HHS Public Access
Author manuscript
AI Ethics. Author manuscript; available in PMC 2024 May 20.

Published in final edited form as:
AI Ethics. 2024 May ; 4(2): 403–418. doi:10.1007/s43681-023-00277-6.

A precautionary approach to autonomous vehicles

David B. Resnik[1], **Suzanne L. Andrews**[2]

[1]National Institute of Environmental Health Sciences, National Institutes of Health, 111 TW Alexander Drive, Mail Drop E1-06, PO Box 12233, Research Triangle Park, NC 27709, USA

[2]Master of Science in Bioethics Student, Columbia University, 203 Lewisohn Hall, 2970 Broadway, Mail Code 4119, New York, NY 10027-6902, USA

Abstract

In this article, we defend an approach to autonomous vehicle ethics and policy based on the precautionary principle. We argue that a precautionary approach is warranted, given the significant scientific and moral uncertainties related to autonomous vehicles, especially higher-level ones. While higher-level autonomous vehicles may offer many important benefits to society, they also pose significant risks, which are not fully understood at this juncture. Risk management strategies traditionally used by government officials to make decisions about new technologies cannot be applied to higher-level autonomous vehicles because these strategies require accurate and reliable probability estimates concerning the outcomes of different policy options and extensive agreement about values, which are not currently available for autonomous vehicles. Although we describe our approach as precautionary, that does not mean that we are opposed to autonomous vehicle development and deployment, because autonomous vehicles offer benefits that should be pursued. The optimal approach to managing the risks of autonomous vehicles is to take reasonable precautions; that is, to adopt policies that attempt to deal with serious risks in a responsible way without depriving society of important benefits.

Keywords

Autonomous vehicles; Precautionary principle; Ethics; Policy; Artificial intelligence; Benefits; Risks; Public health; Environment

1 Introduction

Autonomous vehicles (AVs) are being promoted as a revolutionary advancement in transportation technology that could substantially improve motor vehicle safety and access to social and economic activities for individuals with disabilities [1–4]. While it may take decades before highly automated AVs are deployed widely, that day may come eventually, given the steady advancement of the computing technologies used to control these vehicles and the economic interests of companies who stand to profit from them [1–4]. Additionally,

✉David B. Resnik, resnikd@niehs.nih.gov.

Conflict of interest The authors have no competing interests to declare.

governments are likely to promote AV development for public health, military, intelligence, law enforcement, or other purposes [5, 6].

Private companies have spent or are planning to spend over $100 billion on AV research and development (R&D) [7–10]. In the US, Tesla says it has sold over 800,000 vehicles with advanced autopilot and self-driving functions [11]. Although Tesla claims that its vehicle can drive itself, the human driver must monitor the vehicle's performance and be prepared to take over driving functions, so it is not fully autonomous [12]. In China, Baidu began operating AV "robotaxis" in Beijing and other cities in 2022 [13]. Other companies vying for leadership in AV technology include Nvidia, Waymo (a subsidiary of Alphabet), Zoox, and Argo. ai [14]. Conventional motor vehicle manufacturers, such as Ford, Chrysler, General Motors, Honda, Volkswagen, and Mercedes-Benz have also begun to develop their own AVs or are partnering with AV companies. Most manufacturers are focusing their R&D investments on electric AVs because they have lower fuel costs and fewer environmental impacts than non-electric AVs [15].

Federal and state governments have also invested in AV research and have begun regulating AVs [16, 17]. The National Highway Transportation Safety Administration (NHTSA) spends about $7 million annually on AV safety research. In 2019, the US Department of Transportation (2019) awarded $60 million to winners of AV research grants. The Self-Drive Act, which is under consideration by the US Congress, would establish a federal regulatory framework for AVs and spend $100 million on AV safety research [18–21]. Led by Florida and California, 34 US states now permit testing or deployment of AVs [22].

However, not everyone is enthusiastic about AVs. Consumer advocate Ralph Nader, who mounted a campaign to remove Chevrolet's Corvair from the market in the 1960s, has voiced significant concerns about AV safety and has called on Tesla to remove its "self-driving" software from its vehicles [23]. The NHTSA has been investigating the safety of Telsa's "self-driving" vehicles since August 2021 [24]. Conventional motor vehicle manufacturers, such as Ford and Volkswagen, have scaled back their AV development plans and advised investors that the technology still may be many years down the road [25, 26]. Also, consumers continue to be concerned about AV safety and performance. The American Automobile Association (AAA), a US-based automobile service organization, has polled its members several times about AVs in the last decade. A recent AAA survey found that 85% of respondents were unsure or fearful of self-driving technologies. The survey also found that respondents were much more interested in driver assistance technologies and safety improvements than in self-driving capabilities [27].

Development and deployment of AVs, especially highly autonomous or fully autonomous vehicles, raise significant ethical and policy issues relating to safety, accessibility, and social and economic impacts [2, 28–33]. Moreover, these issues are difficult to resolve because there is considerable scientific uncertainty on matters such as expected progress in AV development and the socioeconomic effects of AV deployment. AV designers and developers cannot accurately or reliably predict when (or if) highly or fully autonomous vehicles will be safer than human-driven vehicles or how widespread use of AVs will affect the labor force [1, 27]. Two general responses to this uncertainty have emerged. Optimists are trying to push this technology forward as fast as they can, while pessimists are trying to put on the brakes [1, 30–33].

Given all this uncertainty and disagreement, how should society make ethical and policy decisions concerning AVs? In other areas of scientific and technological development, the precautionary principle (PP) has been used as a framework for making ethical and policy decisions concerning risks when decision-makers face scientific uncertainty [34].[1]

The PP roughly states that one should take reasonable precautions to avoid, minimize, or mitigate serious risks. One of the main arguments for using the PP is that other methods traditionally used by governments in policymaking, such as risk assessment and cost–benefit analysis, do not apply to situations when scientific evidence is uncertain, as is often the case when one is dealing with issues relating to scientific and technological development [34]. Critics of the PP have argued, however, that the principle impedes progress in science and technology and should therefore not be used to guide public policy [35].

In this article, we will defend a precautionary approach to AV ethics and policy. We will argue that a precautionary approach is warranted, given the significant scientific uncertainties related to AVs, especially higher-level ones. To support our argument, we will defend a version of the PP that balances scientific and technological progress with other important moral concerns, such as protection of public health and the environment. Although we describe our approach as precautionary, this does not mean that we are opposed to AV development and deployment, because AVs offer potential benefits that should be pursued. While some might argue that the best way to deal with the risks of AVs is to not develop this technology, we will argue that a better approach is to adopt policies that attempt to deal with serious risks in a responsible way without depriving society of important benefits. After defending our precautionary approach to AVs, we will offer some recommendations based on the PP.

2 What are autonomous vehicles?

An autonomous vehicle is a transportation machine in which at least some movement functions are controlled by the vehicle as opposed to a human agent [31].[2] The NHTSA distinguishes between six levels of motor vehicle automation ranging from Level 0 (no automation) and Level 1 (Driver Assistance) to Level 4 (high automation) and Level 5 (Full Automation) [36][3] (see Table 1). In the US, Levels 1, 2, and 3 are commercially available. Levels 4 and 5 are being tested but are not yet commercially available.

