European Commission

ETHICS of Connected and Automated Vehicles

Independent
Expert
Report

Ethics of Connected and Automated Vehicles
Recommendations on road safety, privacy, fairness, explainability and responsibility

More information on the European Union is available on the internet (http://europa.eu).

Ethics of Connected and Automated Vehicles

Recommendations on road safety, privacy, fairness, explainability, and responsibility

Independent Expert Group members

Jean-François BONNEFON – Chairman
David ČERNÝ
John DANAHER
Nathalie DEVILLIER
Veronica JOHANSSON
Tatiana KOVACIKOVA
Marieke MARTENS
Milos N. MLADENOVIC
Paula PALADE
Nick REED
Filippo SANTONI DE SIO – Rapporteur
Stavroula TSINOREMA
Sandra WACHTER
Karolina ZAWIESKA

TABLE OF CONTENTS

EXECUTIVE ABSTRACT

This report presents the work of a European Commission Expert Group established to advise on specific ethical issues raised by driverless mobility for road transport. The report aims to promote **a safe and responsible transition to connected and automated vehicles (CAVs)** by supporting stakeholders in the systematic inclusion of ethical considerations in the development and regulation of CAVs.

In the past few years, ethical questions associated with connected and automated vehicles (CAVs) have been the subject of academic and public scrutiny. A common narrative presents the development of CAVs as something that will inevitably benefit society by reducing the number of road fatalities and harmful emissions from transport and by improving the accessibility of mobility services. In contrast, this report applies a **Responsible Research and Innovation (RRI) approach** to CAVs. This approach recognises the potential of CAV technology to deliver the aforementioned benefits but also recognises that technological progress alone is not sufficient to realise this potential. To deliver the desired results, the future vision for CAVs ought to incorporate a broader set of ethical, legal and societal considerations into the development, deployment and use of CAVs.

To this end, this report presents a set of **20 ethical recommendations** concerning the future development and use of CAVs.

These recommendations are **grounded in the fundamental ethical and legal principles** laid down in the EU Treaties and in the EU Charter of Fundamental Rights (briefly described on p. 21).

The recommendations are presented and discussed in the context of three broad topic areas:

- **CHAPTER 1. ROAD SAFETY, RISK, DILEMMAS**

Improvements in safety achieved by CAVs should be publicly demonstrable and monitored through solid and shared scientific methods and data; these improvements should be achieved in compliance with basic ethical and legal principles, such as a fair distribution of risk and the protection of basic rights, including those of vulnerable users; these same considerations should apply to dilemma scenarios.

- **CHAPTER 2. DATA AND ALGORITHM ETHICS: PRIVACY, FAIRNESS, EXPLAINABILITY**

The acquisition and processing of static and dynamic data by CAVs should safeguard basic privacy rights, should not create discrimination between users, and should happen via processes that are accessible and understandable to the subjects involved.

• CHAPTER 3. RESPONSIBILITY

Considering who should be liable for paying compensation following a collision is not sufficient; it is also important to make different stakeholders willing, able and motivated to take responsibility for preventing undesirable outcomes and promoting societally beneficial outcomes of CAVs, that is creating a culture of responsibility for CAVs.

The recommendations are set out in terms that are intended to be actionable by three stakeholder groups in the context of their specific domains: **manufacturers and deployers** (e.g. car manufacturers, suppliers, software developers and mobility service providers); **policymakers** (persons working at national, European and international agencies and institutions such as the European Commission and the EU National Ministries) and **researchers** (e.g. persons working at universities, research institutes and R&D departments).

These recommendations are not intended to provide an exhaustive list of relevant ethical considerations. **Further research and collaboration** with stakeholders is needed on the impact of CAVs on topics such as sustainability, inclusiveness, and employment. Further work in this area should be supported by continual public deliberation to develop a shared collective identity and working culture that promotes the systematic integration of ethical considerations into the potential transition towards driverless mobility.

20 RECOMMENDATIONS

1. Ensure that CAVs reduce physical harm to persons.

To prove that CAVs achieve the anticipated road safety improvements, it will be vital to establish an objective baseline and coherent metrics of road safety that enable a fair assessment of CAVs' performance relative to non-CAVs and thereby publicly demonstrate CAVs' societal benefit. This should be accompanied by new methods for continuously monitoring CAV safety and for improving their safety performance where possible.

2. Prevent unsafe use by inherently safe design.

In line with the idea of a human-centric AI, the user perspective should be put centre-stage in the design of CAVs. It is vital that the design of interfaces and user experiences in CAVs takes account of known patterns of use by CAV users, including deliberate or inadvertent misuse, as well as tendencies toward inattention, fatigue and cognitive over/under-load.

3. Define clear standards for responsible open road testing.

In line with the principles of non-maleficence, dignity and justice, the life of road users should not be put in danger in the process of experimenting with new technologies. New facilities and stepwise testing methods should be devised to promote innovation without putting road users' safety at risk.

4. Consider revision of traffic rules to promote safety of CAVs and investigate exceptions to non-compliance with existing rules by CAVs.

Traffic rules are a means to road safety, not an end in themselves. Accordingly, the introduction of CAVs requires a careful consideration of the circumstances under which: (a) traffic rules should be changed; (b) CAVs should be allowed to not comply with a traffic rule; or (c) CAVs should hand over control so that a human can make the decision to not comply with a traffic rule.

5. Redress inequalities in vulnerability among road users.

In line with the principle of justice, in order to address current and historic inequalities of road safety, CAVs may be required to behave differently around some categories of road users, e.g. pedestrians or cyclists, so as to grant them the same level of protection as other road users. CAVs should, among other things, adapt their behaviour around vulnerable road users instead of expecting these users to adapt to the (new) dangers of the road.

6. Manage dilemmas by principles of risk distribution and shared ethical principles.

While it may be impossible to regulate the exact behaviour of CAVs in unavoidable crash situations, CAV behaviour may be considered ethical in these situations provided it emerges organically from a continuous statistical distribution of risk by the CAV in the pursuit of improved road safety and equality between categories of road users.

7. Safeguard informational privacy and informed consent.

CAV operations presuppose the collection and processing of great volumes and varied combinations of static and dynamic data relating to the vehicle, its users, and the surrounding environments. New policies, research, and industry practices are needed to safeguard the moral and legal right to informational privacy in the context of CAVs.

8. Enable user choice, seek informed consent options and develop related best practice industry standards.

There should be more nuanced and alternative approaches to consent-based user agreements for CAV services. The formulation of such alternative approaches should: (a) go beyond "take-it-or-leave-it" models of consent, to include agile and continuous consent options; (b) leverage competition and consumer protection law to enable consumer choice; and (c) develop industry standards that offer high protection without relying solely on consent.

9. Develop measures to foster protection of individuals at group level.

CAVs can collect data about multiple individuals at the same time. Policymakers, with assistance from researchers, should develop legal guidelines that protect individuals' rights at group levels (e.g driver, pedestrian, passenger or other drivers' rights) and should outline strategies to resolve possible conflicts between data subjects that have claims over the same data (e.g. location data, computer vision data), or disputes between data subjects, data controllers and other parties (e.g. insurance companies).

10. Develop transparency strategies to inform users and pedestrians about data collection and associated rights.

CAVs move through and/or near public and private spaces where non-consensual monitoring and the collection of traffic-related data and its later use for research, development or other measures can occur. Consequently, meaningful transparency strategies are needed to inform road users and pedestrians of data collection in a CAV operating area that may, directly or indirectly, pose risks to their privacy.

11. Prevent discriminatory differential service provision.

CAVs should be designed and operated in ways that neither discriminate against individuals or groups of users, nor create or reinforce large-scale social inequalities among users. They should also be designed in a way that takes proactive measures for promoting inclusivity.

12. Audit CAV algorithms.

Investments in developing algorithmic auditing tools and resources specifically adapted to and targeting the detection of unwanted consequences of algorithmic system designs and operations of CAVS are recommended. This will include development of CAV specific means and methods of field experiments, tests and evaluations, the results of which should be used for formulating longer-term best practices and standards for CAV design, operation and use, and for directly counteracting any existing or emerging ethically and/or legally unwanted applications.

13. Identify and protect CAV relevant high-value datasets as public and open infrastructural resources.

Particularly useful and valuable data for CAV design, operation and use, such as geographical data, orthographic data, satellite data, weather data, and data on crash or near-crash situations should be identified and kept free and open, insofar as they can be likened to infrastructural resources that support free innovation, competition and fair market conditions in CAV related sectors.

14. Reduce opacity in algorithmic decisions.

User-centred methods and interfaces for the explainability of AI-based forms of CAV decision-making should be developed. The methods and vocabulary used to explain the functioning of CAV technology should be transparent and cognitively accessible, the capabilities and purposes of CAV systems should be openly communicated, and the outcomes should be traceable.

15. Promote data, algorithmic, AI literacy and public participation.

Individuals and the general public need to be adequately informed and equipped with the necessary tools to exercise their rights, such as the right to privacy, and to actively and independently scrutinise, question, refrain from using, or negotiate CAV modes of use and services.

16. Identify the obligations of different agents involved in CAVs.

Given the large and complex network of human individuals and organisations involved in their creation, deployment and use, it may sometimes become unclear who is responsible for ensuring that CAVs and their users comply with ethical and legal norms and standards. To address this problem every person and organisation should know who is required to do what and how. This can be done by creating a shared map of different actors' obligations towards the ethical design, deployment and use of CAVs.

17. Promote a culture of responsibility with respect to the obligations associated with CAVs.

Knowing your obligations does not amount to being able and willing to discharge them. Similar to what happened, for instance, in aviation in relation to the creation of a culture of safety or in the medical profession in relation to the creation of a culture of care, a new culture of responsibility should be fostered in relation to the design and use of CAVs.

18. Ensure accountability for the behaviour of CAVs (duty to explain).

"Accountability" is here defined as a specific form of responsibility arising from the obligation to explain something that has happened and one's role in that happening. A fair system of accountability requires that: (a) formal and informal fora and mechanisms of accountability are created with respect to CAVs; (b) different actors are sufficiently aware of and able to discharge their duty to justify the operation of the system to the relevant fora; (c) and the system of which CAVs are a part is not too complex, opaque, or unpredictable.

19. Promote a fair system for the attribution of moral and legal culpability for the behaviour of CAVs.

The development of fair criteria for culpability attribution is key to reasonable moral and social practices of blame and punishment - e.g. social pressure or public shaming on the agents responsible for avoidable collisions involving CAVs – as well as fair and effective mechanisms of attribution of legal liability for crashes involving CAVs. In line with the principles of fairness and responsibility, we should prevent both impunity for avoidable harm and scapegoating.

20. Create fair and effective mechanisms for granting compensation to victims of crashes or other accidents involving CAVs.

Clear and fair legal rules for assigning liability in the event that something goes wrong with CAVs should be created. This could include the creation of new insurance systems. These rules should balance the need for corrective justice, i.e. giving fair compensation to victims, with the desire to encourage innovation. They should also ensure a fair distribution of the costs of compensation. These systems of legal liability may sometimes work in the absence of culpability attributions (e.g. through "no fault" liability schemes).

GLOSSARY OF SELECTED TECHNICAL TERMS

ARTIFICIAL INTELLIGENCE (AI): AI systems are software (and possibly also hardware) systems that, given a complex goal, act in the physical or digital dimension by perceiving their environment through data acquisition, interpreting the collected structured or unstructured data, reasoning on the knowledge, or processing the information, derived from this data and deciding the best action(s) to take to achieve the given goal. AI systems can either use symbolic rules or learn a numeric model, and they can also adapt their behaviour by analysing how the environment is affected by their previous actions[1].

AGENT: A human individual with the power to act on the basis of intentions, beliefs and desires. In this report, the term "agent" (and associated terms such as "agency" and "human agent") is used in this philosophical sense and not in the legal sense of a person who acts on behalf of another. In this philosophical sense, agency is typically understood to be a prerequisite for moral capacity and responsibility. The term is only used in relation to humans and is not used to refer to artificial agents or autonomous systems.

ALGORITHMS: Mechanisms for decision-making based on a set of digital rules and using input/output sources, encompassing Artificial intelligence (AI) algorithms, developed with the intention of mimicking human intelligence. In CAVs, algorithms are embedded in hardware and software, and can be based on other systems besides AI.

AUTOMATED DRIVING SYSTEM: Hardware and software that are collectively capable of performing the dynamic driving task on a sustained basis, regardless of whether it is limited to a specific operational design domain.

BLACK-BOX: In the context of AI and machine learning-based CAV systems, the black-box refers to cases where it is not possible to trace back the reason for certain decisions due to the complexity of machine learning techniques and their opacity in terms of unravelling the processes through which such decisions have been reached.

CONNECTED AND AUTOMATED VEHICLES (CAVS): Vehicles that are both connected and automated and display one of the five levels of automation according to SAE International's standard J3016, combined with the capacity to receive and/or send wireless information to improve the vehicle's automated capabilities and enhance its contextual awareness.

CULTURE: The ideas, practices, beliefs and values of a group of people. This term is used frequently in this report with respect to creating an ethical and responsible set of ideas, practices, beliefs and values among those involved in the manufacture, deployment and use of CAVs.

ETHICS: An academic discipline, a subfield of philosophy. It studies the norms, values and principles which should govern individual and group behaviour in society; grounding, integrating or complementing legal norms and requirements. This report takes a normative ethics perspective, insofar as it aims to guide as opposed to just describe the behaviour of stakeholders, in order to achieve societally positive outcomes in compliance with ethical principles. However, the report does not engage in a philosophical or legal discussion of normative principles but rather endorses the fundamental ethical and legal principles laid down in the EU Treaties and in the EU Charter of Fundamental Rights. Ethics of CAVs is therefore mainly an example of applied ethics insofar as it focuses on the specific, potentially new, normative issues raised by the design, development, implementation and use of CAV technology.

MACHINE LEARNING: The ability of systems to automatically learn, decide, predict, adapt and react to changes, improving from experience, without being explicitly programmed. Types of learning include *reinforcement, supervised, semi-supervised, unsupervised*[2].

MANUFACTURERS AND DEPLOYERS OF CAVS: Companies that build or sell vehicles with Automated Driving Systems (ADS) (second-hand vehicle sellers are not included), give assignments to provide software updates in order to change functionalities of the ADS, convert manually driven vehicles into vehicles with ADS, or companies that deploy CAVs for freight or mobility-related services.

OPERATIONAL DESIGN DOMAIN: The combined, operating conditions under which a given driving automation system (or feature thereof) is specifically designed to function, including, but not limited to, environmental, geographical, and time-of-day restrictions, and/or the requisite presence or absence of certain traffic or roadway characteristics (SAE International's standard J3016).

POLICYMAKERS: Persons working at national, European and international agencies and institutions such as the European Commission and the EU National Ministries, or any other organisation entitled to guide or influence the social and political processes concerning the design, development, implementation, regulation and use of CAVs.

PUBLIC DELIBERATION: Any social or political process through which individuals and groups not part of political or regulation bodies are engaged in discussions or decisions, in this specific report relevant for the design, development, implementation, regulation and use of CAVs.

RESEARCHERS: Individuals and organisations engaged in the professional, industrial, scientific or academic studies of topics relevant to CAVs.

THE PUBLIC: The aggregation of all individuals in society.

