Available online at www.sciencedirect.com

ScienceDirect

Transportation Research Procedia 72 (2023) 2449–2456

Transportation Research Procedia

www.elsevier.com/locate/procedia

Transport Research Arena (TRA) Conference

Ethical and legal challenges of automated driving: The prioritization of socio-political values

Carolina Diaz-Piedra[a1], Belen Liedo[b], Miguel Núñez de Prado-Gordillo[a], Mª Jesus Caurcel[a], Leandro L. Di Stasi[a]

[1] *Mind, Brain, and Behavior Research Center (CIMCYC), University of Granada, Granada, Spain, 2 Institute of Philosophy, Spanish National Research Council (CSIC), Madrid, Spain*

Abstract

Much research on the ethics of automated driving (AD) focusses on moral decision-making processes in extreme traffic situations. However, more likely scenarios involving daily trade-offs among different values in the design and implementation of AD are also worth considering. Our main goal was to provide an in-depth analysis of how diverse experts and stakeholders prioritize different socio-political values. Through a series of workshops, semi-structured interviews, and focus groups, we assessed today's major social, ethical, and legal concerns related to AD, as well as their relative prioritization in specific near-future realistic traffic scenarios. We summarize experts' and lay-people's opinions on the trade-offs among six socio-political values (privacy, autonomy, safety, security, performance, and costs) and present overall recommendations for the design of future AD technologies.

Keywords: Automated driving; Ethical and political values; Human-Machine Integration; Traffic psychology; Algorithmic Ethics.

1. Introduction

While the literature on the ethical aspects of automated driving (AD) continues to grow, much of it focuses on extreme traffic situations where the vehicle "makes" a moral decision. Typically, in such hypothetical scenarios, automated vehicles (AVs) face a trolley-problem-like situation and are forced to choose among several more or less grievous outcomes (Rodríguez-Alcázar et al. 2021). However, an overemphasis on this kind of studies might lead to neglect important and more likely and pressing socio-political, ethical, and legal issues related to the use and design of AD technologies. To begin with, trolley-problem-based studies tend to focus on highly or fully AVs (e.g., SAE

[1]* Corresponding author. Tel.: +34-958-24-6276.
E-mail address: dipie@ugr.es

10.1016/j.trpro.2023.11.745

Level 5, see SAE International 2021), while more likely scenarios will involve vehicles with lower or variable automation degrees (Mörtl 2020). Furthermore, discussions between competing moral theories are their primary focus of interest. Despite the relevance of such theoretical discussions, the kind of emergency situations depicted are expected to be infrequent, precisely, as a result of the likely widespread use of AVs. Other ethical, legal, and socio-political concerns related to AD (e.g., how personal data will be handled, how legal policies and insurances will be regulated, how the automotive industry and road infrastructures will change to accommodate new needs, how laws can make this technology safer for the general people, etc.) seem more urgent to address (Roff 2019). Finally, trolley-problem-like studies typically draw from what Mörtl (2020) refers to as "solipsistic" approaches to AVs design, where the ascertainment of the possible ethical and socio-political implications of AD would happen after AD technologies have been already designed. By contrast, more recent proposals stress the need to adopt a more holistic approach, which entails a proactive perspective and aims to incorporate the results of human factors research into the design process itself. The Holistic Approach for Driver Role Integration and Automation Allocation for European Mobility Needs (HADRIAN) project, which provides the framework for this investigation, is one example of such kind of holistic approach to the development of AVs, with a special focus on how specific socio-political values and needs should shape the design of AD systems (see Fig. 1.).