Driving is a highly complex activity involving perception of and acquisition of information about the environment; management of numerous variables, such as the position and velocity of the vehicle, other vehicles, and pedestrians; driving conditions, such as weather

[1]Since scientific knowledge depends on evidence from observation, experimentations, and other inductive methods, it is never absolutely certain. By "scientific certainty" we therefore mean that scientists have enough evidence to form a rational consensus concerning a hypothesis or theory [34].
[2]In this paper, we will focus on autonomous motor vehicles, such as cars and trucks, but we recognize that autonomous vehicles encompass all kinds of transportation, including trains, airplanes, drones and boats.
[3]The NHTSA scale is similar to the scales developed by Wiseman and the Society of Automotive Engineers (SAE) [37].

and lighting, the location of fixed objects, such as telephone poles; hazards, such as road damage or construction; and road signs and traffic laws. AVs use computer programs and various sensors and communication devices to manage these variables and perform driving functions, such as steering, accelerating, decelerating, and navigating [2]. AV sensory devices, such as cameras, radar, and lidar (i.e., laser-based sensing) gather information about the environment. AVs also use the global navigation system to acquire information about the vehicle's position and velocity from satellites. All this information is electronically relayed to an autonomous driving computing platform, which is located on internet servers outside the AV. The platform uses artificial intelligence (AI), machine-learning (ML) software, such as neural networks, to process the information and make driving decisions, which are communicated to the AV. The information gathered by AVs is recorded and processed by the platform, which enables it to learn from the environment how to improve driving over time [38–40]. AVs also have onboard computers that are used to help control the vehicle and interact with human passengers. However, the onboard computers do not currently have enough computing power or the appropriate software to autonomously control the vehicle.

Although enthusiasm for AVs continues to run high among many highway safety officials and automotive engineers and investors, AVs still face some significant technical challenges, and some AV advocates are revising their earlier predictions concerning the advent of widespread deployment of AVs [10]. One of the key technical challenges is developing computing systems that can manage the complexity and unpredictability involved in driving [1, 2, 40]. Although AV AI/ML systems continue to become better at handling complexity and predicting the behaviors of objects in the driving environment, these systems are susceptible to biases related to the data they have been trained on. For example, an AV AI/ML system may not recognize road signs if they differ from the signs it learned to recognize when it was trained [41]. This problem makes AVs vulnerable to accidents caused by road sign tampering or graffiti [42, 43]. Also, an AV AI/ML system trained in relatively simple and predictable driving environments, such as rural roads, may not be able to adequately respond to environments that are complex and unpredictable, such as city streets [44]. Other important technical challenges include developing sensors that can continue to provide useful information in adverse driving conditions, such as rain, fog, or darkness; providing uninterrupted internet access for AVs, and protecting AVs from hacking and other cybersecurity threats [40, 41, 44–47].

3 Autonomous vehicle safety

One of the most important potential benefits of AVs (see discussion below) is that they could become safer than human-driven vehicles and reduce traffic-related deaths and injuries. Each year, about 1.35 million people die from automobile accidents around the world [46]. In the US, there were nearly 43,000 automotive fatalities, a 16-year high [47]. Fatal and non-fatal automobile injuries cost the global economy about $1.8 trillion each year [46]. 90% of traffic accidents are due to human error [45]. Given these grim numbers, technologies that can improve automotive safety would be a welcome development. AVs represent a huge potential advance in automotive safety because they reduce human error by potentially eliminating the human driver. Unlike human beings, AVs do not drive while inebriated, fatigued, or distracted.

While there have been highly publicized reports of AVs killing pedestrians and motorcyclists, high-quality (i.e., reliable, accurate, complete, and relevant) data on AV safety are not currently available [48–53]. The most important problem with available AV data is lack of context. While manufacturers are required to report data on AV accidents and AV-related deaths to the NHTSA, they are not required to report data that systematically compares AVs to human-driven vehicles. For example, according to the NHTSA, Level 3 to 5 AVs were involved in 130 crashes from July 2021 to May 2022, which resulted in no deaths, one serious injury, three moderate injuries, and twelve minor injuries [53]. However, during that same period Level 1 to 2 AVs were involved in 367 accidents that resulted in 5 deaths. Most of these vehicles were Teslas with driver assist technologies [52]. The data suggest that Level 3–5 AVs are safer than Level 1–2 AVs, but that makes little sense, given what is known about the state of AV technology. These numbers are difficult to interpret because there is no direct comparison with data from human-driven vehicles. To get useful comparative data, one would need to know how many miles AVs were driven. For example, in the US in 2020 there were 1.46 deaths per 100 million miles driven. If Tesla Level 1–2 AVs drove 200 million miles or less during the reporting period, this would indicate that they were safer than the average human-driven vehicle, but there is no information about how many miles these vehicles were driven. Another interpretative problem is that there is no information about the conditions under which these vehicles were driven. For example, 130 Level 3–5 AV crashes from July 2021 to May 2022 does not sound very dangerous, but it might if the vehicles were driven in highly safe, controlled conditions. Finally, the available data do not distinguish between accidents due to malfunctioning AVs or human errors, which is important to know for understanding the safety of Level 1 and 2 AVs, which still rely on humans for driving functions.

Although AV safety data are far from what one would need to make responsible decisions about AV policy and regulation, the NHTSA has determined that Level 4–5 AVs are currently not safer than human-driven vehicles and should not be deployed widely [36]. It is difficult to predict when Level 4–5 AVs will be safer than human-driven vehicles, given the complexity of driving and difficulties with collecting and assessing AV safety data. AV developers and manufacturers who made optimistic predictions continue to scale-back their expectations and have begun expressing skepticism [1, 10, 30].

4 Benefits and risks of autonomous vehicles

AVs have many different potential public health, socioeconomic, and environmental benefits and risks (see Table 2).

The most significant potential public health benefit of AVs is that they could reduce traffic accidents, deaths, and injuries. Indeed, some commentators and organizations view AVs as a type of public health intervention [54, 55]. However, as noted earlier, Level 4–5 AVs are currently not safer than human-driven vehicles and could dramatically increase traffic accidents, deaths, and injuries if they were widely deployed at present. Thus, when it comes to public health, AVs have significant potential benefits but also significant risks, and it is not possible at the present time to accurately compare and assess these benefits and risks, given the uncertainties related to AV technology.

Another important potential public health and socioeconomic benefit of Level 4–5 AVs is that they could improve access to transportation for people who are disabled, elderly, or socioeconomically disadvantaged [2, 56, 57]. Reliable and safe transportation is essential for health care, employment, education, socialization, recreation, and many other important activities. Thus, Level 4–5 AVs could improve health and quality of life for many people. However, today's AVs are not likely to have these beneficial effects since they are so expensive that they are available to only people in higher-income brackets. Currently, Level 3 AVs are too expensive for most people [58]. Although AV prices will probably decline due to improvements in technology, economies of scale, and market competition, one cannot predict when Level 4–5 AVs will be affordable for most people in middle- and lower-income brackets. Using AVs as a form of public transportation or taxi service may help address afford-ability problems because governments or private companies would be purchasing AVs. Even so, AV prices could still be incorporated into the cost of transportation, unless AVs are subsidized by the government. Thus, while AVs may help to reduce socioeconomic inequalities in society by increasing access to transportation, they may not have this effect until they become more affordable.

AVs may also have environmental benefits [59]. As mentioned previously, most AV manufacturers are focusing on electric vehicles (EVs), which do not produce gases, such as sulfur dioxide or carbon dioxide, that pollute the air or contribute to global warming. However, EVs are not environmentally harmless because the electricity used to charge their batteries may be produced by power plants that burn coal or other fossil fuels. Also, EV batteries contain lithium, cobalt and other toxic metals that must be disposed of properly to avoid harming people or the environment, and considerable energy must be used to obtain the metals used to make AV batteries [60]. Another factor to consider is that environmental benefits of AVs may be offset by increased utilization of transportation, since people who do not like to drive may use AVs to go somewhere that they would not have gone otherwise [2]. For example, someone who hates waiting in traffic may take an AV because they know they can pass the time by read or watching television [61]. Thus, while electric AVs may have significant environmental benefits, their impact is far from certain at this point in time.