KEY DOCUMENTS REFERENCED IN THE REPORT

Document	Short name used in the report
AI High Level Expert Group. Ethics guidelines for trustworthy AI. B-1049 Brussels, 2019.	**AIHLEG Guidelines for Trustworthy AI**
Di Fabio, U., M. Broy, and R. J. Brüngger. Ethics Commission. Automated and Connected Driving. Federal Ministry of Transport and Digital Infrastructure of the Federal Republic of Germany, 2017.	**German Ethics Commission's guidelines**
European Commission Expert Group on Liability and New Technologies – New Technologies Formation, Liability for Artificial Intelligence and Other Emerging Technologies, Brussels, 2019.	**Expert Group report on Liability for AI**
European Group on Ethics in Science and New Technologies (EGE) statement on Artificial Intelligence, Robotics and 'Autonomous Systems', Brussels, 2018.	**EGE statement on Artificial Intelligence**
Regulation (EU) 2016/679 Of The European Parliament And Of The Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation).	**GDPR**
Santoni de Sio, F. Ethics and Self-Driving Cars: A White Paper on Responsible Innovation in Automated Driving Systems, Dutch Ministry of Transportation and Infrastructure Rijkswaterstaat, 2016.	**Dutch White Paper on Ethics and Self-driving Cars**
Task Force on Ethical Aspects of Connected and Automated Driving (Ethics Task Force), Report, Federal Ministry of Transport and Digital Infrastructure of the Federal Republic of Germany, 2018.	**Ethics Task Force report**

INTRODUCTION

The need to discuss **ethical issues raised by Connected and Automated Vehicles (CAVs) at European level** was recommended by the Ethics Task Force, a Member State initiative that was set up after the second High Level Meeting of EU Transport Ministers, the European Commission and Industry on Connected and Automated Driving in Frankfurt, September 2017[3]. As such, in its 2018 Communication On the Road to Automated Mobility: An EU Strategy for Mobility of the Future[4], the European Commission announced the creation of a Commission Expert Group to advise on specific ethical issues raised by driverless mobility.

The work of this **Independent Expert Group** started in June 2019, with the goal of providing practical support to relevant researchers, policymakers and CAV manufacturers and deployers in the safe and responsible transition to connected and automated mobility. This Expert Group consisted of **14 experts** from the fields of ethics, law, philosophy and CAVs from all over Europe, working independently and in the public interest.

This report aims to promote a **safe and responsible transition to connected and automated vehicles** by supporting stakeholders in the systematic inclusion of ethical considerations in the development and regulation of CAVs. This report provides **20 recommendations** to support researchers, policymakers and CAV manufacturers and deployers in dealing with a variety of ethical issues raised by connected and automated mobility. From June 2019 to June 2020, the Expert Group had six formal meetings to identify the issues to include in the report, and to discuss, deliberate, and draft the recommendations. One of these meetings took the form of a stakeholder workshop that aimed to foster a **participatory approach** in the preparation of the final report. The workshop gathered a variety of researchers, policymakers, associations and industry experts, who received a draft report with recommendations upon which they could propose revisions. This served as the basis for discussion during the workshop.

The report builds upon existing reports[5], such as the **AI High Level Expert Group Guidelines for Trustworthy AI** (AIHLEG), the **European Group on Ethics in Science and New Technologies** (EGE) statement on Artificial Intelligence, Robotics and Autonomous Systems, the **Ethics Task Force's report** and the **Expert Group report on Liability and New Technologies**, this Expert Group proposes recommendations to include ethical, societal, and legal considerations for the safe and responsible development of CAVs. Some of these recommendations may be used to develop new regulations concerning the development and use of CAVs.

However, this report relies on the idea that legislation alone may be insufficient to ensure that the development, deployment and use of CAVs is aligned with ethical principles and norms. The timely and systematic integration of broader ethical and societal considerations is also essential to achieve **alignment between technology and societal values** and for the public to gain **trust and acceptance of CAVs.** This means that policymakers and CAV manufacturers and deployers should work to ensure that this is achieved by CAVs' demonstrable compliance with as many ethical and societal needs and requirements as possible supported by stakeholder and user involvement in the process of their design, development, testing, implementation, regulation.

Ethical issues related to the emergence of new technologies such as Artificial Intelligence (AI), robotics, and autonomous systems have been discussed in numerous policy and regulatory documents in the last decade. These new technologies create scenarios and issues that are not fully covered by existing regulations, policies, and social practices. Consequently, a broad philosophical, political and societal discussion is required in order to guide the creation of new regulations, policies, and practices.

Specific attention has already been given to some of the novel challenges. For example, there has been widespread discussion of data protection rules, liability for autonomous systems, the protection of vulnerable technology users, and the avoidance of compromises between the use of a technology and a person's dignity and autonomy. The AIHLEG's Guidelines for Trustworthy AI, the General Data Protection Regulation (GDPR), the European Data Protection Board (EDPB), and the Expert Group report on Liability are good examples of recent contributions to this field.

This report frames and addresses some of these **emerging ethical and legal issues in the specific context of CAVs.** In the past few years, a number of other groups have tackled the specific topic of CAV ethics, such as the German Ethics Commission and the Ethics Task Force. These groups have laid out important general principles and recommendations, such as the advice to promote public participation in the development of CAVs, and reinforcing the prohibition against relying on factors such as age and gender in programming crash avoidance algorithms. They have also raised questions that need further exploration: Are there dilemma situations for which ethical recommendations simply cannot be prescribed? How should responsibility be distributed within new complex networks of software and technical infrastructures controlling traffic? What are the requirements in terms of safety, human dignity, personal freedom of choice and data protection that have to be fulfilled before approving automated driving systems?

This report acknowledges existing principles and recommendations, and addresses some of the open questions raised in previous reports. It also raises some new issues and questions, such as how to promote the moral responsibility of manufacturers and deployers of CAVs, and provides some original, specific recommendations to address these issues in law, policy, and social practice.

In the past few years, many ethical questions associated with CAVs have been the subject of scientific and academic scrutiny. They have also been widely covered by

the media and discussed in the public domain. This report **relies on available scientific and academic knowledge** and, where appropriate, tries to demystify some misleading popularised ideas of CAV ethics with the aim of promoting greater understanding and better-informed public debate.

A common narrative presents the anticipated societal benefits of CAVs as something that will inevitably happen, simply through the promotion of technological development and innovation. According to this "solutionist" narrative[6], the development and uptake of CAVs will reduce the number of road fatalities, reduce harmful emissions from transport and improve the accessibility of mobility services. The deployment of CAVs would thus improve the mobility system as a whole, strengthen the competitiveness of European industry and support the Digital Single Market strategy.

In contrast to this "solutionist" narrative, this report applies the **Responsible Research and Innovation (RRI)** approach[7]. This recognises the potential of CAV technology to deliver the aforementioned benefits but also that **technological progress alone is not sufficient** to realise this potential. To deliver the desired results, the future vision for connected and automated driving should include a broader set of ethical, legal and societal considerations throughout the development, deployment and use of CAVs. This will ensure that relevant scientific, technical, societal, and legal challenges are raised and addressed in a timely manner; that the risk of adverse, undesirable outcomes is minimised; and that the expected gains of the technology are realised for society as a whole.

RRI aims to support stakeholders in translating shared moral values into practical requirements and recommendations for the design, development, and use of technology. As a result, stakeholders can systematically and pro-actively integrate these values into their processes and products in a timely fashion, rather than having to adapt to these values, possibly at a late stage of development.

Consider, as a first topic, the **safety of CAVs.** An academic and public debate on so-called "moral dilemmas" with automated vehicles has vividly shown that crash avoidance by CAVs is not only a technical challenge but also an ethical and societal one[8]. Dilemma situations are rare accident scenarios in which a highly automated CAV finds itself confronting an unavoidable crash and yet has the possibility of choosing between the road users that will be harmed by the event (e.g. a group of pedestrians, or the CAV's occupant)[9].

Interesting though they may be, moral dilemmas in crash avoidance are not the only, nor even the most urgent, ethical and societal issue raised by CAV safety[10]. To begin with, lower levels of vehicle automation may create serious and possibly more frequent safety issues too. From this standpoint, various recent crashes involving CAVs on public roads have been a strong wake-up call[11]. This is one reason why this report recommends moving away from a narrow focus on crash avoidance towards a broader focus on a set of ethical considerations required to facilitate the safe transition to automated driving.

Technology should not be over-trusted, and the technical development of CAVs should

be accompanied and guided by, among other things, new suitable safety baselines and models, and **enhanced metrics for measuring traffic safety**. Current safety measurements and predictions, such as those based on a simple comparison of injury and fatality rates per million kilometres may give an incomplete and misleading picture. Relatedly, **responsible standards for open road testing** and careful (re)consideration of traffic rules in the context of CAVs are needed. These may allow us to develop and introduce CAV technologies without creating unreasonable risks for the public. In addition to this, as the recent crashes involving advanced driving assistance systems show, CAV technologies should be designed to reflect the road users' psychological capabilities and motivations, and CAV development should not only reduce the overall number of crashes, but also improve the safety of all road users, on all fronts, including the most vulnerable, such as children, cyclists or elderly persons.

Moreover, bringing attention to the broader set of ethical and societal considerations recommended by the RRI approach, allows us to see that harm and injustice can derive not only from road crashes. They can also come from an insufficiently responsible design of **the collection and processing of CAV and user data,** or from the lack of transparency in the **algorithmic decision-making of CAVs.** It has become evident that the combination of AI and big data in commercial products makes it difficult to ensure that these comply with ethical and legal standards relating to the **respect of privacy and non-discrimination.** This highlights the importance of proactively addressing new specific data related ethical and legal issues with CAVs, from the early

stages of their development. It also highlights the need to devise technical and societal interventions to enhance the **explainability of the processes of acquisition and use of data** by CAVs.

Finally, a broader and more proactive ethical approach will also help to reveal new perspectives on the often-asked question of **who is responsible for the behaviour of CAVs.** CAVs are complex socio-technical systems: many individuals, organisations, and technologies interact in the manufacture, deployment and use of CAVs. Moreover, manufacturers, deployers and users will interact differently and at different stages with these intelligent, AI-equipped systems. One evident consequence, already highlighted by the German Ethics Commission and the Ethics Task Force as well as the Expert Group report on Liability, is that attribution of liability for crashes may become more difficult as traditional moral and legal concepts may not be easily applicable to these new scenarios.

In this report, we will discuss these as issues of "backward-looking" responsibility for CAVs (that is responsibility for past accidents), and will propose some recommendations to address them. However, this report also urges the importance of creating new concepts and tools that facilitate "forward-looking" responsibility for CAVs. This will include principles and recommendations to establish what different human agents should do to **ensure CAVs' safety, the responsible use of data and the accountable development of algorithms** and other technical features, before CAVs are on the road. This report recommends that opportunities and incentives be created for policymakers, manufacturers and deployers

of CAVs, users and other agents in the CAV network to create a shared understanding of their respective responsibilities, and to create a culture of good practices and habits for each of them to be able and motivated to fulfil these (new) responsibilities.

This report is not meant to be the last word on the ethics of CAVs. There are three reasons for this. First, it only covers a selected set of ethical issues: safety, data and algorithm ethics, and responsibility. Other important issues such as the connection between CAVs and environmental sustainability, the future of employment, and transport accessibility are not discussed. The Expert Group views these issues as at least as important as the ones that are discussed in this report. The choice to focus only on the three defined sets of ethical issues was made with the intention of making the most of the expertise within the Group in the limited time available.

Second, the RRI approach requires that stakeholders are actively engaged in translating general principles into practice, based on further empirical evidence and technological insight they may acquire on the field. This means that the recommendations contained in this report can and should be further discussed in **future stakeholder meetings,** on the basis of further data and experience in the development and deployment of CAVs. This should be supported by continual public deliberation to develop a shared collective identity and working culture that promotes the systematic integration of ethical considerations into the potential transition towards driverless mobility.

Third, and more specifically, researchers, policymakers, manufacturers and deployers of CAVs will sometimes have to take the extra step of **bringing the recommendations to their specific policy or industry domains,** and thus identifying the specific tools needed to translate them into living policies and practices.

GUIDING

ETHICAL PRINCIPLES

According to the Responsible Research and Innovation approach, the design and implementation of connected and automated vehicles should be built upon ethical guidelines grounded in **fundamental ethical and legal principles** that have been critically and reflexively adopted by society. In line with the EGE's statement on Artificial Intelligence and the AIHLEG's Guidelines for Trustworthy AI, we propose that our analysis and recommendations be guided by the following ethical and legal principles, as laid down in the EU Treaties and in the EU Charter of Fundamental Rights.

• NON-MALEFICENCE

Primum non nocere. The physical and psychological integrity of human beings ought to be respected. The welfare of other living beings and the integrity of the planet ought also to be protected. In relation to CAVs, this would mean, for example, that the first and foremost ethical requirement for CAVs is not to increase the risk of harm for road users (including users of CAVs or other road users that are in interaction with CAVs) compared to manual driving.

• BENEFICENCE

CAV technology should be designed and operated to contribute positively to the welfare of individuals, including future generations, and other living beings, as long as this is consistent with the *principle of non-maleficence.* CAV operations should not only aim at the minimisation of costs but should also benefit persons. CAVs primary purpose should be to enhance mobility opportunities and bring about further benefits to persons concerned, including enhancing the mobility opportunities of persons with special needs. CAVs should contribute to improved sustainability and environmental friendliness of the transport system. CAVs' social and societal impact ought always to be carefully considered.

• DIGNITY

Every individual human possesses intrinsic worth that should not be violated or traded for the achievement of any other ends. Dignity is the basis of the equality of all human beings and forms the normative point of reference that grounds human rights. In relation to CAVs, respecting human dignity requires that fundamental individual rights (e.g. life) are not infringed upon or sacrificed in the name other societal goods.

• PERSONAL AUTONOMY

Human beings should be seen as free moral agents, who demand respect for the conditions of their agency. This requires that CAVs should protect and promote human beings' capacity to decide about their movements and, more generally, to set their own standards and ends for accommodating a variety of conceptions of a 'good life'. In relation to CAVs, this requires, among other things, protecting users from unreasonable restrictions of their capacity to move and from hidden and aggressive marketing (e.g. mobility data used by third parties for commercial purposes). To this end, the effective application of relevant EU consumer protection and data protection law is a solid starting point for further ethical efforts.

• RESPONSIBILITY

The counterpart of human autonomy is human responsibility. Both individual persons and institutional stakeholders can and should be held morally and legally responsible for the consequences of their actions when it is appropriate to do so. At the same time, they should be given a fair capacity and opportunity to behave according to moral and legal expectations. In relation to CAVs, this means establishing clear moral and legal standards of responsibility, while at the same time providing different actors (CAV users, but also CAV manufacturers and deployers) with sufficient knowledge, capacity, motivation and opportunities to comply with these standards.

• JUSTICE

Justice concerns the question of how we ought to distribute fairly the benefits and burdens of newly emerging technologies. Injustice occurs when the benefits to which an individual is entitled are unjustifiably denied, or when some burden is unduly imposed upon somebody without adequate compensation. In relation to CAVs, that would mean, among other things, that CAVs should provide equality of access to mobility for all and should be calibrated by developers to reduce disparities in exposure to harm between categories of road users.

• SOLIDARITY

Solidarity concerns pro-social actions and practices, as well as institutional and political regulations designed to assist others, against the background of a group's common goals within a mutually supportive community. It requires the protection and empowerment of vulnerable persons or groups and complements the requirements of justice. In relation to CAVs, that would mean, among other things, promoting data-sharing about fatalities and injuries caused by CAVs among the appropriate safety agencies.

- **INCLUSIVE DELIBERATION**

The above principles cannot be applied with a mechanical top-down procedure. They need to be specified, discussed and redefined in-context. Inclusive deliberation ensures that perspectives from all societal groups can be heard, and no one is disregarded. Moreover, tensions between these principles may arise in specific applications. That is why the design and development of CAV systems should be supportive of and resulting from inclusive deliberation processes involving relevant stakeholders and the wider public.