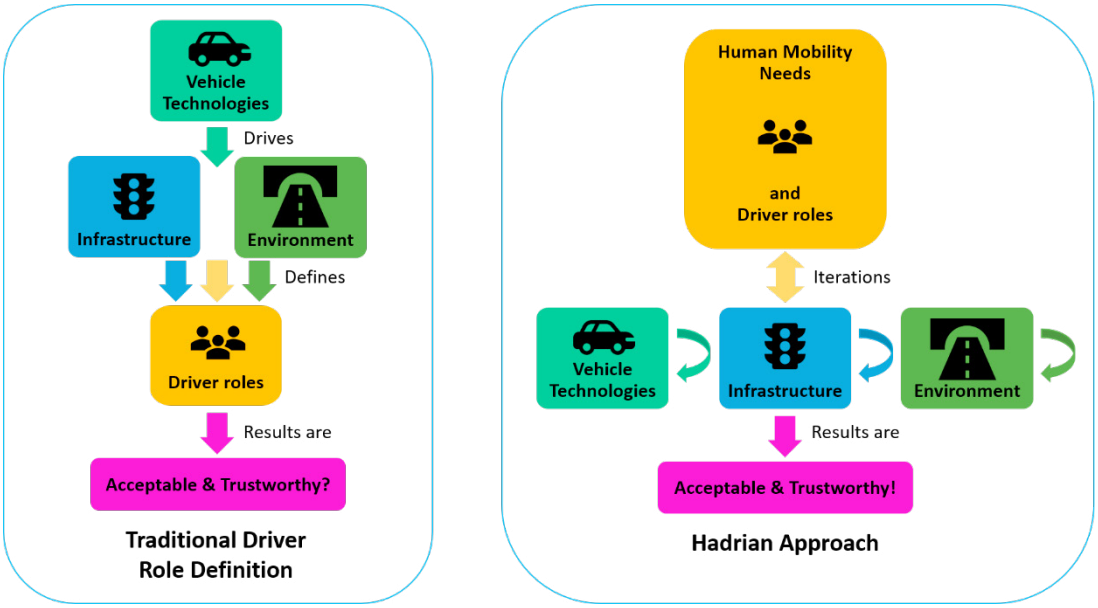

Fig. 1. Traditional approach (left side) vs. HADRIAN project's holistic approach (right side) to the relationships between automated driving technologies design and research on human factors and ethical, legal, and socio-political concerns. Adapted from HADRIAN's website (https://hadrianproject.eu/holistic-vision/).

This work constitutes a first approximation to the establishment of such socio-political values and needs. The shaping of the complex technologies involved in AD requires decisions on the value trade-offs that are inescapable in artificial intelligence (AI) assessment and design (Todt et al. 2010). Although most people would agree that values like users' autonomy, privacy, or safety are very important, considered in isolation, societies may prioritize these values to varying degrees. Hence, our main goal was to provide some insight on how relevant stakeholders, experts, and road users prioritize different socio-political and economic values that may conflict with each other in the design, development, and use of AVs. These insights can be taken into account, as high-level requirements, by lawmakers, carmakers, and software developers.

2. Method

We conducted an in-depth analysis of the main values involved in the design and use of AD-based technologies, as well as of the value prioritization schemes of different relevant stakeholders, experts, and road users. To do so, we employed a combination of descriptive qualitative and quantitative methods.

2.1. Participants and data collection methods

First, we conducted three workshops with a series of experts in ethics, political philosophy, psychology, human factors, and law to discuss and identify potential legal and ethical issues concerning data-flow privacy and security, as well as issues related to drivers' autonomy when using fluid interactions in the context of AD and the HADRIAN project application descriptions (Mörtl et al. 2020). Most experts (4 out of 6) were members of the HADRIAN Consortium and all of them worked at the University of Granada (Spain).

Then, we conducted a series of semi-structured interviews to collect the opinion of 30 qualified representatives of several stakeholders as well as experts in different fields (15 of them were related to the HADRIAN project). The participants came from 26, both public and private, European institutions from eight different countries (Austria, France, Germany, Greece, Italy, Slovenia, Spain, and Sweden). Participants were organized in groups depending on their field of expertise, which included Automotive industry (n = 9), Urban and transportation infrastructure (n = 5), Human factors (n = 3), Ethics (n = 4), Insurance sector (n = 3), Law (n = 4), and Traffic Medicine (n = 2).

Finally, we conducted five focus groups –i.e., facilitated discussions covering a set of specific issues– to gain insight into the trade-off preferences of lay people (n = 27). A cross-section of different road users took part in them, including elderly drivers (n = 6, 63-75 years old, two men), young drivers (n = 6, 21-26 years old, two men), health professionals (elders' caregivers, n = 5, 37-55 years old, all women), law enforcement agents (n = 5, 36-39 years old, all men), and managers of technology companies (n = 5, 27-50 years old, all men).