AVs are likely to have significant economic impacts, both good and bad. While AVs may eliminate thousands of jobs for professional drivers, they may also create jobs for people who service AVs and support and manage AV fleets, so the precise economic effects of AVs are difficult to estimate at this time [62, 63]. For many years, economists and social commentators have warned that technology can contribute to unemployment by eliminating jobs for workers with modest skills and education [64]. For example, robots have eliminated manufacturing jobs and automated tellers have eliminated banking jobs. However, technology can also create new jobs. For example, while personal computers have reduced the secretarial workforce, they have also created jobs for software developers and information technology support people. While unemployment can have traumatic and tragic impacts on individuals, a certain amount of job destruction, creation, and transition (or "churning") is essential to economic growth and development [65]. Providing funds for education and training of displaced workers can help relieve some of the economic pain of churning.

Finally, AVs may contribute to the deskilling of the human population by eliminating cognitive tasks involved in driving [66]. Driving, as discussed earlier, is a difficult and complex cognitive activity that requires the driver to perceive the environment, assess information, manage many different variables, and make decisions with consequences for human life and safety. Some people may partially lose these skills if they are not periodically challenged by the activity of driving. Studies have shown that the use of automation in flying has caused airline pilots' skills to deteriorate [67]. The use of global positioning system (GPS) applications in cars and smartphones has reduced many people's ability to navigate without the use of this technology or even read maps [68, 69]. While this is clearly a significant risk of widespread use of AVs, it is not clear how much AVs are likely to contribute to the cognitive decline of the human population, since various technologies are already contributing to this decline and would be likely to do so even if AVs were never developed [70]. Also, while human beings have clearly lost some cognitive skills as a result of automation, this does not mean they are becoming less intelligent. Various studies have shown that intelligence quotient (IQ) test scores increased in many different countries during the twentieth century [71]. It may be the case that while human beings have lost some cognitive skills, they have gained others. However, the problem of the effect of automation on human cognition bears further investigation.

To briefly summarize the previous three sections of our paper, it seems clear that AVs have significant benefits, risks and uncertainties and that one cannot currently make reliable predictions concerning when (or if) Level 4–5 AVs will be safer than human-driven vehicles, or how AVs will impact society and the economy. AVs predictions are therefore fraught with peril. The lack of high-quality, publicly accessible data concerning AV safety may be viewed as a public policy scandal, given the potentially transformative and wide-ranging impacts of this new technology. By the time a new drug comes to market, regulatory agencies have considerable data concerning its benefits and risks from pre-clinical studies and three phases of clinical trials [34]. The same cannot be said for AVs, which are likely to have far greater impacts on society than any drug on the market today.

5 Autonomous vehicle ethics and policy

Most of the published articles on the ethics of AVs have focused on dilemmas that can occur during automated driving and fewer have addressed the broader ethical and social issues raised by AVs [72–82]. AV ethical dilemmas have drawn considerable interest among computer scientists, philosophers, and psychologists because they have much in common with "Trolley Problem" scenarios [2, 77, 79]. The Trolley Problem is a philosophical thought experiment in which you imagine that you find yourself in a situation in which a runaway trolley is heading toward five people that a criminal has tied up on the tracks. The criminal has also tied up one person on a sidetrack. You can divert the trolley to the sidetrack by pulling the lever. There is not sufficient time to untie the five people. The ethical dilemma is whether you should pull the lever to save the five people but kill one person or stand by and do nothing [78]. A dilemma like the Trolley Problem could arise if a pedestrian suddenly runs in front of a Level 4–5 AV and the AI/ML system that controls the vehicle must choose between hitting the pedestrian or turning away quickly and crashing the vehicle, possibly killing the passenger(s). The AI/ML system would need to make a moral choice of how many lives to save or kill [79, 80].

While the literature on how AVs can and should respond to driving dilemmas is rich, thought-provoking, and technically sophisticated, this paper will not focus on these micro-level ethics problems. Instead, the focus will be on macro-level ethical/social/policy questions related to AVs, such as:

- Should AVs be developed and deployed on roads? If so, how, when, and why?
- If society decides to permit AVs to be developed and deployed, how should the benefits and risks of AVs be balanced and managed?
- How can society protect human rights with respect to AVs?
- How can society promote justice and equity in AV development and deployment?

While some articles, reports, and books in the scholarly literature have addressed macro-level AV ethical and policy issues like those listed above [83–86], none use the PP to guide ethical/policy analysis. As shall be argued further below, the PP can lend some useful insights into AV ethics/policy issues, because the PP applies to situations involving decision-making about serious risks when one is faced with scientific uncertainty, which is clearly the case when it comes to AVs and other applications of AI [87]. However, many scientists, engineers, entrepreneurs, and policy analysts in the AI field view the PP as antithetical to their goals and objectives [88–90]. It is therefore important to defend and interpret the PP and address some objections to it before applying it to AV issues.

6 The precautionary principle

The PP was developed in the 1970s as an alternative to traditional strategies used by governments to make decisions about managing risks. Most government agencies make environmental and public health policy decisions, such as decisions related to the regulation of drugs, pesticides, or pollutants, based on rigorous assessment of scientific evidence related to risks. The risks associated with different policy options, such as approving (or not approving) a new drug or medical device, are considered “reasonable” or “acceptable” when evidence indicates that the benefits outweigh the risks. This approach has an important shortcoming, however: it does not provide policymakers with a coherent way of addressing risks when evidence is incomplete or inconclusive and decision outcomes are uncertain [34]. The only advice the risk assessment approach can offer is to obtain more scientific evidence to reduce uncertainty, but actions may need to be taken in the meantime to prevent or minimize serious harms. Waiting until there is more evidence related to a decision may be too late [34, 91–95].

Since the 1990s, the PP has grown steadily in influence. It has been incorporated into several international treaties, including the United Nations’ Rio Declaration on Environment and Development, the Cartagena Protocol on Biosafety, and the European Commission’s guidance on chemical regulation [34, 96, 97]. Also, scientists and scholars have published articles and books that apply the PP to various environmental and public health topics, including chemical regulation, medical decision-making, electromagnetic radiation, geoengineering, nanotechnology, dual use research in the biosciences, and climate change [34, 93–104].

Despite, or perhaps because of, its influence, the PP has generated considerable controversy. Critics of the principle argue that it is poorly defined, renders unclear or even inconsistent guidance, and is fundamentally opposed to scientific and technological progress [35, 91, 105–107]. Several writers have recently argued that the PP undermines progress in the development of AI and its applications in industry, transportation, and business [88–90]. PP proponents have developed versions of the principle designed to meet these objections [34, 94–104].

While it is important for readers of this article to understand that the PP has critics and supporters, it is not our aim in this article to adjudicate debates about the merits of the PP or evaluate different versions of it. However, we will respond to the charge that the PP undermines scientific and technological progress, because this is a serious critique that needs to be addressed if one is to use the PP as a guide for AV ethics and policy.

To address this critique, we will first point out that many different versions of the PP have appeared in the scientific and scholarly literature, which differ markedly in their degree of risk-aversiveness [34]. Highly risk-aversive versions of the PP: (1) focus on the necessity of avoiding serious risks; (2) do not adequately consider the costs of risk-avoidance, such as foregoing benefits or opportunities; and (3) address risks that are highly speculative and unrealistic, i.e., worst-case scenarios [34]. Less risk-aversive versions (1) recognize that there are often many ways of addressing risks, including avoidance but also minimization and mitigation; (2) consider the costs of risk-avoidance; and (3) do not dwell on speculative risks [34].[4] Putting all this together, the version of the PP that will be used in this paper can be stated as follows:

> In the absence of the degree of scientific evidence required to establish accurate and precise probabilities for outcomes related to a decision, take reasonable precautionary measures to avoid, minimize, or mitigate plausible and serious harms [34, p. 91].