Chapter 1

Road safety, risk and dilemmas

1.1 INTRODUCTION

A central promise of CAVs is to improve road safety by decreasing the frequency of crashes and/or limiting the harm that they cause. We consider the possible steps to ensure that this goal is pursued in an ethically appropriate manner. Recommendations focus on the limitation of physical harm. While they may also help to reduce psychological harm, we leave detailed consideration of psychological harm to future initiatives.

Even in the best-case scenario CAVs will not totally eliminate crashes in the foreseeable future. Consequently, a difficult ethical issue is to define what would be considered an appropriate distribution of the residual risk arising from their use. That is to say: how will the remaining crashes be statistically distributed among different categories of road users? We consider some possible recommendations for an ethically grounded distribution of risk among road users. Our recommendations go beyond the guidelines of the German Ethics Commission and align with the AIHLEG's Guidelines for Trustworthy AI by considering CAVs as a tool to correct current and historic inequalities in the vulnerability of different categories of road users.

A limit case of risk redistribution has already garnered global attention: the case in which a CAV may have to actively decide between one of two possible unavoidable crash outcomes at a given point in time. We consider these "dilemma" cases and how they may be organically solved by the emerging properties of the rules and methods introduced for the distribution of residual risk. As a continuation of the German Ethics Commission, we provide positive recommendations for solving dilemmas, in addition to considering negative recommendations about unethical solutions.

1.2 IMPROVING ROAD SAFETY

Ensure that CAVs reduce physical harm to persons
Recommendation 1

In line with the principle of **non-maleficence**, a minimal requirement for manufacturers and deployers is to ensure that CAVs decrease, or at least do not increase, the amount of physical harm incurred by users of CAVs or other road users that are in interaction with CAVs, compared to the harm that is inflicted on these groups by an appropriately calculated benchmark based on conventional driving. A further requirement, in line with the principle of **justice**[12], is that no category of road user (e.g. pedestrians, cyclists, motorbike users, vehicle passengers) should end up being more at risk of harm from CAVs than they would be against this same benchmark. In line with the principle of **dignity**, other possible benefits of CAVs, such as positive environmental impact or

congestion reduction, should not be considered as compensating for an increase in physical harm to road users. Note that observing an average decrease in harm across all CAVs would not mean that Recommendation 1 has been satisfied — Recommendation 1 applies to each new model or update of CAVs, not to the aggregation of all CAVs produced by a single developer, or the aggregation of all existing CAVs.

In order to pursue Recommendation 1, **manufacturers and deployers,** together with **policymakers** and **researchers** must collaboratively define the metrics and benchmarks that will be used as evidence for the net positive effect of CAVs on road safety. **Researchers** should be supported to develop new methods (possibly on the basis of new measures) to do this in a scientifically sound manner (explored in the discussion of this Recommendation). In the short term, **manufacturers and deployers** should be clear about the benchmarks to which they are comparing their CAV safety metrics. In the long term, **policymakers** will need to define a standard set of benchmarks against which the safety metrics of CAVs will be compared; this may benefit from international collaboration between researchers to develop these benchmarks, building on existing road safety data metrics.

Finally, CAV safety performance should not be assessed as a single snapshot but continuously monitored and improved. The data used for this continuous improvement should include fatality and injury rates, as well as data about near-collisions and crash-relevant conflicts, and function-by-function safety metrics (e.g. traffic light recognition, sudden braking, lane deviations). **Policymakers** should encourage the accessibility of data about collisions and near-collisions for independent crash investigation agencies and for researchers (see also **Recommendation 13** on **accessibility of datasets**).

Discussion of Recommendation 1

Injury and fatality rates per million kilometres are the most commonly described and straightforward metrics of road safety. However, (a) the rarity of these events, (b) the scarcity of CAVs on the road, (c) the lack of details about the exact circumstances of a collision, (d) the possible uncertainty about whether automated driving was activated before a crash, (e) the fact that these vehicles or functions are used in different circumstances or on different types of roads, and (f) the large under-registration of crashes — together mean that conclusive statistics about injury and fatality rates may require decades of testing, registrations and new ways of logging statistics[13].

Furthermore, the safety metrics of CAVs cannot be simply compared to the safety

metrics of human drivers or conventional vehicles[14]. If, for example, CAVs are typically tested under favourable weather conditions, then their safety metrics should not be compared to that of human drivers operating under adverse weather conditions. Likewise, current partially automated functions are primarily used on motorways (or comparable roads, with more consistent traffic dynamics and larger radius curves), in which collision statistics are typically low. Accordingly, the collision statistics when the systems are engaged (under favourable conditions) and when they are not (under less favourable conditions), cannot be directly compared. Similarly, CAVs may be equipped with state-of-the-art safety features that do not relate to their automated driving capacity. As a result, their safety metrics should be compared to that of conventional vehicles benefiting from the same features. Enthusiasm about the promise of CAVs to improve road safety should not obfuscate the complexity of the realisation of this goal: **researchers, policymakers** and **manufacturers and deployers** have an ethical obligation to maintain a scientifically sound and critical approach in this respect.

Prevent unsafe use by inherently safe design
Recommendation 2

CAV users will inevitably use some automated driving functions in an unsafe manner, either intentionally or not. **Manufacturers and deployers,** together with **researchers,** should create intuitive, user-centred systems that are designed to prevent unsafe use. Where relevant, CAV systems should offer unambiguous and timely guidance concerning the possible overrides or handovers required when a system reaches the limits of its operational design domain. **Manufacturers and deployers** should safely and clearly provide users with appropriate in-car guidance by means of an inherently safe interface that shows how the CAV operates and how it is designed to cope with incorrect use and potential misuse. System design should account for known limitations of human performance[15]. These system design recommendations are in line with the principle of **beneficence**.

At the same time, and in line with the principle of responsibility, to the extent to which they were provided with sufficient capacity and fair opportunity to make safe use of automated driving functions, CAV users should bear some of the moral and legal responsibility for obvious incorrect behaviour (see **Recommendation 19** on **attribution of culpability** for more). Systems that monitor the driver's state may provide useful information to support the safe operation of vehicle automation systems: **researchers** should investigate further how driver-monitoring systems can reliably support safe use of CAVs while complying with the requirement of data and algorithm ethics as presented in **Recommendation 7** on

privacy and informed consent; <u>Recommendation 8</u> on **user choice;** <u>Recommendation 10</u> on **transparency of data collection** and <u>Recommendation 11</u> on **non-discriminatory service provision.**

Discussion of Recommendation 2

Generally speaking, careful system design should enable safe use of automation and prevent users from intentionally or unintentionally using such systems unsafely. Handover scenarios (featuring a transition of control from CAV to user) should provide enough time for the user to regain situational awareness[16 17]. Sudden handover requests are unsafe and should always be avoided.

In one simulator study[18], no participant was able to keep the car on the road when they had two seconds to react to a sudden failure just before a curve. However, other simulator studies in the context of truck platooning[19] (i.e., the electronic coupling of several trucks allowing them to maintain a short-gap, tight formation) showed that professional truck drivers responded well to timely handover requests in non-critical situations — in such contexts, drivers must indicate to the system that they are ready to take over after an initial take-over request.

Define clear standards for responsible open road testing
Recommendation 3

In line with the principle of <u>**non-maleficence,**</u> standards for open road testing and the procedures for deciding if a CAV is ready for open road testing must be carefully defined by **policymakers** in a joint effort with **manufacturers and deployers.** New facilities and stepwise testing methods should be devised to promote innovation without putting road users' safety at risk[20]. **Researchers, policymakers** and **manufacturers and deployers** should not fuel unreasonable expectations about the capabilities of CAVs, and should collaborate by contributing to public debates that realistically reflect the state-of-the-art in CAV technology.

Discussion of Recommendation 3

A comprehensive and rigorous framework for open road testing would be most appropriately addressed at the European level, and should identify (a) the levels of testing that should be conducted before testing on open roads, including, for example, the use of simulation, hardware-in-the-loop testing or dedicated automotive proving grounds in a diverse range of driving environments; (b) the mix of audit, self-certification and third-party testing that will lead to certification for open road testing; (c) the measures that must be taken to mitigate risks incurred by uninformed road users, for example the use of geofencing and the presence of a safety driver and (d) the evidence that manufacturers and deployers must provide to show the effectiveness of these risk mitigation measures.

Consider revision of traffic rules to promote safety of CAVs and investigate exceptions to comply with existing traffic rules by CAVs

Recommendation 4

Traffic rules are a means to road safety, not an end in themselves. Accordingly, the pursuit of greater road safety may sometimes require non-compliance with traffic rules, in line with the principle of non-maleficence. **Policymakers** and **researchers** should use data provided by manufacturers and deployers to identify contexts in which it would be more appropriate to (a) change a traffic rule so that CAVs can act safely without engaging in non-compliance, (b) have the CAV handover control so that a human can make the decision to not comply with a traffic rule, or (c) allow the CAV to not comply with a traffic rule if it can explain why it made this decision and leave it to the justice system to decide whether this non-compliance was justified by the pursuit of greater safety. **Researchers** should study the extent to which it is reasonable to expect that an intelligent non-human system is able to engage in the complex process of evaluation of the interpretation of a legal, ethical or societal norm and its balancing with another norm, value or principle.

Researchers should also test the ex-post explainability of these decisions (see **Recommendation 14** on **algorithmic explainability**). The pursuit of comfort or efficiency should not be considered a justification for non-compliance. More generally, **policymakers** may need to consider the modification of traffic rules to accommodate a heterogeneous fleet of CAVs and human driven vehicles[21].

Discussion of Recommendation 4

It may be ethically permissible for CAVs not to follow traffic rules whenever strict compliance with rules would be in conflict with some broader ethical principle. Non-compliance may sometimes directly benefit the safety of CAV users or that of other road users, or protect other ethical basic interests; for example, a CAV mounting a kerb to facilitate passage of an emergency vehicle. This is a widely recognized principle in morality and in the law.

However, the extent to which this principle can and should apply to the behaviour of CAVs should be carefully considered[22]. Uncertainty in the application and interpretation of rules (and the necessity of their violation) may necessitate the involvement of a human operator (the user inside a vehicle, a remote operator, or a worker in a remote centre issuing an authorisation to engage in non-compliance). This transfer of responsibility should only occur if the human operator has sufficient time and information to make responsible control decisions and in no circumstance should the human operator be assigned a task for which humans are unsuited or for which they have not been sufficiently trained (see **Recommendation 17** on **culture of responsibility** and **Recommendation 19** on **attribution of culpability**). Situations in which a CAV chooses not to comply with traffic rules, or transfers control to a human operator should be carefully and extensively studied and discussed, and should be recorded to ensure that the decision can be analysed and justified, although this would require due consideration of privacy concerns as well (see **Recommendation 7** on **privacy and informed consent**, and **Recommendation 8** on **user choice**).

1.3 RISK DISTRIBUTION

Redress inequalities in vulnerability among road users
Recommendation 5

CAVs may offer the opportunity to redress some inequalities in vulnerability among road users, in line with the principle of justice. Researchers can use current traffic collision statistics to reveal which categories of road users incur a disproportionate amount of harm, relative to their road exposure (see the discussion of Recommendation 5, below). CAVs may then be calibrated by manufacturers and deployers to reduce strong disparities in the ratio of harm-relative-to-road-exposure between different road users.

In other words, in order to create greater equality in the safety of all road users, **policymakers**

may require **manufacturers and deployers** to develop and deploy CAVs that behave differently around some categories of vulnerable road users than other less vulnerable users. The ethical and social acceptability of such measures may be a topic of investigation for **researchers** as well as a topic for inclusive deliberation.

Discussion of Recommendation 5

This recommendation amounts to using CAVs to change the focus from vulnerable users needing to adapt to the dangers of the road to CAVs needing to adapt to vulnerable road users. This is in line with the Ethics Guidelines for Trustworthy AI produced by the AIHLEG, which recommends that particular attention should be paid to vulnerable groups, to the historically disadvantaged, or to those who suffer disproportionately under existing asymmetries of power.

For example, assume that cyclists are found to incur a disproportionately high share of fatalities compared to their share of road exposure (e.g. share of person-hours of road use out of the total number of person-hours of road use for all categories of road users), and that car crashes are involved in a substantial proportion of these fatal events. In such a case, **policymakers** may require **manufacturers and deployers** to show evidence as to how their vehicles operates to reduce risk for cyclists so that their harm-relative-to-road-exposure ratio decreases. The means for achieving that goal may include slowing down when cyclists are detected, but also giving cyclists more space, even if that behaviour gives less space to other less vulnerable road users (as long as the total harm to these other road users does not increase either). This programming would not amount to giving greater value to the safety of cyclists—it would rather be an attempt to correct safety inequalities, which partly result from the current behaviour of human drivers.

Such decisions, though, should be carefully debated in line with the **inclusive deliberation** principle. For example, if the hypothesis that some categories of road users are more vulnerable than others is confirmed by scientific research, based on such evidence, **manufacturers and deployers** may program CAVs to be more cautious around users whose behaviour is less predictable, by slowing down and giving them more space. This may be true, for example, for young children whose less predictable behaviour may create greater uncertainty in the calculations of the CAV. This may also be true for road users whose mobility is reduced, for example wheelchair users; for visually impaired users (especially if CAVs are predominantly electric and silent and thus less detectable by visually impaired users) or for

pedestrians walking in a large group, if movements of an individual within the group are more likely to be obfuscated by others in the group, or if their mobility is impeded by the group.

In all these examples, the recommendation to provide greater road safety to a subset of road users must always be premised on evidence that it is technically possible for a CAV to detect and respond to these road users accurately and reliably, that some users' harm-to-exposure ratio is high, that improving road safety for one subset of road users does not raise the total harm inflicted to another category of road users above its current baseline.

1.4 DILEMMAS IN CRASH-AVOIDANCE

Manage dilemmas by principles of risk distribution and shared ethical principles

Recommendation 6

Dilemmas are defined as critical situations in which, at a given point in time, a CAV will inevitably harm at least one road user and/or one group of road users and the CAV's behaviour will eventually determine which group or individual is harmed[23]. In regulating the development and deployment of CAVs, **policymakers** may accept that the behaviour of CAVs in dilemma situations can organically emerge from the adherence to the principles of risk distribution stated in [Recommendation 5](#) on **inequalities**. Adherence to these principles of risk distribution should ensure that the behaviour of the CAV does not conflict with basic ethical and legal principles.

In light of the broad public debate raised by the dilemma-based situations, and the public concerns that CAVs may be programmed by developers to select their collisions based on some non-transparent, or otherwise ethically and socially unacceptable criteria, **researchers, policymakers,** and **manufacturers and deployers** should reassure the general public about Recommendation 6, and engage the general public in an inclusive deliberation process about its possible implications.

Recording and reviewing the outcomes of a dilemma (and other safety critical situations encountered by CAVs (even if they are only identified post hoc)) can still serve as a basis upon which to update CAV software and their future behaviour. In line with the principle of [solidarity,](#) sharing data with appropriate safety agencies, as long as this respects data protection legislation, should be encouraged for that purpose by **policymakers.** Inspiration

can be taken from data sharing policies in the domain of security or air transportation, which may inform policymakers about the best way to give due consideration to the sensitive nature of these data in relation to security, commercial interest, and privacy (see Recommendation 13 on **accessibility of datasets**). In some cases, it might be appropriate for **manufacturers and deployers** to share information extracted from the data, rather than the raw data themselves.

Discussion of Recommendation 6.