2.2. Procedures and data analysis

Throughout the initial workshops, relevant issues were first classified under six major categories: Autonomy, Safety, Security, Privacy, Performance, and Costs, all relevant for the AD context and design. The pertinence and definition of these six socio-political values were assessed by experts from inside and outside the HADRIAN Consortium with regard to their pertinence and their definitions. These six values represented the framework that guided the later debate with experts, stakeholders, and users.

The interviews collected the main concerns and points of view of both experts and relevant stakeholders about different ethical issues on the design and future implementation of AD-based technologies. An interview guide was set up beforehand and it represented the common core structure for each interview session, although it was slightly modified with specific questions to certain profiles of experts. All interviews were carried out through video-conference. Informed consent was obtained from each interviewee before proceeding (University of Granada´s Institutional Review Board approval 1528/CEIH/2020). Before starting the interviews, a series of videos presented both the HADRIAN project and its main target populations (i.e., elderly drivers, transport workers, and office workers who may use AVs as a mobile office, Harold, Sven, and Florence, respectively see Mörtl et al. 2020). These videos served as an introduction to their main characteristics and specific mobility needs. The interview recordings were video- and audio-edited to anonymize interviews. An automated transcription software (Happy Scribe Ltd. n.d.) converted the audio files to text. Then, two judges checked over the transcripts of each interview and edited the transcriptions when necessary. Responses to the same questions were organized by fields of expertise and themes.

The focus groups collected the trade-off preferences of different road users. Each group worked on a particular mobility scenario related to one of the HADRIAN project's main target populations (i.e., elderly drivers, transport workers, and office workers who may use AVs as a mobile office, see Mörtl et al. 2020). We did not look for the moral judgement of individuals concerning their hypothetical reactions to moral dilemmas in extreme AD scenarios, but rather for their prioritization of social, economic, and/or environmental values, taking into account the common

good. A topic guide was employed to guide discussions. Its content and structure was similar for each of the five groups, but slightly modified to address the specific needs of each HADRIAN application descriptions of AD. Focus groups were conducted in person and video-recorded. Informed consent was obtained from each participant before proceeding. After presenting the HADRIAN project, a video presentation of a particular mobility scenario was screened. In the video, participants saw the main information about the primary actor/actress of the hypothetical scenario (Harold, Sven, or Florence), their general characteristics, and main driving challenges. They had the opportunity to ask questions about the content if they needed any clarification. Then, 4 to 6 mobility scenario segments (i.e., specific driving situations faced by the main actor) were presented. Participants filled in the HADRIAN Values Index after each segment. This index consists of a multi-dimensional rating procedure that assesses the importance/ relevance of six independent socio-political criteria while designing an AV: autonomy of the user, safety, security, privacy, performance, and costs. Firstly, participants were asked to rate the importance of each socio-political value on a 0-to-100 response scale, from "not at all important" to "extremely important", in a particular segment. Then each, out of fifteen, possible pair-wise comparisons of the six values were presented to participants in a randomized order. They selected the value that, from their point of view, would be more relevant when designing a Human-Machine Interaction (HMI) system in each particular mobility scenario segment. Finally, a discussion was held. The recordings were video- and audio-edited to anonymize discussions. An automated transcription software (Happy Scribe Ltd. n.d.) first converted the audio files to text. Then, two judges checked over the transcript of the focus groups and edited the transcriptions when necessary. Their insights were used, when relevant, as a counterpoint to the experts' opinions. Also, we conducted an additional descriptive analysis of the relative frequency that a given value was selected over the others in the fifteen possible pair-wise comparisons. Data were normalized using Min-Max normalization within each category of participants.

3. Results and Discussion

The main themes and insights obtained from the three methodologies employed are presented in Table 1, organized by issues and concerns for each value.

Table 1. Six major socio-political values identified in the workshops and their related ethical concerns and issues regarding AD identified in the interviews and focus groups.