The above definition includes several different ideas that require further elaboration. We will now explain these ideas and show how they apply to AVs.

7 The precautionary principle and autonomous vehicles

The first point that requires further elaboration is that the PP applies to decisions where there is insufficient scientific evidence to establish accurate and precise probabilities related to a decision; that is, the outcomes of the decision are uncertain. Clearly, this is the case with respect to AV policy, since there is not sufficient scientific evidence at present to say whether (or when) Level 4–5 AVs will be safer than human-driven vehicles. As AVs continue to

[4]Minimization and mitigation are different concepts. Minimization involves taking steps to reduce the probability that a harm will occur, whereas mitigation involves taking steps to minimize the damage from the harm if (or when) it occurs. For example, keeping a car's tires properly inflated minimizes the risk of having a flat tire and carrying a spare tire mitigates the inconvenience that can result from a flat tire.

improve, they may become safer than human-driven vehicles at some point in time, but one cannot say with confidence when or whether this will happen.

The second point is that harms must be plausible. There is little point in wasting time and mental effort worrying about possible outcomes that lack a solid basis in current science, technology, and engineering. Dwelling on these worst case ("sky is falling") scenarios leads one to succumb to fear and unreasonable caution [34]. When it comes to AVs, one can imagine many bad things that could happen: for example, AVs could play a key role in enabling artificially intelligent machines to take over the world and exterminate the human race [76]. While outcomes like these are worth contemplating in science-fiction novels and movies, the PP focuses on outcomes that are grounded in reality, such as traffic deaths, hacking, job losses, and growing human dependence on automation.

The third point is that there are usually a variety of measures (e.g., strategies, policies, or procedures) that can be used to deal with harms, including various forms of avoidance, minimization, or mitigation.[5] With respect to AV policy, harms could be avoided by completely banning higher level (Level 4–5) AVs. If Level 4–5 AVs are kept off the road, they will never produce traffic deaths. While some people may favor this highly risk-averse policy, it is neither reasonable nor realistic, as will be discussed further below. Rather, the better course of action is to pursue policies that minimize or mitigate harms. These policies recognize that some harms will occur but treat them as unfortunate outcomes that must be accepted if we are to obtain the benefits of AVs.

The fourth point is that precautionary measures must be reasonable. What does it mean for something or someone to be reasonable? The answer to this question is subject to debate and may vary according to the situation one is in [34]. For example, acting reasonably in a courtroom may be very different from acting reasonably at a graduation party, because behavior in a courtroom is strictly controlled while behavior at a party generally is not. In this paper, we will focus on reasonableness with respect to choices concerning risks. According to several PP proponents, a precautionary measure (i.e., a choice or policy) is reasonable if it is judged to be the morally best choice, all things considered [34, 94, 95, 98]. Some key criteria for assessing reasonableness include: (1) proportionality; (2) fairness; (3) epistemic responsibility; and (4) consistency.

Proportionality involves the balancing or weighing of benefits and risks. Because the PP applies to situations where benefits and risks cannot be quantified, due to insufficient scientific evidence concerning those risks or moral agreement about how to evaluate them, risk/benefit assessment is qualitative not quantitative.[6] One must decide whether the risks are worth taking, given the benefits. The bigger the risk, the bigger the potential benefit must be to justify the risk [34, 108, 109]. With respect to AVs, societies must decide whether the risks of AVs, such as traffic deaths due to AVs, hacking threats or job losses, are worth the

[5]Some writers use the term "harm prevention" but this term is ambiguous because there are different ways of preventing a harm. For example, one could try to avoid it or one could minimize the chance that it will occur.
[6]Cost–benefit analysis is a form of quantitative risk/benefit assessment [34]. In cost–benefit analysis, one assigns different probabilities to various outcomes and multiplies that value to create an expected value (or utility).

benefits, such as improvements in transportation safety, or increased transportation access for disabled and elderly people.

Proportionality is essential to counter the charge that the PP is fundamentally opposed to science and technology development, because proportionality requires us to consider potential harms in relation to potential benefits [34]. Excessive precaution can deny society important benefits and opportunities presented by new technologies and scientific discoveries, such as applications of AI. One must always consider the costs associated with precautionary measures in a way that reasonably balances benefits and risks.

As noted earlier, critics of the PP, argue that it undermines the advancement of science and technology [35]. Most of these critiques are mistaken, however, because they focus on excessively risk-averse versions of the PP and therefore make a "straw man" argument against applying the PP to AI/ML issues. The version of the PP used in this paper gives appropriate consideration to benefits and opportunities and is not excessively risk-averse. The PP can slow progress enough to give people time to think clearly and proactively about the issues at stake and decide upon reasonable precautions, but it does not unduly hinder scientific and technological progress, including, progress in AV development and deployment [34].

Fairness refers to the fairness of the distribution of risks and benefits as well as the fairness of the procedures or processes used to distribute benefits and risks. These two senses of fairness are conceptually distinct, and both are important. For example, one could argue that a jury trial was distributively unfair because it allowed a person who was clearly guilty to go free, while admitting that it was procedurally fair because it followed legal procedures for due process.

With respect to AVs, policymakers may need to consider how AVs affect different groups, including poor people, rich people, and disabled people. Egalitarian theories of justice would imply that an AV policy is distributively unfair if it primarily benefits rich people and not socioeconomically disadvantaged members of society, such as poor people and people with disabilities [110, 111]. To promote fairness, AVs should be widely available to all people, especially to those whose transportation needs are not currently being met, such as people who cannot drive due to disability or age [56].

Policymakers also need to ensure that the procedures and processes that affect the distribution of AV benefits and risks are fair. These procedures and processes could include methods for obtaining public input into AV policymaking, such as voting, public forums, surveys, and open meetings. One could argue that an AV policy is unfair if it has been developed and implemented at the behest of powerful industry groups with little public input. Conversely, one could argue that an AV policy is fair because it has received good input from the public and has not been unduly shaped or influenced by business or other special interests. Procedural fairness is especially important for dealing with controversial issues affecting the public, such as AV development and deployment, because people may disagree about how risks and benefits should be balanced and distributed [34].

Epistemic responsibility has to do with (1) making decisions based on the best available evidence and realistic assumptions; and (2) revising plans or policies based on new evidence or better assumptions. Epistemic responsibility also implies a duty to seek out new information by supporting or conducting research on AV safety. To help promote procedural fairness, the data should be transparent and open and available to the public. With respect to AVs, epistemic responsibility implies that government agencies and other key stakeholders should gather data on the risks and benefits of AVs and should conduct and support AV research and development (R&D). AV policies should be based on the best available evidence and sound assumptions. AV policies should be revised, if appropriate, as new evidence becomes available or assumptions are revised. For example, if research indicates that AVs are likely to produce harms that have not been adequately addressed in existing AV policies, those policies should be revised to minimize or mitigate those harms. AV data should also be open and transparent to the public. (These points are discussed further below.)

Finally, consistency refers to ensuring that decisions and policies are consistent with each other. For example, if a government agency restricts the use of a pesticide based on evidence concerning its risk/benefit profile, it should restrict other pesticides with similar risk/benefit profiles. Similar cases should be treated similarly. With respect to AV policy, regulations that are developed to promote safety and other important goals should be consistent. For example, it would be inconsistent to regulate two types of AVs with similar risk/benefit profiles differently.