Providing guidance for the decisions of CAVs in dilemma situations raises major challenges. First, it may be ethically and legally impermissible to let CAVs actively decide to enter in a collision with one or another specific individual in a critical situation[24]. This would go against the principles of **non-maleficence** and **dignity.**

Second, the CAV may be in a considerable state of uncertainty regarding the possible outcomes of its decisions in a dilemma situation[25]. In fact, it may be hard to pinpoint the exact moment at which a CAV transitions from continuous multi-dimensional risk management to a genuine dilemma situation. Accordingly, this report treats dilemmas as a limit case of risk management[26].

Rather than defining the desired outcome of every possible dilemma, it considers that the behaviour of a CAV in a dilemma situation is by default acceptable if the CAV has, during the full sequence that led to the crash, complied with all the major ethical and legal principles stated in this report, with the principles of risk management arising from **Recommendation 5** and if there were no reasonable and practicable preceding actions that would have prevented the emergence of the dilemma. This may be necessary in order to give manufacturers and deployers of CAVs the confidence to deploy their systems, with reduced speed and preventative manoeuvres always being the best solution to decrease safety risks.

Chapter 2

Data and algorithm ethics: privacy, fairness, and explainability

2.1 **INTRODUCTION**

CAV operations require the collection and use of great volumes and varied combinations of static and dynamic data relating to the vehicle, its users, and the surrounding environments. Through algorithms and machine learning, these data are used for CAV operations on different time scales, ranging from second-by-second real-time path planning and decision-making, to longer-term operational parameters concerning choice of routes and operating zones, up to longest-term user profiling and R&D investments.

Consequently, data subjects need to be both protected and empowered, while vital data resources need to be safeguarded and made accessible to specific actors. This can only occur after due consideration of ethical principles of **human dignity** and **personal autonomy.** In this context, these fundamental principles are tied to specific principles concerning privacy, fairness, and explainability.

First, the notion of privacy encompasses each individual's authority to determine a private sphere for personal conduct and self-development, including privacy of communications and the ability to control the terms and conditions of personal information sharing. Privacy is not only an ethical imperative but an enforceable fundamental right in the EU. Standardly, respect for privacy requires a valid legal basis (pursuant to Article 6 GDPR) for any collection, processing, use or exchange of personal data.

Second, fairness and explainability are binding data protection principles that are enshrined in secondary EU law (e.g., the GDPR, the Law Enforcement Directive, and the data protection instruments that apply to the EU institutions). Fairness requires that personal data collection, processing, uses, and outcomes do not discriminate negatively against any individual or group of data subjects. This entails that data-driven CAV operations should be as inclusive as possible, and that equal access and opportunities need to be safeguarded for all parties, particularly for potentially vulnerable persons and groups.

Finally, in line with previous reports[27], explainability (Explainable AI) requires that the objectives, mechanisms, decisions and actions pursued by data- and AI-driven CAV operations should be accessible, comprehensible, transparent and traceable to users and data subjects, in a way that goes beyond a strictly technical understanding for experts.

2.2   PRIVACY AND DATA PROTECTION

Safeguard informational privacy and informed consent
Recommendation 7

In line with the GDPR basic principles regarding data minimisation, storage limitation and the strict necessity requirements of Article 5, **manufacturers and deployers** of CAVs, as those who decide the means and the purposes of personal data processing (referred to as "data controllers" under the GDPR), have to inform data subjects about the predefined purposes for which their data are collected. In the event that **manufacturers and deployers** wish to collect data for purposes that are not necessary for the proper functioning of the CAV, such as advertising, selling products to the CAV users, or sharing data with third parties, they have to seek the data subject's explicit, free, and informed consent. Otherwise, such use is to be prohibited altogether.

Moreover, **manufacturers and deployers** ought to facilitate data subjects' control over their data through the implementation of specific mechanisms and tools for the exercise of their rights, particularly their rights of data access, rectification, erasure, restriction of processing, and, depending on the particular legal basis of the processing, their right to object or right to data portability (e.g. moving to another service provider).

Manufacturers and deployers should actively inform users about the consequences if they do not agree to share their data. The data subject's objection to collecting or sharing of data that is not necessary for the proper and safe operation of the CAV, must not result in a denial of service. **Manufacturers and deployers** ought to take all the necessary measures to ensure that there is reliable and sufficient protection against manipulation, misuse or unauthorised access to either the technical infrastructure or the associated data processes.

Policymakers should set further legal safeguards and enforce the effective application of data protection legislation, notably provisions on organisational and technical safeguards, to ensure that the data of the CAV user are only ever disclosed, or forwarded, on a voluntary and informed basis. **Policymakers** and **researchers** should make sure that the development of such measures is conducted and grounded in responsible innovation processes with a high-level of engagement between stakeholders and the wider public.

Discussion of Recommendation 7

Data-driven CAV technology can technically be used to identify and monitor vehicle passengers through sensors and video monitoring inside the vehicle. It can also be used for personal identification requirements (facial recognition, biometric data, etc.). This data can technically be collected and associated with users, developing their profile over time in conjunction with background information[28].

With these technical possibilities, concerns arise about uncritical or improper fine-grained profiling and its potentially illegal applications, including manipulation and misuse. CAV users should have control over their personal data. This data should only be disclosed, forwarded and used on a voluntary basis to the point that all terms and conditions for data provision to second and third parties have to adhere to the highest standards of free, informed and explicit consent.

Enable user choice, seek informed consent options and develop related best practice industry standards

Recommendation 8

Policymakers, manufacturers and deployers and **researchers** should work together towards formulating more nuanced and alternative approaches to consent-based user agreements for CAV services. The formulation of such alternative approaches should (a) go beyond "take-it-or-leave-it" models of consent, to include agile and continuous consent options, (b) leverage competition and consumer protection law to enable consumer choice, and (c) develop industry standards that offer robust protection without relying solely on consent.

Article 7 of the GDPR prohibits forced consent. **Manufacturers and deployers,** especially mobility service providers, have to comply with this provision and offer agile consent management tools. Public authorities should oversee the implementation and enforcement of this requirement. **Policymakers** should also leverage competition and consumer law to counteract monopolies and enable user choice. One promising example of this could be the elaboration of rules that prevent only one provider from operating in certain zones or for certain types of services. Competition laws should be rapidly developed to combat monopolies and maintain adequate competition conditions for the CAV service market in order to shift power in favour of users.

Finally, user consent may not always be a sufficient measure to gauge a data subject's

privacy rights[29]. Thus **policymakers** must ensure that new industry standards around "reasonable algorithmic inferences"[30] are established. Such best practice standards should address ethical data sharing, transparency and business practices (e.g. with insurers, advertisers or employers) and give guidance on grounds for and boundaries of legally and ethically acceptable inferential analytics (e.g. unlike inferring race or age to offer goods and services). The aim of those standards is to guarantee a high data protection standard without solely relying on users' consent.

Research in the legal, philosophical, technical, and social domains needs to identify alternative and CAV-specific solutions to protect informational privacy and informed consent, and establish best practices for industry.

Discussion of Recommendation 8

Access to and aggregation of personal data, as invoked by or generated in relation to CAV use, can technically be mined and analysed for classification of different user groups, enabling the inference of highly sensitive information about users (e.g. financial status, ethnicity, political views, personal associations, patterns of habit). This can have a great impact on the principles of **dignity**, **personal autonomy**, and also run against the principles of **non-maleficence** and **justice**.

Traditional and legally established consent procedures for personal data collection as defined in the GDPR – emphasising requirements that consent should be free, informed, explicit, and specific – may in some instances of CAV use provide weak ethical protection for users.

Alternative models or options of consent procedures need, in addition, to be explored: an ethical alternative to the "take-it-or-leave it" model of consent could be using data management systems with appropriate software tools for giving individual data subjects the means for choosing strategies for handling their data, thereby eliminating the impractical requirement for individuals to give separate consent on every issue of data use and also ensuring greater data control, traceability, and transparency.

The proper functioning of such management systems should be accompanied by appropriate auditing or certification mechanisms.

Moreover, there are potential risks of abusive exploitation of power imbalances on behalf of CAV-based mobility service providers. A CAV service user can be considered

to be in a vulnerable position, meaning they are temporarily or permanently in a position with limited or no means to choose or negotiate conditions of consent as offered by a service provider. In particular, such conditions may arise if the user is under time pressure; seeking service during off-hours; in an unsafe area; or when other options for mobility do not exist.

Develop measures to foster protection of individuals at group level
Recommendation 9

Policymakers should develop legal guidelines that protect individuals' rights at group levels (e.g. driver, pedestrian, passenger or other drivers' rights) and should outline strategies to resolve possible conflicts between data subjects that have claims over the same data (e.g. location data, computer vision data), or disputes between data subjects, data controllers and other parties (e.g. insurance companies).

As conflicts of this type are rather new, stakeholder and policy actions need to be solidly grounded in work by **researchers** and extensive public deliberation. In particular, there is a need to support and mobilise **researchers** to study the ethically, legally, and socially justifiable resolutions of data-related conflicts of interest.

Policymakers should develop new legal privacy guidelines that govern the collection, assessment and sharing of not just personal data, but also non-personal data, third party personal data, and anonymised data, if these pose a privacy risk for individuals. This is important because machine learning algorithms are able to infer personal private information about people based on non-personal, anonymised data or personal data from group profiles, over which the affected party might not have data protection rights[31]. This is a new and significant privacy risk.

Discussion of Recommendation 9

Significant data collection is necessary for the safe and efficient functioning of CAVs. The vehicle-to-vehicle (V2V), vehicle-to-infrastructure (V2I) or vehicle-to-everything (V2X) communication channels include the potential for a multitude of separate actors vying for general and specific personal data controlled by drivers, in real time or near-real time.

One particular challenge that arises in this context is the privacy protection of multiple concerned individuals (e.g. driver, pedestrian, passenger or other drivers). The use of CAVs can include sharing of rides from similar origins and destinations, between different passengers. In such situations, all passengers sharing the same vehicle, as well as pedestrians and other road users in the vehicle's vicinity could, in principle, be identified. This can occur without the awareness of those affected.

The European data protection rules require any such processing to rely on a valid legal basis and on transparent information about the processing being provided to all individuals concerned[32]. The collection of data in public spaces may conflict with individual informed consent and realistic opt-out choices for data subjects, such as pedestrians, other drivers or passengers.

Another challenge is the invasiveness and disclosive power of non-personal, third party personal data or anonymised data for individuals[33]. These types of data may allow highly privacy-invasive inferences (e.g. disability, ethnicity or sexual orientation) to be drawn. Unfortunately, these types of data are currently not governed under data protection law and thus novel privacy standards should be developed and expanded to govern all types of data that have an effect on individuals[34].

For example, computer vision captures the data of multiple data subjects at the same time (pedestrians, other drivers and road users), and thus may threaten their privacy rights as members of such ad hoc groups. This urges the question of who should be granted rights over data that concerns various people simultaneously. Even though the European strategy on Cooperative Intelligent Transport Systems (C-ITS) concludes that "data broadcast by C-ITS from vehicles will, in principle, qualify as personal data as it will relate to an identified or identifiable natural person"[35], privacy risks remain. Even if an individual chooses to opt-out or exercise other data rights, algorithms can still infer and assume personal information about them based on group profiles, non-personal data, or anonymised data for which privacy rights might not exist.

Both of these elements show that the protection of privacy rights of individuals at a group level needs to be additionally considered and further researched. In situations such as these, the challenges to the principles of autonomy and fairness are significant. To address these challenges, there is a need for further research and policy provisions regarding the scope and application of data protection schemes/models to include all data that could pose risks for individuals.

Develop transparency strategies to inform road users about data collection and associated rights
Recommendation 10

Policymakers should work with **manufacturers and deployers** to develop meaningful, standardised transparency strategies to inform road users, including pedestrians, of data collection in a CAV operating area that may, directly or indirectly, cause risks to their privacy as they travel through such areas. This includes digital and near real-time updates for road users when approaching, entering, and leaving zones where potentially privacy intrusive data collection occurs.

Such communication may occur through in-vehicle or wearable smart-device displays, audio-visual aids on roads (e.g. street signs, flashing icons, beeping sounds), or other minimally privacy-invasive communication modes with textual, visual, audio and/or haptic elements[36]. This allows the communication of privacy risks and rights to a wide and diverse audience.

As these suggested measures are new, stakeholder and policy actions and decisions need to be grounded in evidence from **researchers** and extensive public deliberation. **Researchers** should study ethically, legally, and socially justifiable resolutions of data related conflicts of interest; the design of accessible and user friendly data collection and privacy intrusion related alert terms and symbols; mechanisms to communicate these clearly and efficiently in dynamically shifting and distracted road user situations; and the type of interfaces and notification options that most efficiently support user empowerment in setting preferences, choosing routes, and negotiating terms and conditions. **Policymakers** should consider and, where relevant, apply the outcomes of this research. In all of these activities, there must be compliance with data protection law.

Discussion of Recommendation 10

The complexity of personal data and privacy related conflicts of rights and interests (see **Recommendation 9** on **individual privacy at group level)** are further exacerbated by the mobility characteristics of CAVs. Mobility-induced conflicts of interest are largely unavoidable due to the need for CAVs to move through public spaces, where intentional but non-consensual monitoring and collecting of traffic-related data, and its later use for research and development or other public ends can occur. It is also possible that a CAV ride passes in the vicinity of, or ends at, a private or otherwise non-public space in which data collection occurs on other

grounds and for other purposes, but that the CAV user similarly has not been made aware of or given the opportunity to accept or decline.

For example, because the CAV moves through physical space and may select alternative routes due to dynamic multi-variable real-time calculations without consulting the user prior to such choices, a CAV ride may result in personal data collection that the user could not anticipate from the outset, to which they have not consented, and of which they may never become aware. Thus, individual instances of CAV travel are likely to cause its passenger(s) to intersect with, move through, and thereby be subjected to a great number of geographical and personal spheres with potentially divergent and privacy intrusive personal, private (commercial), public or government related data interests, regulations and requirements.

To a certain extent, the GDPR, in particular Article 5 on purpose limitation and data minimisation, addresses problems of this sort. Furthermore, the stricter the rule stating that only necessary data should be collected is applied, the less warning will be needed. Nevertheless, the numerous exceptions in the GDPR to consent-based personal data collection and processing (Article 6) may leave scope for privacy-intrusive data practices. Furthermore, when legally enforcing the rights concerning purpose limitation and data minimisation, diverging interests between different data subjects, collectors, and stakeholders may be afforded different weight, thus effecting power imbalances at these intersections.

Therefore, developing novel and creative transparency standards (e.g. via textual, visual, audio and/or haptic elements) to communicate those privacy risks effectively and to inform about associated privacy rights (e.g. opt-out, deletion of personal data, data access rights, recourse mechanisms, alternate routes and point destinations) are paramount.

2.3 **FAIRNESS**

Prevent discriminatory service provision
Recommendation 11

The future CAV service market opens possibilities for differential provision of CAV systems, services and products that pose a risk of perpetuated and increased inequalities between individuals and groups in society[37]. This can and should be counteracted on several levels and domains.

In general, EU non-discrimination law needs to be complied with where it applies. In particular, **manufacturers and deployers** should be held responsible for designing and operating CAVs in ways that neither discriminate against individuals or groups of users, nor create or reinforce large-scale social inequalities, while taking proactive measures for furthering the ethical principle of **beneficence**. Here, it is essential that diversity is built into all aspects of the design (including CAV design team diversity)[38], operation, and business models of CAV systems and services. Such diversity should include gender, ethnicity and other socially pertinent dimensions.

To support such design and operation processes, **policymakers** need to set up institutions that continuously monitor, evaluate, and steer **manufacturers and deployers.** Relevant EU and national-level public sector institutions should proactively establish both positive (e.g. legal support, financial contributions and subsidies) and negative (e.g. legal constraints, fees, fines and taxes) regulatory means to steer such development.