Privacy	Autonomy (of the user)	Safety	Security	Performance (of the vehicle)	Costs
Data minimisation	Responsible agents	HMI	Means for avoiding hacking	Explainability	Infrastructure readiness
Satisfactory consent	Safety of a situation	Activities allowed	Reaction in case of attack	Avoiding algorithm discriminatory bias	Accessibility
Data sovereignty	Awareness of a situation	Responsibility in emergencies	Satisfactory informed consent about security risks	Meaningful human control	Sponsoring and other economic interests
Legal coverage	Quality of the experience	Accountability and liability			Mobility as a service
Special data		Cultural differences			
Personal data security		System security			
Strategic data as public goods					

3.1. Privacy

AVs collect more or less 1 GB of data every second (Collingwood 2017) in order to perform driving. Since such data may not only refer to the driving environment, but also to the state/behaviour of different road users, individual rights concerning data management need to be protected. Both exhaustive and clear regulations and effective and reliable means of monitoring this data management are essential conditions on AV.

In this sense, the main issues and concerns referred to the need for a definition of (1) different kinds of data and (2) the means and tools for a correct and secure data management, including legal coverage and satisfactory consent. Firstly, it needs to be specified what "necessary data" amounts to, so HMI systems can be designed to meet the principle of data minimization, according to which all data collected have to be necessary and must be deleted when they are no longer needed. Other sensitive data (e.g., health-related data), as well as publicly relevant data (e.g., data which might help to improve overall road security) also need to be properly defined. Secondly, there was a consensus regarding the need for proper methods of acquiring the user's consent to the use of their data, in particular regarding the avoidance of asymmetric and cumbersome consent systems. Two related and relatively consensual requirements were data sovereignty - i.e., granting users' ability to always remain in control of how their data are being processed

- and minimal AI-literacy - i.e., promoting users' knowledge of how AI-driven data management systems work. The latter was deemed particularly relevant for elderly drivers. On the other hand, there were clear disparities as to whether privacy should be prioritized over other values (namely, security and safety). While some experts stated that privacy should be preserved at almost all costs, others viewed privacy as a minor issue. Lay-people's views on this matter seem to conform to the latter opinion (see Fig. 2). These discrepancies were particularly relevant for the case of transport workers, whose commuting data could be useful for enhancing road safety and determining liability in case of accident, but whose labour rights might be threatened by the possibility of constant surveillance. In any case, all experts considered that, once anonymised, data should be shared among all agents aimed to improve road safety.

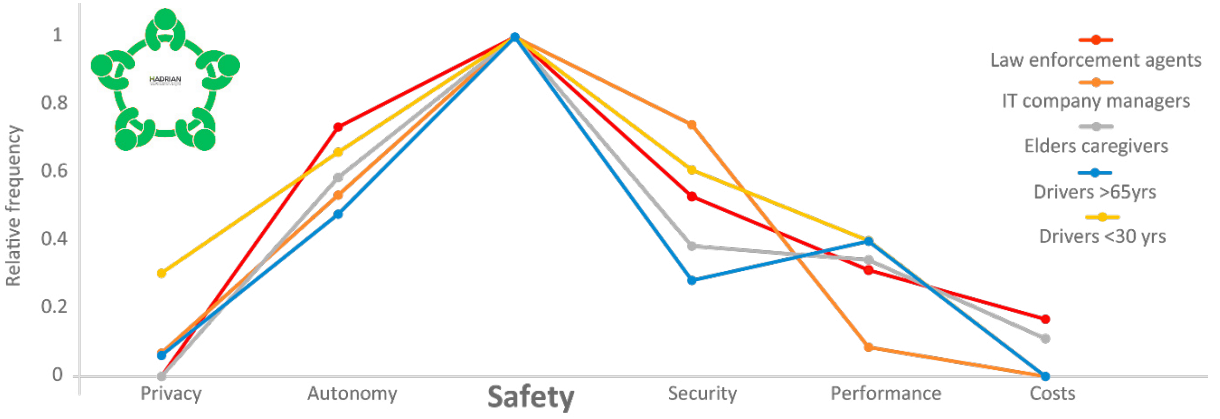

Fig. 2. Average percentage of times (relative frequency) that a value was selected over the others. Each socio-political colored line represents the results of a different focus group involving either law enforcement agents; IT company managers, elders' caregivers, elderly drivers, and young adult drivers.