8 Applying the precautionary principle to autonomous vehicle ethics and policy

Now that we have considered, in general, the relationship between the PP and AVs, we would like to discuss some specific implications of the PP for AV ethics and policy (see Table 3). In thinking about these implications, the first step is to consider general strategies societies have at their disposal. Once a general strategy has been chosen, more specific ethical and policy decisions can be made based on it. The paper will focus on Level 4–5 AVs, since these raise the biggest concerns. Three general strategies for dealing with Level 4–5 AVs include:

A. Permit the use and marketing of Level 4–5 AVs without any regulation or governmental control (i.e., a *Laissez Faire* approach).

B. Permanently ban Level 4–5 AVs.

C. Regulate and control/oversee Level 4–5 AVs.

To the best of our knowledge, most of the world's governments are pursuing some form of Option C. For example, the US, Australia, China, Germany, South Korea, and the UK are moving forward with AV regulation and oversight [112].

The PP would rule out Options A and B because they do not balance benefits and risks proportionally. The *Laissez Faire* approach is unreasonable because it does nothing to minimize or mitigate the risks of Level 4–5 AVs, which are substantial. A *Laissez Faire* approach is a reasonable option for dealing with the development of low-risk technologies, such as automated vacuum cleaners, but Level 4–5 AVs do not fall within this category. Indefinitely banning Level 4–5 AVs prevents harms but deprives society of important potential benefits of Level 4–5 AVs, such as reducing traffic accidents and deaths or increasing access to transportation for disabled or elderly people. Indefinitely banning Level 4–5 AVs is also epistemically irresponsible because it assumes, unrealistically, that bans can be enforced, which is not likely, given the various economic and political factors which are likely to drive higher-level AV development. As mentioned earlier, businesses and governments have already invested heavily in AV R&D. Even if one country could enforce a ban, this would not stop higher-level AV development around the globe. A higher-level AV ban would be no more successful than a ban on automobiles, which were once considered a dangerous and expensive technology [113]. However, temporary bans (or moratoria) on Level 4–5 AVs would not face such problems and could be a part Option C. For example, Level 5 AVs could be temporarily banned from public roads until there is substantial evidence concerning their risks.

Option C is more reasonable than A and B because it allows society to take advantage of the benefits and opportunities presented by higher-level AVs but also includes measures to minimize or mitigate harms. Floridi's concept of envelopment provides a useful way of thinking about Option C. Envelopment involves the establishment of physical, legal, and social boundaries around a technology to ensure that its uses are safe and appropriate [114, 115]. An enveloped technology is part of society, but it does not corrupt or control society. Some examples of envelopment include the use of AI in robot vacuum cleaners, face and fingerprint recognition, magnetic resonance imaging, and internet search engines.[7] So how can society envelop AVs? The devil, as they say, is in the details, and the details depend on data, which, unfortunately, are in short supply. To successfully envelop a technology, one must know how it works, what it does, and how it may be used [115]. Thus, the first two recommendations supported by the PP address the need for data.

Recommendation 1

Develop a publicly accessible, comprehensive, and systematic database concerning AV safety. While this recommendation may seem obvious, that does not mean it is trivial or that it is currently being met. Indeed, as noted earlier, one of the key obstacles to AV policymaking is the lack of good data about AV safety. Data are essential for government officials and the public to make decisions about how to weigh the benefits and risks of AVs when making decisions about AV laws, regulations, policies, and ethical guidelines. Since the public needs to be involved in this process, the database should include data which can be understood by non-experts [116]. The data should include information needed to make judgments concerning the proportionality of risks to benefits of AVs and the fairness of the distribution of risks. The data should therefore include information concerning accidents, injuries, deaths, miles driven, manufacturer, make, model, driving conditions, AV level (including Level 0 for comparison to human-driven vehicles). It should

[7]While some technologies can be enveloped successfully, others are so all-encompassing and powerful that they are difficult to keep within proper boundaries. For example, society has not successfully enveloped social media, which continues to create many problems that were not anticipated initially. It remains to be seen whether AVs can be successfully enveloped, but hopefully they can be.

also include information about who is harmed by AV accidents, such as driver/passenger/pedestrian/bystander, and their age, gender, and race/ethnicity, because this information is important for making judgments about the distribution of risks. Finally, the database should also include technical information, such as information about AV design, sensors, cybersecurity, software, support systems, and so on, which experts can use in providing advice to government officials and the public.

Public databases operated by the Centers for Disease Control and Prevention (CDC) provide a useful precedent for an AV database. The CDC makes a vast amount of data available to the public on disease prevalence, incidence, mortality, and morbidity, including historical trends and demographics, such as age, state of residence, gender, and race. The CDC's website also includes useful articles and essays that explain information about signs, symptoms, and causes of diseases, disease prevention and treatment, and health promotion. The CDC provides valuable data and information for scientists, doctors, public health officials, and laypeople to use in research, professional practice, and public health policy formation [117].

While few people would disagree with the recommendation to develop an AV database that can be used by experts, government officials, and the public, it does raise questions about who will collect the data and pay for data collection and maintenance of the database. This brings us to our second recommendation.

Recommendation 2

AV companies, academic institutions, and the government should work together to support and conduct basic and applied AV research and collect, store, secure, manage, share, analyze, and interpret AV performance and safety data. This recommendation is already being met, to a certain extent, but clearly much more work could be done [17, 118]. More money could be spent by private companies and the government on AV research and data collection, and this seems likely to happen. Private companies and government agencies should also work together to support AV databases.

While funding is important for promoting AV research, it is not sufficient. Research needs to be appropriately organized, managed, and executed to meet standards of rigor, reproducibility, quality, and integrity. Because the government plays a key role in overseeing industry and counterbalancing its influence on the research agenda, the relationship needs to be managed properly to avoid bias and conflict of interest [119].

Biomedical research provides a useful precedent here. In the US, the National Institutes of Health (NIH) supports basic and applied (or translational) biomedical research conducted on its campuses and at dozens of universities and medical centers across the country. Most of the clinical trials in the US are sponsored by private pharmaceutical and biotechnological, which collect data on products they are developing and planning to market. Companies submit their data to the Food and Drug Administration (FDA), which approves new drugs, biologics, and medical devices. While NIH researchers often collaborate with private companies, these relationships are carefully managed to avoid bias and conflict of interest. Although the FDA is a regulatory agency that does not support research, it works closely with private companies on issues related to study design, safety monitoring, and data reporting [119]. Biomedical research databases, such as PubMed, ClinicalTrials.gov, and GenBank, are supported by the federal government.

Recommendation 3

Level 4–5 AV development and deployment should proceed incrementally and cautiously to allow sufficient time for research, data collection, analysis, risk assessment, regulatory oversight, and public and community engagement.

This recommendation is supported by the PP's emphasis on risk minimization and Floridi's envelopment idea. It is already being followed, to a certain extent, by many states and nations, but it is worth emphasizing here because of its importance [118]. Proceeding incrementally and cautiously helps to minimize risks by ensuring that there is sufficient time to obtain the information needed to regulate and control AVs and for the public to have meaningful input into AV policies. Level 4–5 AV deployment should start with small pilot programs in easily navigable, sparsely populated areas (e.g., rural counties and roads) and then expand into more difficult areas (e.g., suburbs, cities, highways) as Level 4–5 AVs increase in safety. Level 4–5 AVs could be deployed on pre-determined, standard routes, such as trips to an airport or shopping mall, before they are used in novel driving situations. Incremental deployment should correspond to AV levels. For example, Level 5 AVs should be deployed widely only after Level 4 AVs have been widely deployed [18, 86, 89]. Investments in transportation infrastructure, such as roads, clearly marked road signs, traffic lights, and bridges can help promote safety in AV testing and deployment and improve safety for other vehicles. Investments in infrastructure required by AVs, such as reliable internet and GPS access, will also be important [120].