Specific steering actions and oversight procedures should be developed through cross-sectoral expert engagement[39] from public and non-governmental institutions working for the protection and implementation of human rights and ethical principles in relevant areas (see also **Recommendation 16** on **obligations** and **Recommendation 17** on **culture of responsibility**).

In addition, the wider public should be actively and continuously engaged in deliberation about design and evaluation of CAV systems and services, throughout the innovation lifecycle (e.g. through in-person[40] and online[41] citizen forums). **Researchers** should focus their efforts on developing further public engagement mechanisms regarding CAVs, drawing from existing good practices from urban and transport planning.

Discussion of Recommendation 11

Some modest versions of legal user profiling and segmentation practices (including ranking and scoring) on individual or group level constitute established and legitimate forms of positive special treatment. This includes offering advantageous deals to high-frequency users, or subsidised fares to vulnerable groups. More severe, ethically – and sometimes legally – illegitimate examples of the same would be cases of negative special treatment and algorithmic bias that would constitute discriminatory acts in violation of the **principles of justice, human dignity** and **non-maleficence.**

Examples of negative special treatment in the CAV context include the risks of

different individuals or groups of users receiving: (a) unequal access to products or services; (b) discriminatory forms and quality of service such as reprioritisation, deprioritisation, and even denial of access to products and services in periods of high demand; and (c) discriminatory differential pricing strategies for CAV access. Such expressions of discriminatory service and product access in CAV contexts may further be introduced as either a conscious strategy on behalf of service and product providers and operators, or as unintended consequences of algorithmic bias or biased data in machine-learning models. Due to the opacity and black-box characteristics of certain AI and automated systems, discriminatory practices and consequences of both origins may be difficult to identify and prove[42] (see Recommendation 12 and section **2.4 Explainability** below).

If such discriminatory user segmentation and differential service provision is connected to individuals or groups of users on grounds directly or indirectly related to, for example, race, gender, social class, income, religion, sexual orientation, place of residence, citizenship, political conviction or religious belief, there will be legal, as well as ethical, violations of principles of non-discrimination and fairness. Where this happens on the basis of grounds that are prohibited under EU non-discrimination law, this is already considered illegal. In addition to this, however, proactive measures following the principle of **beneficence,** should be taken for developing means to decrease existing social inequalities and discriminatory structures, and to increase equitable and inclusive access to mobility services for all.

Audit CAV algorithms

Recommendation 12

Besides making available necessary aspects of public information for trustworthy CAV systems, investments in developing algorithmic auditing tools and resources specifically adapted to and targeting CAV systems and applications are recommended. Algorithmic auditing of AI systems in general is dedicated to primarily experimental methods for detecting and diagnosing unwanted consequences of algorithmic system designs and operations[43].

Algorithmic auditing adapted and applied to CAV systems will include development of CAV specific means and methods of field experiments, tests and evaluations, the results of which should be used for developing longer-term best practices and standards for CAV design, operation and use, and for directly counteracting any existing or emerging ethically and/or legally unwanted applications.

To steer this development, **policymakers** should establish independent bodies that include representatives of consumer or user organisations, to systematically conduct audits on specific algorithmic applications in CAV operation and use. The findings should be used in part to directly monitor and correct wrongful or unwanted designs and operations, and in part for the development of long-term standards and good practice recommendations to be communicated to manufacturers and deployers (see also chapter 3 on Responsibility).

Manufacturers and deployers should implement associated suggested measures that increase the users' awareness of potential risks of bias. Initial examples might include "warning flags", labelling remedies, compensatory information architecture solutions and diversity requirements when presenting users with options. In addition, **policymakers** should establish responsible regulatory institutions and mechanisms to enforce the adherence to the above standards. **Researchers** should continue the development of state-of-the-art CAV adapted algorithm auditing tools and practices, and work in close collaboration with other stakeholders on fostering their adoption to CAV specific innovation processes.

Discussion of Recommendation 12

Opacity[44] in CAV technology development and algorithm-based operations, service, and product offerings risks critically undermining the principles of **dignity**, **autonomy** and **justice**. In order to achieve fairness in this context, the underlying data, algorithms and models that operate on them should represent all conceivable groups of CAV users, as well as other road users as broadly, neutrally, and correctly as possible without discrimination.

Although acknowledged and addressed in other areas, the algorithmic bases for CAV systems and operations evoke unique variables that alone or combined give rise to a number of CAV specific concerns regarding black-boxed processes and biased outcomes. We know from numerous prior examples in AI that the prevalence of social biases in data sets, combined with limitations in sensing systems and automated machine learning models are highly likely to reproduce and reinforce biases, such as negatively representing women, children, or people with darker complexions. In addition, algorithm-based operations can produce false correlation assumptions and deductions, resulting in biased associations to certain objects or areas around the vehicle.

For example, bias in datasets might lead CAV algorithms and machine learning models

to associate teenagers or residents in certain geographical areas with higher risks of damage to the vehicle, thus deciding to avoid or block such users, to suggest routing of the vehicle through only some areas, or to introduce safety hazards due to misclassification in object recognition. In order to avoid unnecessary opacity and reinforcement of biases, there is a need for targeted developments in algorithm auditing and related approaches.

Identify and protect CAV relevant high-value datasets as public and open infrastructural resources
Recommendation 13

Emphasising fairness in the CAV context also demands the identification and protection of certain data as free and open resources. Particularly useful and valuable data for CAV design, operation and use, include but are not limited to: geographical data, orthographic data, satellite data, weather data, data on crash or near-crash situations including and not including CAVs, and data on mobility, traffic patterns and participants (see also Recommendation 1 on **reducing harm**, Recommendation 6 on **dilemmas,** and Recommendation 12 on **algorithm auditing**). Data of these sorts should be identified and kept free and open, insofar as they can be likened to infrastructural resources that support free innovation, competition and fair market conditions in CAV related sectors[45].

The successful identification and protection of such free, open and high-value datasets for CAV design, deployment, and use will require a number of efforts involving multiple stakeholders. First, **policymakers** should detail what sorts of data could and should be deemed high-value in the CAV context, and therefore be kept free and open. They should do this in dialogue with **manufacturers** and **deployers** of CAVs, as well as third party data stakeholders. Further, for functional open access, data formats and processing requirements will need to be harmonised and standardised in accordance with non-commercial, platform-neutral schemes and taxonomies. The successful establishment of these schemes will need to rely on extensive **research** and cooperation with open source and standardisation organisations.

Second, **policymakers** should, to the extent that this is possible, and in full compliance with personal and data privacy legislation, lead and support the establishment of high-quality high-value data infrastructure resources. Such infrastructural resources are essential for the creation of optimal conditions for analysis, response, decision-making, innovation and fair competition. This recommendation embraces the Open

Data Directive[46] within the Digital Single Market initiative of Member States, and prepares for potential CAV data related consequences of this new Open Data Directive.

Third, **policymakers** need to identify specific obligations for state, public and private actors to provide certain types of data as open data, in the interests of transparency, fair competition, financial and industrial development, and competitiveness.

Discussion of Recommendation 13

The EU's Digital Single Market initiative, with its related directives and policies, has drawn attention to the legal possibilities of regulating the balance between data protection, ownership and open access in the best interests of individuals, society, and commercial actors. Central to these ideas is the emphasis on the public good and competitive value of safeguarding certain, particularly critical, data as open access resources[47]. This initiative is founded upon the recognition that certain data have particularly high value, in that they may function as infrastructural resources for other (data-based) activities.

The most common types of data with infrastructural values are temporal and geospatial data, crucial for countless analyses and applications of other types of data (maps, time series). The Open Data Directive, by convention, positions public sector information in this realm but its latest revision opens up the detection and safeguarding of additional, previously undefined and unregulated types of data with values that make them worthy of protection from proprietary or otherwise protectionist restraints – so called "high-value data sets".

From an ethical perspective, open, free and equal access to information – in the sense of informational infrastructures and raw material for knowledge and innovation – constitutes a requirement for fair market competition, consumer empowerment, transparency and accountability in the shared interest of citizens, consumers, industry and states[48]. As such, these initiatives and re-evaluations concerning open data resources demand attention in the CAV context: what data would benefit the greatest number of CAV actors and stakeholders, and thereby be exempt from proprietary claims; in whose interests; and with what possible gains?

Setting up, enriching, and protecting high-quality, free and open data infrastructure resources would be a way to honour the principles of **fairness, beneficence** and

solidarity. There are risks from unduly keeping such data out of the public realm. Such risks are especially important if this data would help optimal analysis, decision-making, fair competition and responsible innovation; and if such data would benefit society and the planet in a more fundamental sense. Certainly, we can anticipate that stakeholders across various concerned sectors and industries will have different views, needs, and interests concerning the same vehicle-generated and/or mobility-relevant data.

While some stakeholders call for open data in the interests of fair competition, research, development, public transparency, scrutiny, and accountability, others could aim to pursue strategic partnerships for harnessing, enclosing and safeguarding proprietary data as business opportunities. Thus, there are challenges emanating from the forces that drive commercialisation and privatisation of CAV-relevant data and tools that are, or could be, seen as being a public good.

Ultimately, and in full compliance with personal data protection legislation, approaching data as a public good will not only ensure direct benefits for CAV technology development (e.g., safety improvements by independent crash investigations, optimal routing for minimising emissions), but also protect other high-value datasets which are necessary for fair competition. This will help to balance the power relationships between manufacturers and deployers of CAVs and users of the technology.

2.4 EXPLAINABILITY

Reduce opacity in algorithmic decisions
Recommendation 14

Manufacturers and deployers should develop and implement user-centred methods and interfaces for the explainability of relevant CAV applications of algorithm and/or machine learning based operational requirements and decision-making. They should ensure that the methods and vocabulary used to explain the functioning of CAV technology are transparent and cognitively accessible, the capabilities and purposes of CAV systems are openly communicated, and the outcomes traceable. This should ensure that individuals can obtain factual, intelligible explanations of the decision-making processes and justifications made by these systems, particularly in the event of individually or group-related adverse or unwanted consequences.

The fast-growing research area concerning explainable AI (XAI) and Fairness, Accountability and Transparency (FAT) in algorithmic systems should be encouraged by **policymakers,** promoted by Member States and through EU investments. This could include fostering measures for enhancing public engagement in associated R&D innovation processes (see also **Recommendation 18** on **accountability/duty to explain**).

Researchers should aim to develop explainability-enhancing technologies in relation to data collection and algorithms used for CAV decision-making. They should formulate methods for designing CAV systems which guarantee that datasets and algorithms are thoroughly documented, meaningfully transparent and explicable in a way that is adapted to the expertise of the parties concerned (e.g., individual users, policymakers, etc.) More broadly, further empirical, technical, normative/philosophical and legal research is needed to explore methods and safeguards of explainable AI that help to mitigate against biases and discrimination risks.

Discussion of Recommendation 14

AI/machine learning and other algorithm-based CAV systems and functionalities may operate as "black-boxes" that do not allow cognitive access to how they have arrived at a particular output, or what input factors or a combination of input factors have contributed to the decision-making process or outcome[49]. Counterfactual explanations are a safe and easy way to assess and investigate why a CAV or CAV related mobility service has behaved in a certain way (e.g. stopped or swerved[50] or responded differently to individuals' request for mobility services). Such explanations should, for example, be able to explain why the classifier identified (or failed to identify) an object as a dog or a bike or pedestrian, or on what grounds otherwise comparable requests for service would receive different responses from a CAV service provider.

Automated decisions are shown to have a negative tendency to replicate and reinforce old biases or generate new ones. This creates space for unintentional (as well as intentional) harmful and discriminatory practices, in violation of the principles of **dignity, autonomy** and **non-maleficence**. Discrimination can enter into AI-systems depending on how they are built, on the data with which their algorithms have been trained, how they are developed, and how they are used[51]. Training data can be biased because they represent discriminatory human perceptions and decisions, whether intentional or unintentional. Biased training data or biased samples can thereby induce discriminatory outcomes, such as illegitimately privileging one group

of users over another, or discriminating against people of certain racial backgrounds or vulnerable groups[52].

Without adequate explanation, the outputs or decisions made by such systems cannot be contested and scrutinised by affected parties, especially when consequences are erroneous or inaccurate, or have a significant negative impact on people's lives. Such explanation is therefore as important for allocating responsibility for system failures, as it is for the ability to scrutinise and question algorithmic logics that result in discriminatory actions (see also **Recommendation 18** on **accountability**). Without adequate means of access, the role of human agency and oversight is severely weakened or hindered and risks undermining the principles of **human dignity** and **autonomy**, with the consequence of critically eroding public trust in these fast-developing technologies.

The requirement of explainability, encompassing the demands for traceability, transparency, intelligibility and accountability, thereby constitutes a significant factor in building public trust and a pillar for supporting the principles of **beneficence** and **solidarity**.

Promote data, algorithmic, AI literacy and public participation

Recommendation 15

As the main stakeholders in and beneficiaries of CAV deployment, individuals and the general public need to be adequately informed and equipped with the necessary tools to exercise their rights, such as the right to privacy, and the right to actively and independently scrutinise, question, refrain from using, or negotiate CAV modes of use and services.

Policymakers have a responsibility to inform and equip the public with the capacity to claim and exercise their rights and freedoms. Policymakers should formulate explicit roles and obligations for government, public and educational institutions, to adopt strategies and measures to inform and educate the public on literacy in relation to AI, algorithms and data, and to better equip persons of all ages with abilities to act as conscious users, consumers and citizens. Furthermore, they have the responsibility to foster active public engagement and facilitate the involvement of all stakeholders for responsible innovation of CAV technology (see also **Recommendation 16** on **obligations** and **Recommendation 17** on the **culture of responsibility**).

Manufacturers and deployers have the responsibility to ensure the development and deployment of technical and non-technical methods for fostering clear and proactive communication of information to all stakeholders, facilitating training, AI literacy, and public deliberation.

Researchers should investigate both the nature of the cognitive and technical challenges that users will face as part of CAV interactions, and what cognitive and material tools (information, knowledge, skills, choices and interaction possibilities, interfaces, and modes of communication) will, in the best ways possible, help them surmount these challenges.

Discussion of Recommendation 15

Explanations of privacy related terms and agreements, as well as algorithmic operations and decisions, may pose significant cognitive and technical challenges to users. Even with the most scrupulous implementation of user-oriented explainability requirements among manufacturers, service providers and public authorities, the explanations offered are likely to pose significant challenges to individuals with various degrees of prior knowledge and skill.

Moreover, users may have unequal opportunities to acquire the necessary knowledge and competencies to understand the explanations offered. Supporting and upholding the principles of **autonomy**, **justice** and **inclusive deliberation**, requires targeted investments in many areas to promote equal opportunities for individuals to develop the necessary knowledge and competencies. This is often described in terms of data literacy, information literacy, AI literacy, and algorithm literacy[53].

Chapter 3

Responsibility

3.1 **INTRODUCTION**

It is sometimes assumed that in supporting or replacing driving tasks, CAVs will reduce the burdens and demands on human actors, and thereby reduce their responsibilities. This is not the case. Rather than reducing or eliminating human responsibility, the use of CAVs will redistribute responsibilities across the network of human individuals and organisations involved in their manufacture, deployment, and use. This may result in increased demands being placed on some actors (manufacturers and deployers, policymakers), as well as different demands being placed on others (users, passengers). New research and policies are needed to regulate this shift in demands.

Two assumptions support the recommendations that follow: (a) that there are several *types of responsibility* rather than only one; and (b) that developing guidelines and regulations for ethical design, deployment and use of CAVs is insufficient to ensure that these tasks are performed in an ethically responsible manner. Instead, we need to foster a *culture of responsibility* that involves a bottom-up approach and engagement among stakeholders. This goes beyond merely enforcing compliance with a given set of regulations.

