding author, email address: b.khajehpour@tudelft.nl

[a] Delft University of Technology, Faculty of Technology, Policy, Management, Delft, The Netherlands
[b] Haskoning, Advisory Group Sustainable Mobility, Amersfoort, The Netherlands
[c] Nederlandse Organisatie voor toegepast-natuurwetenschappelijk onderzoek (TNO), Integrated Vehicle Safety group, Helmond, The Netherlands
[d] Radboud University, Donders Institute for Brain, Cognition and Behaviour, Nijmegen, The Netherlands

Keywords: road safety, artificial intelligence, intelligent transport systems, AI ethics, transport policy, road traffic crash, safe system approach

Long Abstract

Each year, road traffic crashes (RTCs) cause approximately 1.2 million fatalities and an estimated 50 million injuries worldwide (WHO, 2023a, 2023b), making them ranked among the top three leading causes of death for individuals aged 5 to 50 for the past quarter century (IHME, n.d.). Apart from the immeasurable emotional toll caused by the immediate loss of lives and physical harms, the associated financial costs of RTCs (incl. lost income, medical expenses, and judicial proceedings) are estimated to amount to 12% of the global Gross Domestic Product (Chen et al., 2019). Given this extensive impact, preventing RTCs has been a central concern in safety science (Safarpour et al., 2020) and has driven significant technological innovations.

Technology has long played a key role in improving road safety. Conventional safety technologies such as helmets, seat belts, and anti-lock braking systems are shown to be integral in preventing RTCs and reducing injury severities (European Commission, 2020). Recent advancements in AI have created new opportunities for enhancing road safety (Ehsani et al., 2023), with applications for road user assistance and vehicle safety to road infrastructure planning and traffic management (Eskandari Torbaghan et al., 2022; Sohail et al., 2023). A notable example is advanced driver assistance systems (ADASs), such as lane-departure warning or driver distraction detection, that help mitigate human errors in

performing driving tasks. On the infrastructure level, AI-driven risk mapping tools and road surface monitoring technologies also inform road authorities to make better targeted safety interventions. Similarly, intelligent traffic monitoring cameras and traffic flow prediction models support traffic management authorities in preventing RTCs and reducing their harmful impact.

With such a wide range of applications, the full integration of AI into transport systems holds considerable promise for improving road safety. However, as with other emerging technologies, opportunities often come with risks and unintended consequences, such as harmful environmental impacts, potential human rights violations, or adverse socio-economic effects. Therefore, the extent to which these tools are societally beneficial depends on their ability not only to fulfill their core purpose (here, road safety) but also to reflect on and address the broader challenges they may introduce. Since AI has been acknowledged as a technology with high-risk applications for the transport domain (*EU Artificial Intelligence Act*, 2024), it is essential to critically examine the ethical implications of its use for road safety.

Research in AI ethics that has engaged with road safety technologies has largely kept its focus on CAVs (Lin, 2016; Nyholm, 2022) and their crashing-related ethical issues (Lundgren, 2021) even though their widespread adoption is not expected before 2035 due to technical, regulatory, and financial barriers (World Economic Forum, 2025). Road safety research, on the other hand, has acknowledged some ethical concerns common to AI, such as bias (Xu et al., 2024), misuse (Perrine et al., 2019), and unintended impact (Abebe et al., 2024), with regard to a broader range of AI-driven safety technologies (such as road traffic cameras, traffic signals, and safety investment planning tools) but has marginally explored associated normative concerns and governance challenges that shape their real-world deployment and effectiveness (Wandelt et al., 2025).

This study contributes to understanding these challenges by presenting a task-oriented categorization of AI applications for road safety, namely future forecasting, correlation analysis, detection, group formation, and optimization, and by examining how these functions give rise to ethical concerns related to responsibility, explainability, human autonomy, justice, and privacy. It argues that these concerns are directly linked to the regulatory design, public legitimacy, and safety performance of intelligent transport systems (ITS).

The ethical concerns are not incidental but emerge from specific characteristics of the road safety domain, combined with some inherent technical features of AI. These include the pervasive presence of safety technologies in public space, their use by ordinary users without receiving adequate training (Murtaza et al., 2023), their gradual hidden influence on

human action and decision-making (Nallur et al., 2025), their dependence on collective action, their operation in changing environments with multiple influential factors, their involvement in morally sensitive decisions (Braun et al., 2021; Rebentisch et al., 2019), and their promise to serve a societally shared goal. Some of the influential technical features of AI are also extensive data collection and inference capabilities (Ziakopoulos and Yannis, 2024), opaque and complex internal workings (Kuznietsov et al., 2024), high accuracy in identifying correlations (Mosier and Skitka, 2018), reliance on complex actor networks (Santoni De Sio and Mecacci, 2021), and limitations in training data (Suresh and Guttag, 2021). The above-mentioned characteristics together can activate some unexpected dynamics that undermine the effectiveness of AI tools for road safety.