While it is important to incorporate Level 4–5 AVs into public transportation systems to promote access to transportation for people who lack it due to disability, age, or income, caution should be taken with this step, because accidents involving public transport vehicles (e.g., buses, shuttles) are potentially more disastrous than those involving passenger cars [121–123]. Clearly, some types of public transportation, such as shuttling people between parking lots and airports, involve straightforward, low-risk routes that lend themselves to automation. Other types of public transportation, such as taking children to and from school or transporting people around a busy city, do not. Using AVs in public transportation also raises important fairness issues, because poor people tend to use public transportation more than rich people, and they should not be disproportionately burdened with AV risks. Likewise, children should not be unfairly burdened with risks created by the use of automated school buses.

Caution should also be taken with introducing Level 4–5 AVs into the trucking industry. Although Level 4–5 AVs could play a major role in shipping products and materials and relieve the shortage of truck drivers, trucking also raises major safety issues because a malfunctioning AV truck could produce much more damage than a small vehicle. As with public transportation, some applications are safer than others. For example, driving a truck on an interstate highway lends itself more readily to automation than driving a truck in a busy city. Human drivers who operate Level 4–5 AV trucks must be ready to take over operation of the vehicle at any time [124, 125].

The “proceed with caution” strategy defended here also helps to promote procedural fairness by giving communities and the public enough time to have meaningful input into Level 4–5 AV policy. If a majority of the members of a community strongly opposes implementation of a Level 4–5 AV pilot program where they live, their wishes should be respected. Exposing people to significant risks created by novel technologies without appropriate consent is unfair and can lead to public backlash [34, 126].

The cultivation of genetically modified (GM) crops in Europe provides an instructive historical lesson for AV development [34, 126]. GM crops were introduced in the mid-1990s before there was sufficient time for public dialog and discussion about this new agricultural technology. Although scientific evidence at that time indicated that GM foods were safe to eat and that GM crops would not significantly harm the environment, many Europeans rejected this position and called these new products “Frankenfoods.” Due in large part to this public backlash, the European Union banned GM foods and crops from 1998 to 2007. Many European countries still severely restrict GM crop cultivation or ban it [34]. AV companies would be wise to learn from the mistakes that biotechnology companies made in foisting GM crops on a skeptical public.

Recommendation 4

Level 4–5 AVs should include manual overrides to avoid or minimize harms due to malfunctions, tampering, or hacking. Engineers are taught to think about what can go wrong with a machine or system and to prepare for it, and this piece of wisdom most certainly applies to AVs. The PP would favor manual overrides as a reasonable way of avoiding or minimizing harms. Although Level 4–5 AVs are designed to operate autonomously, human beings should have the ability control the vehicle [127, 128]. For example, if the AV’s AI/ML system malfunctions or someone hacks into the AV and takes control of it, the passenger, or possibly a remote AV operator or public safety officer, should have the ability to override the automation and stop the vehicle. Since overriding the AV may be difficult for people due to disability, age, or other factors, it may be necessary to restrict solo AV usage to people who have the ability to override the AV and take effective action to promote safety if something goes wrong. They need not be fully qualified, competent drivers but they would need to be able to safely stop the vehicle if necessary. For example, a young child or an adult who has quadriplegia or severe dementia or is inebriated should not be allowed to ride in an AV by themselves. They would need to be accompanied by someone who could safely stop the vehicle if necessary.

Recommendation 5

Develop systems of legal and ethical oversight of AVs, including statutes, regulations, forms of liability, and ethical guidelines, to minimize and mitigate the risks of AVs and control their uses. This recommendation is already being pursued by various governments around the world [16, 17, 22, 87, 112, 129–132]. The PP and the envelopment concept support legal oversight as a form of risk minimization and mitigation because laws can help to promote AV safety, dependability, and performance and protect rights to privacy and autonomy. Epistemic responsibility requires that regulations should be based on up-to-date, publicly available information and data pertaining to AV safety, risks, design, manufacturing, use, etc. Consistency requires that AV regulations should be consistent with regulations for non-autonomous vehicles and that there should not be significant variation in AV regulation across different states or regions within a country. The US federal government could promote consistency across states by developing model regulations and guidelines. Procedural fairness requires that all affected stakeholders, including the general public, AV users, and AV manufacturers, have meaningful input into the development of AV regulations.

While much of the current legal discussion has focused on statutes and regulations pertaining to AVs, it is also important to ensure that the legal system provides appropriate civil and criminal liability for AV accidents. Because human drivers are not in full control of Level 4–5 AVs, accidents with these vehicles raise novel issues concerning how to apportion legal responsibility among the driver, the manufacturer, an AV transportation company, and other parties. It is important to ensure that these issues are properly addressed, so that people who are harmed by Level 4–5 AVs are compensated fairly and wrongdoers are punished appropriately. Financial compensation for AV accidents also helps to promote distributive fairness by ensuring that injured parties do not bear an unfair burden of AV-associated harms. Legal liability for AV accidents can also provide manufacturers with incentives to improve safety, which helps to minimize harms [129–132].

It is also important for oversight systems to include ethical guidelines for AV developers and manufacturers to address situations that are not covered by the legal system. Ethical guidelines can address such issues as AV design and operations, the privacy and consent of AV users, and decision-making by AI/MLs that control AVs [83–87, 133]. Key principles of AV ethics could be based on principles of AI ethics [134]. Some of these include honesty, accountability, transparency, respect for human dignity and rights, justice, and social responsibility.

Recommendation 6

Educate and inform AV users and the public about AV operations and safe use. Educating and informing passengers and the public about AVs can help to minimize risks and promote autonomy by giving people the information they need to decide whether they want to take risks associated with AVs. Understanding how AVs work can also help drivers and passengers take effective action to protect themselves or others from harm. Drivers need to be informed about how the automated functions work and how to use them. Passengers in Level 4–5 AVs should be informed about: operations and safety features; how to disable the automated system, if necessary; how the AV will behave when it encounters situations involving moral choices related to saving or taking human life; and privacy protections [126]. Also, the public should be informed about Level 4–5 AV deployment. Like driver education vehicles, Level 4–5 AVs should be clearly marked so that other drivers and pedestrians can act accordingly and protect themselves from harm. Education and information sharing, it should be stressed, should be accessible, non-technical, iterative, thorough, and meaningful and go way beyond providing data and specifications in owner's manuals or warning labels. AV companies could operate websites that address frequently asked questions and allow the user to contact a human being for additional information. AV manufacturers could also engage in public information campaigns on television and social media to inform the public about their products.

9 Conclusion

In this paper, we have applied the PP to ethical and policy issues related to developing and deploying autonomous vehicles. The PP can provide a valuable perspective on these topics, given the high degree of scientific uncertainty and moral disagreement related to Level 4–5 AVs. Contrary to what some might claim, the PP does not unduly interfere with AV development and deployment. Instead, using the PP as a decision-making framework can help slow down the advancement of AV technology and its applications enough to give stakeholders and the public time to think clearly and proactively about how best to manage the risks associated with AVs and ensure that AV policies are reasonable, responsible, consistent, and fair. If (or when) scientific uncertainty and moral disagreement about AVs decrease, it may be appropriate to move from a precautionary framework toward a risk assessment/management approach. We welcome and encourage further discussion of the reasons for or against applying the PP to AV policy, as well as specific recommendations that would be supported by the PP.

Acknowledgements

This research was supported, in part, by the Intramural Program of the National Institute for Environmental Health Sciences (NIEHS), National Institutes of Health (NIH). It does not represent the views of the NIEHS, NIH, or US federal government.

Funding

David Resnik: National Institute of Environmental Health Sciences, ziaes102646-10.

Data availability

There is no data associated with this publication.