3.2. Autonomy of the driver

In manual driving, the driver is always responsible for responding to different moral choices and their consequences. Usually, decisions made under in emergency situations under time pressure can be inadequate or unfair. In the case of AVs, decisions before emergencies may be pre-defined, hence more thoughtful and well-pondered, which could increase overall road safety. Nevertheless, it also raises some new questions that need to be addressed.

In this regard, discussions were primarily about (1) who are the agents who ought to take the decisions concerning the behaviour of AVs, (2) the need for drivers to understand the type of support that AVs can (and cannot) provide, (3) the customer's interests as a potential source of a successful commercialization, and (4) the inclusion of an in-vehicle recording system to make responsibilities more easily detectable (akin to the flight recorders used in aviation). Regarding the first question, the main concern was whether decision-making processes should be pre-defined in accordance with strict regulations or whether drivers should be granted the possibility to choose how the vehicle should react in a given situation among a pre-arranged range of ethical options. There was a general consensus about the need for specific AD regulations, guaranteed by the state, although some stressed that excessive regulations could have a negative impact on the quality of the driver's experience, which is related to the third concern. Regarding the second concern, there was a general agreement regarding the need for securing the drivers' understanding of their duties as drivers/users of AVs, particularly in the case of partially automated devices (e.g., SAE Levels 2-4, SAE International, 2021). Human factors experts deemed the issue of awareness particularly important, given that a decrease in awareness and the driver's ability to respond effectively to emerging safety hazards is to be expected in partial AD scenarios. In this regard, the implementation by design of specific warning functions would be important for responsibility and liability assessments. Relatedly, many advocated for the presence of in-vehicle "black box" systems to facilitate liability allocation. Finally, experts from different fields underlined the necessity of establishing a specific legal framework for liability litigations before deploying AVs on public roads.

3.3. Safety

Improving general road safety is one of the main arguments in favour of AVs (Jenkins 2016). Most accidents are caused by human error or negligence, which could in principle be avoided by AD. However, as one expert pointed out, there is no evidence for this claim so far, and safety is still one of the main concerns of potential users (Kyriakidis et al. 2015). AVs stakeholders thus need to achieve a trustworthy level of safety. In any case, it is expected that safety as a whole would increase if fully AVs are a commonality. In the meantime, AVs, in their different automation levels, will share the road with human drivers. Mixed driving raises some new safety risks that need to be foreseen (Goodall 2014). Above all, safety should be improved for all road users, not just AV drivers.

In this regard, safety was the value considered more important for different road users and all users tended to privilege safety over the other values (see Fig. 2). Discussions among experts were mainly about (1) vehicle reactions in the face of emergencies, (2) the need for a new liability system, or (3) how to incorporate different cultural safety climates. There was a strong agreement about the idea that to achieve a higher road safety, especially during mixed driving scenarios, future HMI systems would need to be clear and fast enough to handle any contingency, and AVs drivers/users would need detailed information regarding activities allowed, the control of the vehicle, etc., but also regarding their responsibility beyond ownership. In this regard, many experts highlighted the need for promoting an accurate understanding of the functioning of HMI systems, the reduction of mismatched expectations and the need for standardized HMI design or intensive training programmes on each specific HMI/AV. This issue was deemed particularly important for the case of elderly drivers, whose ability to understand the functioning of AVs might be compromised. Regarding who should be in charge of decision-making in high-risk scenarios, there was a strong disagreement; while some experts considered that it would be safer to trust AVs decision-making, others considered the driver as the most reliable agent. Some considered that, in the near future, where mixed driving scenarios will be more common and where AVs will still be learning, human drivers will be more reliable. On the other hand, AVs are expected to make safer choices in fully AD systems. In addition, the issue of adapting the functioning of AVs to each cultural safety climates was raised. A majority of experts supported the idea that, at least in mixed driving scenarios, AVs would need to adjust to the cultural context of different countries and regions.