CAVs are complex socio-technical systems with many individuals and organisations involved in their manufacture, deployment and use. Moreover, the human actors involved in these processes will interact with many automated and intelligent systems equipped with AI in doing so. It may, consequently, become unclear who is responsible for ensuring that CAVs comply with ethical and legal norms and standards, and who should be responsible for a CAV's harmful behaviour. This raises the possibility of the emergence of so-called 'responsibility gaps'. Understanding and addressing potential responsibility gaps is crucial to promote the safety of CAVs and facilitate the fair transition to CAVs, but it is important to realise that there is more than one type of responsibility gap and thus more than one desirable way to plug those gaps.

In this section, *two broad dimensions of responsibility* are identified: forward-looking (taking responsibility for things that might happen in the future) and backward-looking (being held responsible for things that happened in the past). Five specific forms of responsibility are then discussed: two forward-looking: *obligation* and *virtue*; and three backward-looking: *accountability* (duty to explain), *moral culpability* (being open to blame), and *legal liability* (facing legal consequences)[54].

It is important for all stakeholders to move beyond a narrow conception of responsibility for CAVs as involving purely backward-looking responsibility (legal liability or culpability) for accidents and mistakes, towards a broader, forward-looking conception of responsibility as a culture that sustains and shapes the development, introduction, and use of CAVs in a way that promotes societal values and human well-being.

This section provides stakeholders and the wider public with a better understanding of the different responsibilities involved in the transition to automated and connected driving, and the capacity to address potential responsibility 'gaps' that might arise in this transition. In order to achieve these goals, the extent to which each of the above five forms of responsibility should be promoted is considered, and recommendations to prevent unwanted "responsibility gaps" from arising are proposed.

The table summarises the five types of responsibilities that guide the following recommendations, their moral value, and the potential sources of associated responsibility gaps.

Type of Responsibility	Moral Function	Cost of Gap	Cause of Gap
Obligation	Avoiding future harm/ risk mitigation	Increases risks associated with CAVs; ethical concerns overlooked; undermines trust	Novelty of technology leading to a failure to specify obligations or duties of manufacturers, deployers or users.
Virtue	Developing moral agency, role-specific virtues, and enabling autonomy/mastery	Loss of moral agency, autonomy and empathy; encourages moral recklessness	Failure to build and reward a culture of responsibility within relevant organisations and among individuals.
Accountability (duty to explain)	Confirming and affirming moral agency; holding together the moral community; fostering trust and public legitimacy	Denial of moral agency; corrosion of trust; loss of public legitimacy	Failure to create mechanisms or tribunals in which actors are asked to give an explanation of their actions.
Culpability	Ensuring retribution, deterrence, or rehabilitation	Undermines the goals of retribution, deterrence, and rehabilitation	Lack of clarity as to who was responsible for an error or accident.
Liability	Compensating victims/ Corrective justice	Undermines the value of corrective justice; unfairly disadvantages victims	Incomplete legal code; exploitation of legal loopholes; failure to use existing legal rules appropriately.

3.2 **RESPONSIBILITY AS OBLIGATION**

Identify the obligations of different agents involved in CAVs

Recommendation 16

In line with the principle of <u>inclusive deliberation</u>, **policymakers** should create an adequate institutional, social and educational environment to promote dialogue between all the key stakeholders involved in the development of CAVs. The purpose of this would be to enable these actors to identify, decide and accept their respective obligations with respect to CAVs.

For example, **manufacturers and deployers** should arguably recognise their obligation to comply with <u>Recommendation 1</u> on **reducing harm** and <u>Recommendation 2</u> on **safe design**, while **policymakers** should arguably recognise their obligation to comply with <u>Recommendation 3</u> on **open road testing.** In addition to this, **policymakers** should establish fora at the national and international level, where engineers, researchers, industry representatives, other practitioners and wider public among others can deliberate in order to develop a schedule of their respective obligations in relation to CAVs, and identify clear shared moral principles that clarify what each agent is responsible for and to whom (if anyone) they are responsible.

At the same time, **policymakers** should create policies that promote, encourage, and when needed enforce respect for these obligations (see e.g. <u>Recommendation 3</u> on **open road testing**). **Policymakers** should promote the activities of **researchers** on the ethics of the design of CAVs (and connected and automated systems more generally), in order to establish this as a solid field of academic research, comparable to, for instance, medical ethics[55].

Policymakers and **manufacturers and deployers** should create a strong system of professional ethical education and accreditation for developers of CAV, and promote the introduction of engineering ethics programmes in engineering curricula. Also, **policymakers** should create an adequate educational environment to promote citizen education on the obligations of different stakeholders, including users of CAVs.

Discussion of Recommendation 16

As happens with all technological and societal changes, it can be difficult for manufacturers and deployers, policymakers, users and others to recognise their (new) obligations in relation to the development and use of CAVs. Moreover, persons and groups coming from different industrial domains (e.g. automotive engineering versus

software engineering)[56] may have a different understanding of the social responsibility of developers. This opens up an obligation gap: a situation in which we as a society recognise the necessity of taking responsibility for the ethical development of CAVs but there is a lack of certainty over what must be done and upon whom the obligation falls to mitigate the risks and promote the societal benefits associated with the deployment of CAVs. This can result in a lack of compliance with important moral obligations and an increase in the risks associated with the introduction of CAVs. Society thereby loses confidence and trust CAVs, eventually leading to missed opportunities for their beneficial use.

A combination of public deliberation, research, education and effective regulation is needed to produce awareness of the respective obligations. The first step towards the creation of a culture of responsibility is the study, deliberation and agreement on the different responsibilities of different stakeholders involved in the process. This should be done in line with the RRI approach that encourages responsiveness to society's needs and engagement with society, where the key to successful RRI is dialogue between a variety of different stakeholders involved.[57]

3.3 **RESPONSIBILITY AS VIRTUE**

Promote a culture of responsibility with respect to the obligations associated with CAVs

Recommendation 17

Knowing your obligations does not amount to being able and motivated to discharge them. Thus in addition to imposing legal obligations and supporting the identification of broader moral obligations as proposed in **Recommendation 16**, **policymakers** should promote publicly-funded interdisciplinary research centres and institutes in which **researchers** with backgrounds in engineering, law, philosophy, social sciences and other disciplines can work together to create and share good practices that promote ethical and societal responsibility. **Policymakers** should work to create an adequate institutional, social and educational environment to facilitate the implementation of such good practices. This will help realise actual changes in the corresponding values and behaviours of key stakeholders.

Policymakers and **manufacturers and deployers** should create mechanisms that reward individual people for proactively taking responsibility within organisations or professional societies responsible for the ethical design and deployment of CAVs. **Policymakers** in collaboration with **manufacturers and deployers** should make adjustments to training and licensing procedures to make users aware, able and motivated to discharge the new tasks and responsibilities that come with increased automation.

Discussion of Recommendation 17

Cultural change cannot be simply legislated[58] and a culture of responsibility cannot be easily planned and imposed via a top-down legislative or regulatory effort. These can certainly assist in fostering a culture of responsibility but they are not sufficient. Values, norms, beliefs, behaviours and practices are also shaped by cultural and educational activities and the creation of a strong sense of one's (professional) role and ethical identity.

Therefore, similar to what has happened, for instance, in aviation in relation to the creation of a culture of safety and in the medical profession in relation to a creation of a culture of care, concern and respect for patients, a culture of responsibility should be fostered in relation to the design and use of CAVs. Without the creation of such a culture, it can be difficult for developers, manufacturers, deployers, and users to actively take responsibility for the ethical design and use of CAVs. There are several reasons for this: they might not have the relevant (technical) knowledge and skills, and they may not feel able to raise ethical concerns they have about the technology, nor feel that ethical sensitivity and awareness is appreciated or rewarded within their organisations.

Manufacturers and deployers should create this culture within their companies since they are, in the end, accountable for the vehicles that are being used (see also Recommendation 18 on **accountability/duty to explain**). This is not least because users of CAVs may not have the necessary abilities, skills, and/or motivation to comply with their new role as driver/keeper of CAVs[59].

All this may give rise to a *virtue gap*: agents operating within the responsibility network are not sufficiently motivated to act according to ethical and societal values, they fear the repercussions of raising ethical concerns (e.g. losing their job), and/or they simply lack the knowledge and skills required to discharge their obligations (e.g. drivers of automated vehicles). When this happens, agents within the responsibility network are not encouraged to take responsibility for what they are doing, nor to express the traits and dispositions we would associate with responsibility. When combined with obstacles relating to the compliance with moral and legal obligations, this can undermine the value of moral autonomy, agency and empathy: actors may not be sufficiently able to see themselves as responsible or virtuous agents, and/or may not be able to perceive sufficiently the value of others' interests. This needs to be prevented.

3.4 **RESPONSIBILITY AS ACCOUNTABILITY**

Ensure accountability for the behaviour of CAVs (duty to explain)

Recommendation 18

"Accountability" is here defined as a specific form of responsibility, specifically the obligation to explain something that has happened and one's role in its occurrence. **Manufacturers and deployers** should be held accountable for creating an innovation culture in which (some of) the people developing these systems are trained and informed to have both the appropriate technical competence over the systems they create and sufficient moral awareness of their role as potential targets of moral, social or legal requests of explanation in the case of misbehaviour of the systems.

This does not mean that each action of the system should be causally traceable to an individual human action, but rather that each action of the system should be understandable by and explainable to the relevant persons or organisations via reference to the choices and actions of at least one human person along the chain of design, control, and use[60]. This does not entail that these actors ought to be also held morally, criminally or civilly culpable or liable for these actions. Accountability can be separated out from culpability and liability (see **Recommendation 19** on **attribution of culpability** and **Recommendation 20** on **compensation**).

A fair system of accountability (duty to explain) requires that: (a) relevant formal and informal fora and mechanisms of accountability are created (spaces where questions about the design and use choices about CAVs can be posed and answered); (b) different actors should be aware in advance of their respective potential need to explain and justify the operation of the system to the relevant fora, and to acknowledge failures if needs be; (c) the socio-technical system of which CAVs are a part is not too complex, opaque, or unpredictable and relevant actors have sufficient insight into its functioning and their role in it.

In addition to the obligations for data controllers following the accountability principle under EU data protection law, **policymakers** should formally arrange for the accountability of manufacturers, deployers, and users of CAVs. In addition to this, **policymakers** should provide proper information and training for the public at large as to the functioning of autonomous machines to facilitate public trust and a correct understanding of the functioning and limitations of CAVs (see also **Recommendation 15** on **AI literacy and participation**).

One specific form of accountability is the public accountability of **manufacturers and deployers,** who should ensure that the logic behind sensitive decisions made by CAVs are transparent and explainable to the public (see also **Recommendation 12** on **algorithm**

audit and <u>Recommendation 14</u> on **explainability**). In this respect, the regulatory guidance that has emerged regarding explainable automated decision-making under the GDPR should be of assistance. Moreover, the Expert Group report on Liability for AI made recommendations on how to clarify the burden of proof for such systems and suggested the creation of a logging by design obligation for producers. However, legal, philosophical and sociological **researchers** should identify the best ways to maintain accountability in increasingly automated and connected (traffic) socio-technical systems[61].

This connects with, but also goes beyond the technical research in so-called Explainable AI (see <u>Recommendation 14</u> on **algorithmic explainability** and <u>Recommendation 15</u> on **AI literacy and participation**): the present recommendation concerns also ethical and social structures within organisations and in society more generally, which encourage and incentivise the capacity and motivation of persons to give meaningful explanations in relation to CAVs.

Discussion of Recommendation 18

Consider a vehicle operated by a driver D1, with the assistance of an automated driving system, produced by the car manufacturer X, powered with digital systems developed by the company Y, possibly including some form of machine learning developed by the company Z, and enriched by data coming from different sources, including the driving experience of drivers D2, D3...D_n; vehicles in this system are in principle subject to standardisation processing done by the agency S, the traffic is regulated by the governmental agency G, drivers are trained and licensed by the agency L, and so on. This complexity and interaction may create accountability gaps: situations in which it is not clear which (if any) of the agents in the responsibility network can be held to explain the behaviour of the vehicle.

This is particularly concerning in the event that something goes wrong. This reduction in the capacity of explanation on the part of individual human actors and agencies involved in complex networks – a phenomenon similar to the so-called "problem of many hands" in complex organisations[62] – can lead to the denial of moral autonomy by the key players within the responsibility network[63] and the corrosion of trust within the broader moral community and in relation to CAVs more specifically.

3.5 **RESPONSIBILITY AS CULPABILITY**

Promote a fair system of attribution of moral and legal culpability for the behaviour of CAVs

Recommendation 19

In addition to promoting knowledge about new ethical and societal obligations coming with CAVs (see **Recommendation 16** on **obligations**), the motivation and capacity to comply with these new obligations (**Recommendation 17** on **culture of responsibility**) and the capacity of different agents to explain what happened when something goes wrong with CAVs (**Recommendation 18** on **accountability**), **policymakers** should collaborate with **researchers** and **manufacturers and deployers** to develop clear and fair criteria for assigning culpability to individual actors or organisations in the event that something goes wrong with CAVs, typically when something is damaged or someone is injured or killed in a crash due to an unjustifiable and inexcusable mistake of some human actor.

Culpability means being open not only to requests for explanation but also to stronger moral, social, and legal responses, such as blame, shame, punishment, pressure to improve one's behaviour in the future and an obligation to provide compensation or support to the persons damaged. In order to achieve this, **policymakers** and **manufacturers and deployers** should ensure that, in general: (a) different actors in the network of CAVs – manufacturers and deployers, regulators, users – are sufficiently aware of the extent to which they are the potential target of moral (and legal) reactions in the event of an unwanted outcome deriving from their mistake and misbehaviour (see **Recommendation 16** on **obligations**); (b) different actors are given a fair opportunity to acquire and develop the knowledge, capacities, skills and motivation to prevent such mistakes and misbehaviour (see **Recommendation 17** on **culture of responsibility**).

One specific challenge to culpability attribution comes from the opacity, complexity and interactivity of the technology embedded in CAVs, in particular AI, as well as the social-organisational context in which the technology is embedded (developing companies, traffic networks etc.). Therefore, in addition to (a) and (b), in order to prevent unwanted culpability gaps, **policymakers** and **manufacturers and deployers** of CAVs should also design companies, organisations, networks, and technologies in such a way that for each of these organisations, companies, networks and technologies there is always at least one human person (ideally more) along the chain of design, development, control, regulation, and use who has sufficient power, knowledge, and expertise about the system and sufficient moral awareness of his role as a potential target of moral (and legal) reaction in case of an unwanted outcome[64].

This means that **policymakers** and **manufacturers and deployers** should provide all the

relevant persons within companies, organisations and networks with the right knowledge, skills, motivation and capacity to develop this moral awareness; should create a clear chain of responsibility within their organisations; and should design and adopt technologies that fit the person's capacities (see also **Recommendation 14** on **algorithmic explainability** and **Recommendation 15** on **AI literacy and participation**).

Conversely, this also means that **policymakers, manufacturers, and deployers** should not deploy technology that does not fit the human capacity available in the organisations and/or otherwise hinders the attribution of human responsibility. To assist policymakers, manufacturers, and deployers in this regard, **researchers** in legal, philosophical, psychological, sociological, and technical disciplines should investigate the conceptual and technical conditions under which human moral and legal culpability for the behaviour of CAVs can and should be maintained, with a strong focus on technical and institutional design solutions for the preservation of human responsibility in complex socio-technical networks.