References

1. Shladover SE: 'Self-Driving' cars begin to emerge from a cloud of hype. Scientific American, Sept 25, 2021. https://www.scientificamerican.com/article/self-driving-cars-begin-to-emerge-from-a-cloud-of-hype/
2. Mervis J: Are we going too fast on driverless cars? Science, December 14, 2017. Available at: https://www.sciencemag.org/news/2017/12/are-we-going-too-fast-driverless-cars (2017). Accessed 14 Dec 2017
3. Parekh D, Poddar N, Rajpurkar A, Chahal M, Kumar N, Joshi GP, Cho WA: Autonomous vehicles: Progress, methods and challenges. Electronics 11, 2162 (2022)
4. Winton N: Computer driven autos still years away despite massive investment. Forbes, February 27, 2022. https://www.forbes.com/sites/neilwinton/2022/02/27/computer-driven-autos-still-years-away-despite-massive-investment/?sh=34f9695618cc

5. Skelley KD: Testing of the Army's first autonomous vehicle speeds ahead. DEVCOM Aviation & Missile Center Public Affairs, August 11, 2022. https://www.army.mil/article/259244/testing_of_the_armys_first_autonomous_vehicle_speeds_ahead

6. Winston C: Autonomous vehicles could improve policing, public safety, and much more. Brookings Institution, August 25, 2020. https://www.brookings.edu/blog/techtank/2020/08/25/autonomous-vehicles-could-improve-policing-public-safety-and-much-more/

7. Insights CB. Autonomous driving startups raise record funding as the push for commercialization begins. CB Insights, October 20, 2021. https://www.cbinsights.com/research/autonomous-vehicle-tech-funding-trends/

8. Fannin R: Where the billions spent on autonomous vehicles by U.S. and Chinese giants is heading. CNBC Disruptor, May 21, 2022. https://www.cnbc.com/2022/05/21/why-the-first-autonomous-vehicles-winners-wont-be-in-your-driveway.html

9. Wilkes W: Volkswagen to invest $2.3 billion in China autonomous driving venture. Bloomberg, October 13, 2022. https://www.bloomberg.comnews/articles/2022-10-13/vw-to-invest-2-3-billion-in-china-autonomous-driving-venture?leadSource=uverify%20wall

10. Chafkin M: Even after $100 billion, self-driving cars are going nowhere. Bloomberg, October 6, 2022. https://www.bloomberg.com/news/features/202210-06/even-after-100-billion-self-driving-cars-are-going-nowhere?leadSource=uverify%20wall

11. Krisher T: Tesla cars involved in 10 of the 11 new crash deaths linked to automated-tech vehicles. Fortune, October 19, 2022. https://fortune.com/2022/10/19/tesla-cars-involved-in-10-of-the-11-new-crash-deaths-linked-toautomated-tech-vehicles/

12. Tesla. Autopilot and full self-driving capability. Tesla, November 21, 2022. https://www.tesla.com/support/autopilot

13. Reuters. Baidu bags China's first fully driverless robotaxi licenses. Reuters, August 7, 2022. https://www.reuters.com/business/autos-transportation/baidubags-chinas-first-fully-driverless-robotaxi-licenses-2022-08-08/

14. Schoppa C Top autonomous vehicles companies to watch in 2022. AT Time Journal, September 26, 2022. https://www.aitimejournal.com/autonomousvehicles-companies-to-watch

15. Global X Research Team. April 17, 2018. https://www.globalxetfs.com/future-of-transportation-is-autonomous-electric/

16. Husch B, Teigen A: Regulating autonomous vehicles. National Conference of State Legislatures 25, 13, April 2017. https://www.ncsl.org/research/transportation/regulating-autonomous-vehicles.aspx

17. Taeihagh A, Lim HS: Governing autonomous vehicles: emerging responses for safety, liability, privacy, cybersecurity, and industry risks. Transp. Rev 39(1), 103–128 (2019)

18. National Highway Traffic Safety Administration. U.S. Department of Transportation announces expansion of AV test initiative, press release. January 11, 2021. https://www.nhtsa.gov/press-releases/us-departmenttransportation-announces-expansion-av-test-initiative

19. National Highway Traffic Safety Administration. Budget estimates. November 21, 2022. https://www.nhtsa.gov/sites/nhtsa.gov/files/documents/fy_2021_nhtsa_congressional_justification.pdf

20. National Highway Traffic Safety Administration. NHTSA announces $260 million in grants for highway safety programs made possible by the bipartisan infrastructure law, press release. December 16, 2021. https://www.transportation.gov/briefing-room/nhtsa-announces-260-million-grants-highway-safety-programs-made-possible-bipartisan

21. National Highway Traffic Safety Administration. Research. November 21, 2022. https://www.nhtsa.gov/research

22. Insurance Institute for Highway Safety. Autonomous vehicle laws. November 21, 2022. https://www.iihs.org/topics/advanced-driver-assistance/autonomousvehicle-laws

23. Beckford A: Ralph Nader begs NHTSA to pull Tesla's "full self-driving" software. Motortrend, August 10, 2022. https://www.motortrend.com/news/ralph-nader-nhtsa-recall-tesla-fsd/

24. Shephardson D: 'Extensive' Tesla Autopilot probe proceeding 'really fast', U.S. official says. Reuters, January 9, 2023. https://www.reuters.com/technology/us-agency-working-really-fast-nhtsa-autopilot-probe-2023-01-09/(2023)

25. Jones S: Ford CEO puts the brakes on the company's plan for fully autonomous cars, saying there's a long way to go before it can develop them at scale. Insider, October 27, 2022. https://www.businessinsider.com/ford-fully-autonomous-self-driving-cars-ceo-development-2022-10(2022)
26. Naughton K, Raymunt M, Bloomberg. Ford and Volkswagen pull the plug on robocar unit Argo AI in major setback to their self-driving plans. Fortune, October 27, 2022. https://fortune.com/2022/10/27/ford-volkswagen-pull-plug-robocar-unit-argo-ai-major-setback-self-driving-plans/
27. Gross A: Third time is not a charm as driving assistance tech continues to underperform. AAA Newsroom, May 12, 2022. https://newsroom.aaa.com/2022/05/consumer-skepticism-toward-active-driving-features-justified/
28. Shariff JF, Rahwan I: The social dilemma of autonomous vehicles. Science 352(6293), 1573–1576 (2016) [PubMed: 27339987]
29. Mordue G, Yeung A, Wu F: The looming challenges of regulating high level autonomous vehicles. Transp. Res. A Policy Pract 132, 174–187 (2020)
30. McFarland M: The backlash against self-driving cars officially begins. CNN, January 10, 2017. https://money.cnn.com/2017/01/10/technology/new-york-self-driving-cars-ridesharing/index.html.
31. Hurst L: Elon Musk has grand plans for Tesla self-driving cars, but the safety debate is getting ugly. Euronews, August 30, 2022. https://www.euronews.com/next/2022/08/30/elon-musk-has-grand-plans-for-tesla-self-driving-cars-but-the-safety-debate-is-getting-ugl
32. Beckford A: Ralph Nader begs NHTSA to pull Tesla's "full self-driving" software. Motortrend, August 10, 2022. https://www.motortrend.com/news/ralph-nader-nhtsa-recall-tesla-fsd/
33. The Intelligent Transportation Society of America. Multi-association letter to US Congress on automated vehicles, March 31, 2022. https://itsa.org/advocacy-material/multi-association-letter-to-us-congress-on-automated-vehicles-avs/
34. Resnik DB: Precautionary Reasoning in Environmental and Public Health Policy. Springer, Cham (2021)
35. Sunstein CR: Laws of Fear: Beyond the Precautionary Principle. Cambridge University Press, Cambridge, UK (2005)
36. National Highway Transportation Safety Administration. Automated Vehicles Safety. November 21, 2022. https://www.nhtsa.gov/technologyinnovation/automated-vehicles-safety
37. Wiseman Y: Autonomous vehicles. In: Mehdi Khosrow-Pour D (ed.) Encyclopedia of Information Science and Technology, 5th edn., pp. 1–11. Information Science Reference, Hershey, PA (2020)
38. Gadam S: Artificial intelligence and autonomous vehicles. Data Driven Investor, April 19, 2018. https://medium.datadriveninvestor.com/artificial-intelligence-and-autonomous-vehicles-ae877feb6cd2
39. Khvoynitskaya S: 3 types of autonomous vehicle sensors in self-driving cars. I Transition, February 11, 2020. https://www.itransition.com/blog/autonomous-vehicle-sensors
40. Dimitrakopoulos G, Tsakanikas A, Panagiotopoulos E: Autonomous Vehicles: Technologies, Regulations, and Societal Impacts, 1st edn. Elsevier, Amsterdam (2021)
41. Keymakr. Overcoming bias in data annotation for autonomous vehicles. Keymakr, May 12, 2021. https://keymakr.com/blog/overcoming-bias-in-data-annotation-for-autonomous-vehicles/
42. Edelstein S: Researchers fool self-driving cars with stickers on street signs. The Drive, August 7, 2017.
43. Eliot L: Pranking driverless cars, dangerously foolish acts. AI Insider, April 1, 2019. https://lance-eliot.medium.com/pranking-driverless-cars-dangerously-foolish-acts-3acaecc80b5e(2019)
44. Schwarting W, Alonso-Mora J, Rus D: Planning and decision-making for autonomous vehicles. Annu. Rev. Control Robot. Auton. Syst 1(1), 187–210 (2018)
45. Vargas J, Alsweiss S, Toker O, Razdan R, Santos J: Sensors and their vulnerability to weather conditions. Sensors 21, 5397 (2021) [PubMed: 34450839]
46. Boberg C: Five things you need to know about connected autonomous vehicles. Ericsson.com, November 15, 2021. https://www.ericsson.com/en/blog/2021/11/five-things-you-need-to-know-about-connected-autonomous-vehicles