3.4. Security

AD systems will involve high connectivity and data traffic (Klaassen & Szuprycinski 2019). AVs optimal performance will lay on this high connectivity, both vehicle to vehicle (V2V) and vehicle to infrastructure (V2I). Complexity and connection to the Internet are two of the main indicators of susceptibility to hacking attacks (van Roermund 2019). Nowadays, achieving a complete immunity from hacking attacks is impossible. Software providers, though, must minimize exposure to hacking and prepare the adequate response in case it happens. Security will not be guaranteed, but it should always be maximized as much as possible.

Regardless of their field of expertise, most of the interviewed experts agreed that security was one of the main concerns regarding AVs and tended to consider it as a priority for all stakeholders. Discussions mainly focused on (1) the availability of both technical and legal means for preventing cybersecurity hazards, as well as the potential provision of specific consent procedures informing about AD security risks, and (2) the need to define legal means for prosecuting potential cybercrimes once they have occurred, as well as for determining how users will be indemnified for loses or harms derived from a security failure. Regarding the first concern, anonymization was raised as a primary technical mean for personal data protection. As to legal prevention means, most agreed that AVs will need to comply with functional safety standards and regulations used to protect the confidentiality of health and personal information and to ensure a safe use of the vehicle. Regarding the second concern, most experts agreed that, unfortunately, it may be never possible to fully secure AVs because of their complexity, and that therefore proper legal means for prosecuting hacking and other cybercriminal activities must be provided. The issue of liability arised again in this discussion, in relation to AV users' right to compensation in case of security failures. The issue of security was deemed as particularly important for office workers who might use sharing vehicle services to continue working while commuting, although most experts did not find that espionage on AVs more likely to occur than through other external connected devices.

3.5. Performance

Performance of the vehicle refers to the AV's effectiveness and functionality, its level of emissions and environmental impact, its contribution to fluidity of traffic, and its failure rates. Since AVs' performance will impact society in many ways, society has the right to have a voice in this topic, which will very likely express itself through legislation. It is necessary to understand who are the relevant agents that are going to decide AVs performance, and which are the objectives to be pursued via AD technologies. In this sense, disagreements between different agents are likely to occur, which is why it is important to establish clear decision-making processes. Probably, States will need to regulate the minimum performance requirements demandable for an AV model to be marketed. User profile diversity should also be taken into account in decisions about performance.

In this regard, the two major concerns were related to (1) the relevance of transparency, explainability, and the presence of meaningful human control across AVs design and implementation, and (2) the need to avoid algorithm discriminatory biases. Regarding the former, some experts highlighted the relevance of transparency in AD, not only in relation to "auditable evidences trails" regarding decisions in emergency situations, but at all times, so the driver knows how the AV was programmed. Explainable Artificial Intelligence (XAI) devices, whose decisions are traceable and explainable in natural language (Arrieta et al. 2020), should be of use in this regard. Regarding the second concern, all experts agreed that biased algorithms might introduce new issues in AD algorithms' decision-making. This would create new safety risks and perpetuate discrimination based on the ethnic group, gender, etc. The lack of guidance on ethical algorithmic systems that relevant stakeholders in the AD industry might receive is thus a pending issue to address.

3.6. Costs

Although AVs are expected to be expensive, both private and public sectors seem to consider investment in AD worthy.

In this regard, most discussions focused on (1) the issue of accessibility, (2) the issue of infrastructure availability, (3) the need to reconsider our transportation model, and (4) the use of sponsorship for reducing costs. Regarding the first concern, due to its likely high cost, AD technologies may not reach all population segments. Public authorities could intervene in this regard by introducing measures to enhance AD accessibility. Relatedly, since the adaptation of road infrastructures to AD will require major public investment, the support of this investment should be warranted through trustworthy public decision-making tools; thus, States should participate on AVs design and implementation to secure their social utility (for instance, by prioritizing green technologies). However, if individual vehicle ownership remains the preferable model, it may be too costly for States. One possible solution for addressing this problem could be to encourage Mobility as a Service (MaaS) initiatives, which promote the development of public and accessible means of commuting that are not necessarily linked to private vehicles (Hensher et al. 2020). In turn, this could help to achieve important societal goals like sustainability and accessibility. Finally, most experts considered sponsorship and advertising opportunities as a plausible help to reduce the cost burden. However, this might compromise drivers' autonomy and important social goals like sustainability, hence the need for specific legislation in this regard.