Discussion of Recommendation 19

Due to the complexity of the socio-technical system of which CAVs are a part, traditional, moral and legal criteria for the attribution of culpability to individual human agents may not easily apply to the behaviour emerging from the human interaction with intelligent systems like CAVs. This may create a culpability gap, a situation in which someone is wronged or harmed by the behaviour of the system, but no human actor may legitimately be held culpable for it[65]. For example, a developer may not have been able to reasonably foresee a particular interaction of a CAV, and the user of that CAV may not have possessed the capacity or skill to govern that dangerous behaviour.

The development of fair criteria for culpability attribution is key to ground reasonable moral and social reactions to accidents and other undesirable events involving CAVs. We concur with the opinion of many philosophers and most lawyers and laypersons that (fair) social and legal practices of attribution of culpability should also be maintained and promoted. They should not be replaced with a system of social and psychological education or therapy[66]. But this is only true insofar as these practices: represent the legitimate expression of appropriate moral sentiments by the wronged individuals and society at large[67]; reinforce the social commitments to shared norms[68]; and, possibly most importantly, they contribute to control and reduce undesirable behaviour.

Large gaps in culpability for the behaviour of technological systems may feed helplessness and moral scepticism towards the possibility of understanding and rectifying

wrongdoing. It may also fuel the desire to find a scapegoat to satisfy these feelings[69]. In line with the principles of **fairness** and **responsibility**, **policymakers, manufacturers and deployers** with the support of researchers should devise organisational structures and regulatory measures that prevent both impunity for avoidable harm (i.e. culpability gaps) and "scapegoating", (i.e. imposition of culpability on agents who were not given a fair capacity and opportunity to avoid wrongdoing).

An example of the latter would include 'pushing' culpability onto end users for a crash caused by a split-second handover of control or pushing it onto individual developers for choices ultimately taken by their employer. More studies by legal, philosophical and psychological researchers on human moral and legal culpability for the behaviour of automated (driving) systems should be promoted by policymakers.

3.6 RESPONSIBILITY AS LEGAL LIABILITY

Create fair and effective mechanisms for granting compensation to victims of crashes involving CAVs

Recommendation 20

Policymakers, in collaboration with r**esearchers** and **manufacturers and deployers,** should establish clear and fair legal rules for assigning liability in the event that something goes wrong with CAVs. This could include the creation of new insurance systems. These rules should balance the need to avoid culpability gaps (see **Recommendation 19** on **attribution of culpability**) against the requirements of corrective justice (i.e. giving fair compensation to victims without hindering innovation).

They should also ensure a fair distribution of the costs of compensation. We refer to the Expert Group report on Liability for AI and the EU reports on Liability for CAVs and on Safety and Liability for AI[70] for a broader discussion of proposals to address potential liability gaps with CAVs. **Policymakers** and **researchers** should investigate the extent to which liability rules may protect the interests of potential victims as well as the desire of companies to innovate, while at the same time being compatible with the goals of the recommendations in this chapter.

Discussion of Recommendation 20

Building on the recommendations within this chapter, further complicating practical factors may make it difficult to legitimately compel anyone to pay compensation to the victims of an incident involving CAVs. This can give rise to a liability gap and undermine the goal of corrective justice, leaving potential victims unfairly disadvantaged. The Commission Expert Group report on liability for new and emerging technologies made recommendations on how to design liability regimes regarding new technologies while recognizing it is impossible to come up with a single solution suitable for the entire spectrum of risks.

From a broader ethical perspective, it is also important to emphasise that any legal solution that grants victims compensation should also be compatible with and supported by the needs of the other forms of responsibility discussed above. For instance, an effective system of insurance, or the willingness of manufacturers to accept in advance the risks of paying compensation for accidents, should never allow or indirectly incentivise any stakeholders to avoid their responsibility to prevent undesirable outcomes, to discharge their duty to explain, or to accept blameworthiness for avoidable mistakes.

The risk of liability gaps occurring can be reduced by the fulfilment of other responsibilities discussed in this report. For example, forward-looking obligations and virtues with respect to CAVs may reduce the risks of accidents or other undesirable outcomes; and designing CAVs which allow for a clearer and fairer attribution of moral and legal culpability systems may simplify the task of attributing liability.

CONCLUSION

AND FUTURE WORK

This report provided **20 recommendations** to deal with ethical issues across three topic areas: **road safety, risk and dilemmas; data and algorithm ethics; and responsibility.** The recommendations are made actionable for three stakeholder groups: **manufacturers and deployers** (e.g. car manufacturers, suppliers, software developers and mobility service providers); **policymakers** (persons working at national, European and international agencies and institutions such as the European Commission and the EU National Ministries) and **researchers** (e.g. persons working at universities, research institutes and R&D departments).

Each recommendation contains different sub-recommendations for individual stakeholder groups. The following table recapitulates the 20 recommendations and the suggested first actions of the primary stakeholder groups to whom they apply. It should be emphasised that these are just suggested first steps. Stakeholders may and are encouraged to engage in any other actions that are in line with the recommendations.

	Primary Target group		
Suggested first actions keyed to recommendations	Manufacturers and deployers	Policymakers	Researchers
Recommendation 1 on reducing harm			
Develop new reliable benchmarks and metrics for CAV safety	✓	✓	✓
Develop new methods for road safety analysis			✓
Promote accessibility of collision / near-collision data	✓	✓	
Recommendation 2 on safe design			
Create intuitive, user-centred systems which reflect human psychological capabilities	✓		✓
Provide appropriate, clear guidance on safe CAV use	✓		
Study how driver monitoring systems can support safe CAV use in compliance with data ethics			✓
Recommendation 3 on open road testing			
Define responsible procedures to determine when a CAV is ready for open road testing	✓	✓	
Offer realistic expectations of CAV technology	✓	✓	✓
Consider traffic rule changes to manage a heterogeneous fleet of CAVs and non-CAVs		✓	

	Primary Target group		
	Manufacturers and deployers	Policymakers	Researchers
Recommendation 4 on traffic rules			
Identify contexts and appropriate actions in situations where CAVs may contravene traffic rules		✓	✓
Study the ability of CAV systems to balance norms and principles in decision-making and the explainability thereof			✓
Recommendation 5 on inequalities			
Identify road user groups that incur a disproportionate harm to exposure ratio			✓
Develop and deploy CAVs to reduce strong disparities in harm to exposure ratio	✓	✓	
Study the ethical and social acceptability of such measures			✓
Recommendation 6 on dilemmas			
Explore acceptability of CAV behaviour in dilemmas based on adherence to principles of risk distribution		✓	
Engage with the public in an inclusive process on CAV behaviour in dilemmas	✓	✓	✓
Encourage data sharing of CAV behaviour in dilemma situations	✓	✓	
Recommendation 7 on privacy and informed consent			
Seek informed consent on CAV data collection, facilitate data subjects' control over their data and prevent unauthorised access to technical infrastructure and associated data	✓		
Enforce and enhance data protection legislation to protect CAV user data		✓	✓
Recommendation 8 on user choice			
Develop agile consent-based user agreements for CAV-based services	✓	✓	✓
Leverage competition and consumer law to counteract monopolies and enable user choice for CAV services		✓	
Develop industry standards around algorithmic inference addressing ethical data sharing, transparency and business practices and protecting informational privacy and informed consent		✓	✓
Recommendation 9 on individual privacy at group level			
Develop ethical and legal guidelines that protect individuals' rights at group level		✓	✓

	Primary Target group		
	Manufacturers and deployers	Policymakers	Researchers
Outline strategies to resolve possible conflicts between data subjects' claims over same data or disputes between data subjects', data controllers and other concerned parties		✓	✓
*Recommendation 10 on **transparency of data collection***			
Develop meaningful, standardised transparency strategies to inform road users (including pedestrians) of data collection in a CAV operating area	✓	✓	
Study (and apply) resolutions of conflicts of interest related to data; the design and communication of user-friendly data and privacy terms, symbols, interfaces and notifications		✓	✓
*Recommendation 11 on **non-discriminatory service provision***			
Design and operate CAVs to neither discriminate against individuals or groups of users nor create or reinforce large-scale social inequalities	✓		
Establish institutions that continuously monitor, evaluate, and steer CAV manufacturers and deployers in relation to non-discrimination and inclusion		✓	
Develop public engagement mechanisms regarding CAVs based on urban / transport planning best practices in relation to non-discrimination and inclusion		✓	✓
*Recommendation 12 on **algorithm audit***			
Implement measures to increase users' awareness of potential risks of bias	✓	✓	
Establish independent bodies to analyse data, algorithmic and machine learning bias and deduce standards and good practice recommendations, enforced by regulation		✓	
Develop state-of-the-art CAV-specific algorithm auditing tools and practices			✓
*Recommendation 13 on **accessibility of datasets***			
Identify high-value CAV data to be kept free and open	✓	✓	✓
Establish high-quality, high-value data infrastructure resources		✓	
Identify specific obligations to assign data as open in the interests of transparency, fair competition, financial and industrial development, and competitiveness		✓	

	Primary Target group		
	Manufacturers and deployers	**Policymakers**	**Researchers**
*Recommendation 14 on **algorithmic transparency***			
Develop and implement user-centred methods and interfaces for the explainability of AI	✓		
Encourage and undertake further research on explainable AI, and fairness, accountability and transparency in algorithmic systems		✓	✓
*Recommendation 15 on **AI literacy and participation***			
Inform and equip the public with the capacity to claim and exercise their rights and freedoms in relation to AI (in the context of CAVs)		✓	
Ensure the development and deployment of methods for communication of information to all stakeholders, facilitating training, AI literacy, as well as wider public deliberation	✓	✓	
Investigate the cognitive and technical challenges users face in CAV interactions and the tools to help them surmount these challenges			✓
*Recommendation 16 **on obligations***			
Promote dialogue between key stakeholders involved in the development of CAVs to identify, decide upon and accept their respective CAV-related ethical obligations	✓	✓	
Create policies to promote, encourage, and enforce (when needed) respect for CAV obligations		✓	
Promote research on the ethics of CAVs development and use		✓	✓
Create a system of education and accreditation for CAV developers and promote ethics programmes in engineering curricula combined with citizen education on the obligations of different stakeholders, including users of CAVs	✓	✓	
*Recommendation 17 on **culture of responsibility***			
Create and share good practices that promote interdisciplinary ethical and societal responsibility		✓	✓
Update training and licencing procedures to make users aware, able and motivated to discharge the tasks and responsibilities associated with increased automation	✓	✓	
Promote a culture of responsibility in relation to the design and use of CAVs	✓	✓	

	Primary Target group		
	Manufacturers and deployers	Policymakers	Researchers
*Recommendation 18 on **accountability (duty to explain)***			
Promote an innovation culture in which people developing systems are trained and informed with appropriate technical competence and moral awareness of the full implications of their role	✓		
Arrange the formal accountability of manufacturers, deployers, and users of CAVs towards relevant actors and institutions and provide proper information for the public on CAV operations		✓	
Ensure that the logic behind sensitive decisions made by CAVs are transparent and explainable to the public	✓		
Identify the best technical and socio-psychological mechanisms to maintain accountability in increasingly automated and connected (traffic) socio-technical systems			✓
*Recommendation 19 on **attribution of culpability***			
Ensure that all CAV stakeholders are aware of the extent to which they are the potential target of moral (and legal) reactions in the event of an unwanted outcome and have fair opportunity to acquire and develop the knowledge, capacities, skills and motivation to prevent such outcomes	✓	✓	
Investigate the conceptual, psychological, and technical conditions under which human moral and legal culpability for the behaviour of CAVs can and should be maintained			✓
Devise organisational structures and regulatory measures that prevent impunity for avoidable harm (i.e. culpability gaps) and "scapegoating", (i.e. imposition of culpability on agents who were not given a fair capacity and opportunity to avoid wrongdoing)	✓	✓	✓
*Recommendation 20 on **compensation***			
Establish rules for assigning liability in the event that something goes wrong with CAVs, protecting the interests of potential victims and companies' desire to innovate	✓	✓	✓

The recommendations reflect the scientific and academic expertise of the Expert Group members, and are meant **to support not to replace the work of stakeholders** engaged in the design, development and regulation of CAVs. Furthermore, the **report is not an exhaustive list of relevant ethical considerations,** and the Expert Group recommends that further research and collaboration with stakeholders will address other considerations such as sustainability, inclusiveness, and socioeconomic impacts.

In line with the Responsible Research and Innovation approach endorsed in the report, stakeholders should further collaborate with experts in the operationalisation and translation into practice of the general principles and recommendations identified in the report, based on their professional expertise. This means that the recommendations contained in this report can and should be further discussed in **future expert and stakeholder meetings and research projects,** at the national, European and international level, on the basis of further accumulated knowledge, information and experiences. New meetings and projects at the academic, policy, and professional level should therefore be organised in the future.

At European level, the Cooperative, Connected and Automated Mobility (CCAM) Single Platform and in particular the future European Partnership on CCAM can play an important role in following up on many of the recommendations of this report. The European Partnership will bring together all the actors of the complex cross-sectoral value chain of CCAM and will develop a shared, coherent and long-term research and innovation (R&I) agenda. This agenda will address R&I actions in the area of CCAM technologies and infrastructure, but also in relation to social aspects, user acceptance and ethical issues.

These fora can ensure continuous dialogue between the entire CAV network (and the three target groups of this report) so that a common identity and culture can be fostered. They could also support in setting up the institutional, social and educational environment necessary for all stakeholders to integrate the ethical considerations laid out in this report. Finally, set at EU-level, these fora have the potential to broaden public debate and go beyond top-down legislative or regulatory efforts by engaging with users and the public on many of the issues that have been discussed here, since such involvement is needed at all levels of these further discussions.

In any case, policymakers, researchers, manufacturers and deployers of CAVs will often have to make the extra step of bringing the recommendations to their specific policy or industry domains, defining the terms and time of a feasible implementation, and identifying the specific tools needed to translate them into living policies and practices.

The recommendations are intended to contribute to the responsible acceleration of progress towards safer, cleaner and more efficient European transport systems, based on the promise of CAVs. They should provide guidance to **policymakers** in the development of regulations and topics

requiring further research; give confidence to **manufacturers and deployers** in the development of CAV technology in ways that are ethically defensible and provide direction to **researchers** towards productive areas of study associated with CAVs. Recognising that there are areas of ethical consideration not covered by the Expert Group and that the development of ethical guidance is a continuous process that interacts with concurrent social and technological developments, this report and its recommendations provide a constructive platform upon which future CAV research, development and deployment and further discourse on the associated ethical matters can flourish.

Endnotes

1. Modified and integrated from AIHLEG Guidelines for Trustworthy AI, p.36

2. Samoili, S., López Cobo, M., Gómez, E., De Prato, G., Martínez-Plumed, F., and Delipetrev, B., *AI Watch. Defining Artificial Intelligence. Towards an operational definition and taxonomy of artificial intelligence,* EUR 30117 EN, Publications Office of the European Union, Luxembourg, 2020, ISBN 978-92-76-17045-7, doi:10.2760/382730, JRC118163, p. 12

3. Task Force on Ethical Aspects of Connected and Automated Driving (Ethics Task Force), Report, Federal Ministry of Transport and Digital Infrastructure of the Federal Republic of Germany, 2018

4. European Commission, On the road to automated mobility: An EU strategy for mobility of the future. Brussels, 17.5.2018. COM(2018) 283 **https://ec.europa.eu/transport/sites/transport/files/3rd-mobility-pack/com20180283_en.pdf**

5. See table **page 6** above for a complete summary of the main documents referred to in this report.

6. Stilgoe, J. "Machine Learning, Social Learning and the Governance of Self-Driving Cars" *Social Studies of Science, 2017* DOI: 10.1177/0306312717741687; See also Morozov, E. *To Save Everything, Click Here: Technology, Solutionism, and the Urge to Fix Problems that Don't Exist,* 2014, London: Penguin.