AI-driven tools do more than change road user behavior or assist policy implementation; rather, they actively participate in defining what counts as risk, whose safety matters most, how trade-offs are resolved, and how obligations are assigned. Group formation and detection systems, used in driver assistance and profiling, classify driver behavior into risk categories, promoting new safe driving norms. Correlation analysis and forecasting systems, used for identifying crash hotspots or planning infrastructure maintenance, privilege certain road users, locations, and forms of harm, thereby shaping how limited resources are allocated. Optimization systems, used for safety investment planning, weigh competing objectives to achieve a desired balance. Decision-support and behavior-change technologies redistribute responsibility across road users, system designers, technology providers, and public authorities, sometimes creating ambiguity.

These dynamics are not merely technical concerns or detached ethical considerations but governance issues, as they challenge current legal frameworks, particularly regarding liability and accountability, take part in shaping public trust and policy legitimacy, and affect system performance. Utilizing AI for road safety decisions affects how they can be justified, contested, and held accountable in practice. While current legal frameworks need to evolve to address these complexities, this concern is mainly acknowledged in policy efforts pertinent to CAVs (Li et al., 2019) and overlooked with regard to other intelligent road safety systems. Moreover, people are more supportive of safety policies that are derived from a transparent decision-making process and perceived as legitimate and fair across different road users (Martínez-Buelvas et al., 2025; Ombagi et al., 2023). The safety gains of AI-driven tools, however, can be distributed unequally among different population groups, and they can be used for illegitimate purposes. Therefore, public trust and policy legitimacy are at risk. Lastly, the unforeseen behavioral and systematic responses to AI-driven tools can risk safety, with negative implications for their adoption and use.

This study contributes to transport policy discourse by reframing the relevant ethical concerns of ITS as the issues that transport policy should address, instead of regarding them as merely subsidiary challenges or extra improvement opportunities for otherwise sound systems that are successful in delivering their main objective, here safety. From a policy and practice standpoint, governing AI for road safety requires moving beyond technical validation and harm reduction metrics toward institutional arrangements that make value assumptions explicit, support meaningful human control, raise public awareness and digital literacy, allow contestation and co-design, and facilitate cross-sector collaboration and reflexive learning over time. This insight suggests that leveraging AI to enhance road safety should not be treated as a purely technical endeavor left to program developers and road safety professionals; rather, it requires sustained engagement with ethical and governance expertise.

Funding: This work was supported by the European Union's Horizon Europe research and innovation program under the Marie Skłodowska-Curie Actions Doctoral Networks, through the project IVORY – AI for Vision Zero in Road Safety [Grant Agreement No. 101119590].

References

Abebe, H.G., Belin, M.-Å., Björnberg, K.E., 2024. Equity and Social Justice considerations in road safety work: The case of Vision Zero in New York City. Transp. Policy 149, 11–20.

Braun, L.M., Le, H.T.K., Voulgaris, C.T., Nethery, R.C., 2021. Healthy for whom? Equity in the spatial distribution of cycling risks in Los Angeles, CA. J. Transp. Health 23, 101227. https://doi.org/10.1016/j.jth.2021.101227

Chen, S., Kuhn, M., Prettner, K., Bloom, D.E., 2019. The global macroeconomic burden of road injuries: estimates and projections for 166 countries. Lancet Planet. Health 3, e390–e398. https://doi.org/10.1016/S2542-5196(19)30170-6

Ehsani, J.P., Michael, J.P., MacKENZIE, E.J., 2023. The future of road safety: challenges and opportunities. Milbank Q. 101, 613.

Eskandari Torbaghan, M., Sasidharan, M., Reardon, L., Muchanga-Hvelplund, L.C.W., 2022. Understanding the potential of emerging digital technologies for improving road safety. Accid. Anal. Prev. 166, 106543. https://doi.org/10.1016/j.aap.2021.106543

EU Artificial Intelligence Act, 2024.

European Commission, 2020. Next steps towards 'Vision Zero': EU road safety policy framework 2021 2030. Publications Office of the European Union.

IHME, n.d. GBD Compare [WWW Document]. URL http://vizhub.healthdata.org/gbd-compare (accessed 12.11.24).

Kuznietsov, A., Gyevnar, B., Wang, C., Peters, S., Albrecht, S.V., 2024. Explainable AI for safe and trustworthy autonomous driving: A systematic review. IEEE Trans. Intell. Transp. Syst.

Li, S., Sui, P.-C., Xiao, J., Chahine, R., 2019. Policy formulation for highly automated vehicles: Emerging importance, research frontiers and insights. Transp. Res. Part Policy Pract. 124, 573–586. https://doi.org/10.1016/j.tra.2018.05.010

Lin, P., 2016. Why ethics matters for autonomous cars. Auton. Driv. Tech. Leg. Soc. Asp. 69–85.