4. Conclusions

The ethical, legal, and socio-political implications of the future AD for data management, opacity, responsibility for bad outcomes, etc. have been central in several discussions with experts and lay-people. Some problems found for the future HMIs are new versions of old AI problems, but may be quantitatively more worrisome (i.e., the impact of marketing in a driver's free choice) and some problems are new (i.e., potential dangers in driver-vehicle interactions). There are also specific puzzles for overall AD as, for example, the need of new traffic regulations. Our work summarizes both the opinion of experts and the points of view of general non-expert population and presents overall recommendations on ethical and legal issues concerning the design of present and future AD-based technologies.

Acknowledgements

This study was funded by the European Union's Horizon 2020 research and innovation programme under grant agreement No. 875597 - HADRIAN (Holistic Approach for Driver Role Integration and Automation Allocation for European Mobility Needs) project. This document reflects only the author's view, the European Climate, Infrastructure and Environment Executive Agency (CINEA) is not responsible for any use that may be made of the information it contains.

Authors thank Prof. J. Rodriguez Alcazar (University of Granada) and Prof. F. Sanchez Barrilao (University of Granada) for their contributions during the first steps of the studies presented here.

References

Arrieta, A.B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., García, S., Gil-López, S., Molina, D., Benjamins, R., Chatila, R., Herrera, F., 2020. Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. Information Fusion, 58, 82–115.

Collingwood, L., 2017. Privacy implications and liability issues of autonomous vehicles. Information & Communications Technology Law, 26, 32–45.

Goodall, N. J., 2014. Machine ethics and automated vehicles, in "Road vehicle automation, lecture notes in mobility". In: Meyer, G., & Beiker, S. (Eds.). Springer International Publishing Switzerland, pp. 93–102.

Hensher, D. A., Mulley, C., Ho, C., Wong, Y., Smith, G., Nelson, J. D., 2020. Understanding Mobility as a Service (MaaS): Past, present and future. Elsevier.

Jenkins, R., 2016. Autonomous vehicles, ethics and law. Towards an overlapping consensus. New America.

Klaassen, M. & Szuprycinski, T., 2019. Security for V2X, in "Automotive systems and software engineering". In: Dajsuren, Y., & van den Brand, M. (Eds.). Springer Nature, pp. 283–294.

Kyriakidis, M., Happee, R., & de Winter, J., 2015. Public opinion on automated driving: Results of an international questionnaire among 5000 respondents. Transportation Research Part F: Traffic Psychology and Behaviour, 32, 127–140.

Mörtl, P., 2020. Human-Systems Integration for Driving Automation Systems: Holistic approach for driver role integration and automation allocation for European mobility needs, in "International Conference on Human-Computer Interaction". In: Krömker, H. (Ed.). Springer, Cham, pp. 78–88.

Mörtl, P., Trösterer, S., Cyril, M., Santuccio, E., Neuhuber, N., Zachaeus, C., Katrakazas, C., 2020. Initial automated driving application descriptions. Final version as of 29.05.2020 of deliverable D1.1 of the HORIZON 2020 project HADRIAN. EC Grant agreement no: 875597, ViF: Virtual Vehicle Research GmbH, Graz, Austria

Rodríguez-Alcázar, J., Bermejo-Luque, L., Molina-Pérez, A., 2021. Do automated vehicles face moral dilemmas? A plea for a political approach. Philosophy & Technology, 34, 811–832.

Roff, H. M., 2019. Artificial intelligence: Power to the people. Ethics & International Affairs, 33, 127–140.

SAE International, 2021. Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles: (Nr. J3016_202104; Surface Vehicle Recommended Practice).

Todt, O., Rodríguez Alcázar, J., Luján, J. L., 2010. Practical values and uncertainty in regulatory decision-making. Social Epistemology, 24, 349–362.

van Roermund, T., 2019. In-vehicle networks and security, in "Automotive Systems and Software Engineering". In: Dajsuren, Y., & van den Brand, M. (Eds.). Springer Nature Switzerland, pp 265–282.