7. Van den Hoven, J. "Value Sensitive Design and Responsible Innovation" In: Owen, R., Bessant, J. and Heintz, M. eds.). *Responsible Innovation.* 2013 Chichester, UK.

8. Lin, P. "Why Ethics Matters for Autonomous Cars" In: Maurer, M., Gerdes, J.C., Lenz, B., Winner, H. (eds.). *Autonomes Fahren, Technische, rechtliche und gesellschaftliche Aspekte.* 2015 Heidelberg.

9. Bonnefon, J-F. et al. "The social dilemma of autonomous vehicles". Science, 2016. 1573-1576.

10. We here concur with the Ethics Task Force (p. 12) and the Dutch Ministry White paper (F. Santoni de Sio, Ethics and Self-driving cars, cit., p. 7-8) that moral dilemmas though important should not assume a dominant position in the ethical debates on CAVs.

11. Calvert, S.C., Mecacci, G., van Arem, B, Santoni de Sio, F., Heikoop, D.D. and Hagenzieker, M. "Gaps in the Control of Automated Vehicles on Roads". 2020 *IEEE Intelligent Transportation Systems Magazine* 2020 DOI:10.1109/MITS.2019.2926278

12.   Goodall, N. J. "Away from trolley problems and toward risk management." *Applied Artificial Intelligence,* 2016. 810–821; Bonnefon, JF et al. "The trolley, the bull bar, and why engineers should care about the ethics of autonomous cars." *Proceedings of the IEEE,* 2019. 502–504.

13.   Kalra, N. and Paddock, S.M. "Driving to safety: How many miles of driving would it take to demonstrate autonomous vehicle reliability?" *Transportation Research Part A: Policy and Practice*, 2016. 182–193.

14.   Noy, I. et al. "Automated driving: safety blind spots." *Safety Science,* 2018. 68–78.

15.   Compare Ethics Task force report on this point, pp. 20–22, also referring to Global Forum for Road Traffic Safety (UNECE): Report of the 75th session. **http://www.unece.org/fileadmin/DAM/trans/doc/2017/wp1/ECE-TRANS-WP1-159e_new.pdf** (2017).

16.   Flemisch, F. et al. "Uncanny and unsafe valley of assistance and automation: First sketch and application to vehicle automation." *Advances in Ergonomic Design of Systems, Products and Processes.* Springer, Berlin, Heidelberg, 2017. 319–334.

17.   Zhang, B. et al. "Determinants of take-over time from automated driving: A meta-analysis of 129 studies." *Transportation Research Part F: Traffic Psychology and Behaviour,* 2019. 285–307.

18.   Flemisch, F. et al. "Cooperative control and active interfaces for vehicle assistance and automation." *FISITA World automotive congress,* 2008. 301–310.

19.   Zhang, B. et al. "Transitions to manual control from highly automated driving in non-critical truck platooning scenarios." *Transportation Research Part F: Traffic Psychology and Behaviour,* 2019. 84–97.

20.   Compare Dutch White Paper, p. 19 on "special testing zones".

21.   Nyholm, S and J. Smids, *Automated Cars Meet Human Drivers: Responsible Human-Robot Coordination and the Ethics of Mixed Traffic, Ethics and Information Technology,* 2018, https://doi.org/10.1007/s10676-018-9445-9.

22.   Santoni de Sio, F. "Killing by autonomous vehicles and the legal doctrine of necessity." *Ethical Theory and Moral Practice,* 2017. 411–429.

23.   Bonnefon J-F. et al. "The social dilemma of autonomous vehicles". Science, 2016. 1573-1576; and Awad E. et al. "The moral machine experiment". *Nature,* 2018. 59-64.

24.   Santoni de Sio, F. "Killing by autonomous vehicles and the legal doctrine of necessity". *Ethical Theory and Moral Practice,* 2017.411-429.

25.   Nyholm, S. and J. Smids, "The Ethics of Accident-Algorithms for Self-Driving Cars: an Applied Trolley Problem?" *Ethical Theory and Moral Practice,* 2016, 10.1007/s10677-016-9745-2

26.   Compare Goodall, N.J. Away from trolley problems and toward risk management, *Applied Artificial Intelligence.* 2016. 810-821

27.   AIHLEG Guidelines for Trustworthy AI; German Ethics Commission guidelines.

28.   Hildebrandt, M. "Profiling and the Rule of Law", *Identity in the Information Society,* 2009. 55-70, at p. 64.

29.   Solove, D.J., "Privacy Self-Management and the Consent Dilemma" *Harvard Law Review.* 2013. 1880 ; Schwartz, P.M. "Internet privacy and the state", *Connecticut Law Review.* 1999. 815; Van Eijk, N. et al., "Online tracking: Questioning the power of informed consent", *Information,* 2012. 57–73.

30.   Wachter, S. and B. Mittelstadt, "A Right to Reasonable Inferences: Re-Thinking Data Protection Law in the Age of Big Data and AI", *Columbia Business Law Review,* 2019. 494-620.

31.   Mittelstadt, B. "From Individual to Group Privacy in Big Data Analytics", *Philosophy and Technology* 2017. 475–494, at pp. 476 and 485.

32.   Mantelero, A. "Personal data for decisional purposes in the age of analytics: From an individual to a collective dimension of data protection", *Computer Law and Security Review,* 2016. 238–255; Bygrave, L.A., *Data Protection Law: Approaching its Rationale, Logic and Limits,* 2002. The Hague: Kluwer Law.

33.   Hildebrandt, M. and Gutwirth, S. (eds) *Profiling the European Citizen: Cross-Disciplinary Perspectives,* 2008. Dordrecht: Springer.

34. Wachter, S. "Data Protection in the Age of Big Data" Nature Electronics, 2019. 6-7

35.   European Commission, *A European strategy on Cooperative Intelligent Transport Systems, a milestone towards cooperative, connected and automated mobility,* Brussels 30.11.2016 COM(2016) 766, at 8, available at https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52016DC0766&from=EN

36.   Wachter, S. "The GDPR and the Internet of Things: A Three-Step Transparency Model", *Law, Innovation and Technology,* 2018. 266-294

37.   Blyth, P. L., Mladenovic, M. N., Nardi, B. A., Ekbia, H. R., & Su, N. M.. "Expanding the design horizon for self-driving vehicles: Distributing benefits and burdens" I*EEE Technology and Society Magazine,* 2016. 35(3) 44-49. Mladenović, M. N., Lehtinen, S., Soh, E., & Martens, K. "Emerging Urban Mobility Technologies through the Lens of Everyday Urban Aesthetics: Case of Self-Driving Vehicle". *Essays in Philosophy,* 2019. 146-170

38.   Lee, N.T. "Detecting racial bias in algorithms and machine learning", Journal of Information, Communication and Ethics in Society, 2018. 252-260.

39.   Lyons, G. *Driverless cars – a great opportunity for society? Final report of the Driverless Cars Emulsion initiative.* 2020, Bristol: University of the West of England and Mott MacDonald

40.   Cerema. *Autonomous mobility and vehicles: what are citizens' expectations for tomorrow?* 2019. Bron: Cerema, "Connaissances" series. ISBN : 978-2-37180-376-3

41.   Mladenović, M. N. "How should we drive self-driving vehicles? Anticipation and collective imagination in planning mobility futures". In Finger, M. and Audouin, M. (eds) *The governance of smart transportation systems,* 2019. Cham: Springer.

42.   Wachter, S, Mittelstadt, B. and Russell, C., "Why Fairness Cannot Be Automated: Bridging the Gap Between EU Non-Discrimination Law and AI", March 3, 2020. Available at SSRN: https://ssrn.com/abstract=3547922 or http://dx.doi.org/10.2139/ssrn.3547922

43.   *Auditing Algorithms: Adding Accountability to Automated Authority (n.d.). (Research project website)* **http://auditingalgorithms.science/?page_id=89** ; Chen, L., Mislove, A., & Wilson, C. (2015). Peeking beneath the hood of Uber. IMC 2015; Kroll, J. A., Barocas, S., Felten, E. W., Reidenberg, J. R., Robinson, D. G., & Yu, H. "Accountable algorithms". *University of Pennsylvania Law Review,* 2016. 633; Kim, P. T. "Auditing algorithms for discrimination". University of Pennsylvania Law Review. 2017, 166, 189; Sandvig, C., Hamilton, K., Karahalios, K. & Langbort, C. "Auditing algorithms: research methods for detecting discrimination on Internet platforms". *64th Annual Meeting of the International Communication Association,* May 22, 2014.

44.   Surden, H., & Williams, M. A. "Technological opacity, predictability, and self-driving cars". *Cardozo Law Review,* 2016. 121.

45.   Kitchin, R. *The Data Revolution: Big Data, Open Data, Data Infrastructures and Their Consequences.* 2014. London: Sage.

46.   Directive (EU) 2019/1024 of the European Parliament and of the Council of 20 June 2019 on open data and the re-use of public sector information, PE/28/2019/REV/1, OJ L 172, 26.6.2019, p. 56–83.

47.   European Commission. *A European Strategy for Data.* Brussels, 19.2.2020 COM(2020) 66 final. Available at: **https://ec.europa.eu/info/sites/info/files/communication-european-strategy-data-19feb2020_en.pdf**;
European Commission. *Building a European Data Economy.* Brussels, 10.1.2017 COM(2017) 9 final. Available at: **https://ec.europa.eu/digital-single-market/en/news/communication-building-european-data-economy**;
Directive (EU) 2019/1024 of the European Parliament and of the Council of 20 June 2019 on open data and the re-use of public sector information. PE/28/2019/REV/1, OJ L 172, 26.6.2019, p. 56–83. Available at: **https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1561563110433&uri=CELEX:32019L1024**
European Data Portal (EDP) (2020). *Analytical Report 15. High-value datasets: understanding the perspective of data providers.* (Last update: 20.01.2020). European Commission. Available at: **https://www.europeandataportal.eu/sites/default/files/analytical_report_15_high_value_datasets.pdf**

48.   Hesse, C. (2002). The Rise of Intellectual Property, 700 B.C.-A.D. 2000: An Idea in the Balance. *Daedalus,* 2002. 26-45.

49.   Mittelstadt, B., Russell, C. and Wachter, S. "Explaining Explanations in AI". In *Proceedings of the Conference on Fairness, Accountability, and Transparency (FAT* '19).* 2019. New York, NY: US. DOI:**https://doi.org/10.1145/3287560.3287574**; Miller, T. "Explanation in Artificial Intelligence: Insights from the Social Sciences" Artificial Intelligence, 2019. 1-38.

50.   Wachter, S. Mittelstadt, B. and Russell, C. "Counterfactual Explanations without Opening the Black Box: Automated Decisions and the GDPR" *Harvard Journal of Law & Technology,* 2018. 841-887.

51.   Barocas, S. and Selbst , A.D., "Big Data's disparate impact", *California Law Review,* 2016. 671-732.

52.   Lee, N.T. (2018), "Detecting racial bias in algorithms and machine learning", *Journal of Information, Communication and Ethics in Society,* 2018. 252-260.

53.   See, for example, Head, A. J., Fister, B. & MacMillan, M. *Information literacy in the age of algorithms: Student experiences with news and information, and the need for change,* 2020. The Algo Study Report. Project Information Literacy. **https://www.projectinfolit.org/uploads/2/7/5/4/27541717/algoreport.pdf**

54.   Van de Poel, I. and Sand, M.. "Varieties of responsibility: two problems of responsible innovation", *Synthese,* 2018. DOI: 10.1007/S11229-018-0195107; Santoni de Sio, F. and Mecacci, G. "Four responsibility gaps with autonomous systems: why they matter and how to address them", forthcoming. Existing policy documents and reports have mainly focused on accountability (AIHLEG Guidelines for Trustworthy AI, p19-20), and liability (Expert Group report on Liability for AI).

55.   Véliz, C. "Three things digital ethics can learn from medical ethics". *Nature Electronics,* 2019. 316–318 **https://doi.org/10.1038/s41928-019-0294-2**

56.   Stilgoe, J. "Machine Learning, Social Learning and the Governance of Self-Driving Cars". *Social Studies of Science,* 2018. 25–56. DOI: 10.1177/0306312717741687.

57.   McBride, N. and B. Stahl, "Developing responsible research and innovation for robotics". *IEEE International Symposium on Ethics in Science, Technology and Engineering,* 2014. Chicago.

58.   Schiebinger, L. *Gendered Innovations in Science and Engineering.* Stanford University Press, 2008, p. 13.

59.   Heikoop, D. et al. "Human behaviour with automated driving systems: a quantitative framework for meaningful human control", *Theoretical Issues in Ergonomics Science,* 2019. 711-730.

60.   Mecacci G. and F. Santoni de Sio. "Meaningful Human Control as Reason-responsiveness: the case of dual-mode vehicles". *Ethics and Information Technology,* 2020. 103-115. **https://doi.org/10.1007/s10676-019-09519-w**

61.   One suggestion in this direction is to promote a "logging by design approach" (Expert Group report on Liability for AI, §20-23).

62.   Bovens, M. The Quest for Responsibility, 1998. Cambridge, UK: Cambridge University Press; and van de Poel, I., Royakkers, L. and Zwart, S.,
Moral Responsibility and the Problem of Many Hands, 2015. London: Routledge.

63.   Danaher, J., The rise of the robots and the crisis of moral patiency. *AI & Society,* 2019. 129-136.

64.   Santoni de Sio, F. and Van den Hoven, J. "Meaningful Human Control Over Autonomous Systems: A Philosophical Account". *Frontiers in Robotics and AI,* 2018. **https://doi.org/10.3389/frobt.2018.00015**.

65.   Matthias, A. "The responsibility gap: ascribing responsibility for the actions of learning automata". *Ethics and Information Technology,* 2004. 175-183.

66.   Morse, S. J. "Moral and Legal Responsibility and the New Neuroscience". In J. Illes (Ed.), *Neuroethics in the 21st Century: Defining the Issues in Theory, Practice, and Policy,* 2006. Oxford University Press.

67.   Strawson, P. F. "Freedom and Resentment". *Proceedings of the British Academy,* 1962.

68.   Sie, M. *Justifying Blame: Why Free Will Matters and Why it Does Not.* 2005, Rodopi.

69.   Danaher, J. "Robots, Law, and the Retribution gap", *Ethics and Information Technology,* 2016. 299-309.

70.   Commission Report on the safety and liability implications of Artificial Intelligence, the Internet of Things and robotics. **https://ec.europa.eu/info/sites/info/files/report-safety-liability-artificial-intelligence-feb2020_en_1.pdf**

Connected and Automated Vehicles (CAVs) are emerging as a new technology and new form of mobility in Europe. Expectations are high: these vehicles can bring down the number of road fatalities near zero; increase accessibility of mobility services; and help to reduce harmful emissions from transport by making traffic more efficient.

However, technological progress alone will not be able to bring about the full potential of CAVs. The timely and systematic integration of ethical and societal considerations, from inception to use, will be essential to ensure their ethical and positive impact.
With its strategy on Connected and Automated Mobility, the European Commission aims to make Europe a world leader in the development and deployment of CAVs.

To tackle ethical issues, the Commission formed in 2019 an independent Expert Group to provide practical support to researchers, policymakers, manufacturers and developers in the safe and responsible transition towards connected and automated mobility.

The 20 recommendations presented in this report consider ethical principles and shared moral values as stimuli in shaping CAV innovation, rather than be perceived as an obstruction to their progress.

Research and Innovation policy