Lundgren, B., 2021. Safety requirements vs. crashing ethically: what matters most for policies on autonomous vehicles. AI Soc. 36, 405–415. https://doi.org/10.1007/s00146-020-00964-6

Martínez-Buelvas, L., Rakotonirainy, A., Grant-Smith, D., Haque, M.M., Oviedo-Trespalacios, O., 2025. Public perceptions and support for policies enhancing safety and justice in connected and automated vehicles and vulnerable road users interactions. Case Stud. Transp. Policy 22, 101571. https://doi.org/10.1016/j.cstp.2025.101571

Mosier, K.L., Skitka, L.J., 2018. Human decision makers and automated decision aids: Made for each other?, in: Automation and Human Performance. CRC Press, pp. 201–220.

Murtaza, M., Cheng, C.-T., Fard, M., Zeleznikow, J., 2023. Preparing drivers for the future: Evaluating the effects of training on drivers' performance in an autonomous vehicle landscape. Transp. Res. Part F Traffic Psychol. Behav. 98, 280–296. https://doi.org/10.1016/j.trf.2023.09.013

Nallur, V., Renaud, K., Gudkov, A., 2025. Nudging Using Autonomous Agents: Risks and Ethical Considerations, in: Collier, R., Ricci, A., Nallur, V., Burattini, S., Omicini, A. (Eds.), Multi-Agent Systems, Lecture Notes in Computer Science. Springer Nature Switzerland, Cham, pp. 283–296. https://doi.org/10.1007/978-3-031-93930-3_17

Nyholm, S., 2022. The Ethics of Transitioning Toward a Driverless Future. Test-Driv. Future Auton. Veh. Ethics Technol. Change 59.

Ombagi, C., Minja, D., Muna, W., 2023. Effect of Policy Network Legitimacy on Policy Process Outcomes in the Road Transport Sector in Nairobi City County, Kenya.

Perrine, K.A., Levin, M.W., Yahia, C.N., Duell, M., Boyles, S.D., 2019. Implications of traffic signal cybersecurity on potential deliberate traffic disruptions. Transp. Res. Part Policy Pract. 120, 58–70. https://doi.org/10.1016/j.tra.2018.12.009

Rebentisch, H., Wasfi, R., Piatkowski, D.P., Manaugh, K., 2019. Safe Streets for All? Analyzing Infrastructural Response to Pedestrian and Cyclist Crashes in New York City, 2009–2018. Transp. Res. Rec. 2673, 672–685. https://doi.org/10.1177/0361198118821672

Safarpour, H., Khorasani-Zavareh, D., Mohammadi, R., 2020. The common road safety approaches: A scoping review and thematic analysis. Chin. J. Traumatol. 23, 113–121. https://doi.org/10.1016/j.cjtee.2020.02.005

Santoni De Sio, F., Mecacci, G., 2021. Four Responsibility Gaps with Artificial Intelligence: Why they Matter and How to Address them. Philos. Technol. 34, 1057–1084. https://doi.org/10.1007/s13347-021-00450-x

Sohail, A., Cheema, M.A., Ali, M.E., Toosi, A.N., Rakha, H.A., 2023. Data-driven approaches for road safety: A comprehensive systematic literature review. Saf. Sci. 158, 105949. https://doi.org/10.1016/j.ssci.2022.105949

Suresh, H., Guttag, J., 2021. A Framework for Understanding Sources of Harm throughout the Machine Learning Life Cycle, in: Equity and Access in Algorithms, Mechanisms, and Optimization. Presented at the EAAMO '21: Equity and Access in Algorithms, Mechanisms, and Optimization, ACM, -- NY USA, pp. 1–9. https://doi.org/10.1145/3465416.3483305

Wandelt, S., Zhang, A., Antoniou, C., Budd, L., Castro-Nuno, M., Ding, Y., Eliasson, J., Goessling, S., He, S.Y., Ison, S., Jiang, C., Lam, J.S.L., Li, K., Liu, Z., Loo, B.P.Y., Ma, X., Ortúzar, J. de D., Oviedo-Trespalacios, O., Wang, D., Sun, X., 2025. The Hitchhiker's Guide to challenges in transport policy research: Towards ANSWERing questions regarding life, mobility, and everything. Transp. Res. Part Policy Pract. 200, 104650. https://doi.org/10.1016/j.tra.2025.104650

WHO, 2023a. Road traffic injuries [WWW Document]. URL https://www.who.int/news-room/fact-sheets/detail/road-traffic-injuries (accessed 1.20.25).

WHO, 2023b. Global status report on road safety 2023. Geneva: World Health Organization.

World Economic Forum, 2025. Autonomous Vehicles: Timeline and Roadmap Ahead.

Xu, W., Smart, M., Tilahun, N., Askari, S., Dennis, Z., Li, H., Levinson, D., 2024. The racial composition of road users, traffic citations, and police stops. Proc. Natl. Acad. Sci. 121, e2402547121. https://doi.org/10.1073/pnas.2402547121

Ziakopoulos, A., Yannis, G., 2024. Key Artificial Intelligence and Digitalization Solutions Towards Vision Zero in Road Safety, in: Sładkowski, A. (Ed.), Using Artificial Intelligence to Solve Transportation Problems, Studies in Systems, Decision and Control. Springer Nature Switzerland, Cham, pp. 1–26. https://doi.org/10.1007/978-3-031-69487-5_1