47. Kim K, Kim JS, Jeong S, Park JH, Kim HK: Cybersecurity for autonomous vehicles: Review of attacks and defense. Comput. Secur 103, 102150 (2021)
48. Centers for Disease Control and Prevention. Road traffic injuries and deaths—a global problem. November 21, 2022. https://www.cdc.gov/injury/features/global-road-safety/index.html#:~:text=Each%20year%2C%201.35%20million%20people,on%20roadways%20around%20the%20world.&text=Every%20day%2C%20almost%203%2C700%20people,pedestrians%2C%20motorcyclists%2C%20or%20cyclists.
49. National Highway Transportation Administration. Newly released estimates show traffic fatalities reached a 16-year high in 2021, press release, May 17, 2022. https://www.nhtsa.gov/press-releases/early-estimate-2021-traffic-fatalities
50. Pavia W: Driverless Uber car 'not to blame' for woman's death. The Times, March 21, 2018. https://www.thetimes.co.uk/article/driverless-uber-car-not-to-blame-for-woman-s-death-klkbt7vf0
51. NPR. A Tesla driver is charged in a crash involving Autopilot that killed 2 people. NPR, January 18, 2022. https://www.npr.org/2022/01/18/1073857310/tesla-autopilot-crash-charges
52. McFarland M: Tesla Autopilot's safety questioned after latest fatal motorcycle crash. CNN, October 17, 2022. https://www.cnn.com/2022/10/17/business/tesla-motorcycle-crashes-autopilot
53. Kanchwala H: Are autonomous cars really safer than human-driven cars? Science ABS, July 8, 2022. https://www.scienceabc.com/innovation/are-automated-cars-safer-than-human-driven-cars.html
54. Fleetwood J: Public health, ethics, and autonomous vehicles. Am. J. Public Health 107(4), 532–537 (2017) [PubMed: 28207327]
55. Rojas-Rueda D, Nieuwenhuijsen MJ, Khreis H, Frumkin H: Autonomous vehicles and public health. Annu. Rev. Public Health 41, 18.1–18.17 (2020)
56. Emory K, Douma F, Cao J: Autonomous vehicle policies with equity implications: Patterns and gaps. Transp. Res. Interdiscip. Perspect 13, 100521 (2022)
57. Wu X, Cao J, Douma F: The impacts of vehicle automation on transport-disadvantaged people. Transp. Res. Interdiscip. Perspect 11, 100447 (2021)
58. Pritchard T: Self-driving cars will be way too pricey for most people — and this proves it. Tom's Guide, October 06, 2021. https://www.tomsguide.com/opinion/self-driving-cars-will-be-way-too-pricey-for-most-people-and-this-proves-it
59. Pan S, Fulton LM, Roy A, Jung J, Choi Y, Gao HO: Shared use of electric autonomous vehicles: Air quality and health impacts of future mobility in the United States. Renew. Sustain. Energy Rev 149, 111380 (2021)
60. Ajanovic A, Haas R: Electric vehicles: solution or new problem? Environ. Dev. Sustain 20(Suppl 1), 7–22 (2018)
61. Circella G, Hardman S: Self-driving cars could encourage a lot more driving—and more carbon emissions. Fast Company, February 5, 2022. https://www.fastcompany.com/90718971/self-driving-cars-more-driving-more-carbon-emissions#:~:text=More%20miles%2C%20more%20carbon%20emissions,%2C%20energy%20consumption%2C%20and%20pollution
62. Balakrishnan A: Self-driving cars could cost America's professional drivers up to 25,000 jobs a month, Goldman Sachs says. CNBC, May 22, 2017. https://www.cnbc.com/2017/05/22/goldman-sachs-analysis-of-autonomous-vehicle-job-loss.html
63. Coppola G: Driverless cars are proving to be job creators, at least so far. Bloomberg, August 10, 2021. https://www.bloomberg.com/news/newsletters/2021-08-10/driverless-cars-are-proving-to-be-job-creators-at-least-so-far
64. Vermeulen B, Kesselhut J, Pyka A, Saviotti PP: The impact of automation on employment: Just the usual structural change? Sustainability 10(5), 1661 (2018)
65. Mintor S: Why the economy needs churn. Industry Week, May 15, 2012. https://www.industryweek.com/the-economy/article/22010108/why-the-economy-needs-churn
66. Vallor S: Moral deskilling and upskilling in a new machine age: Reflections on the ambiguous future of character. Philos. Technol 28, 107–124 (2015)
67. Casner S: Dumbing it down in the cockpit. Slate, December 12, 2014. https://slate.com/technology/2014/12/automation-in-the-cockpit-is-making-pilots-thinking-skills-duller.html

68. Goodyear S: GPS, smartphones, and the dumbing down of personal navigation. Bloomberg, January 24, 2012. https://www.bloomberg.com/news/articles/2012-01-24/gps-smartphones-and-the-dumbing-down-of-personal-navigation
69. Batts V: Often crucial to survival, the skill of map reading is becoming a lost art. Natural News, June 9, 2019. https://naturalnews.com/2019-06-09-the-skill-of-map-reading-is-becoming-a-lost-art.html
70. Frischmann M: Is smart technology making us dumb? Scientific American, December 27, 2018. https://blogs.scientificamerican.com/observations/is-smart-technology-making-us-dumb/ (2018)
71. Bratsberg B, Rogeberg O: Flynn effect and its reversal are both environmentally caused. Proc. Natl. Acad. Sci. USA 115(26), 6674–6678 (2018) [PubMed: 29891660]
72. Robotics RS: Ethics of artificial intelligence. Nature 521, 415–418 (2015) [PubMed: 26017428]
73. Liao S (ed.): Ethics of Artificial Intelligence, 1st edn. Oxford University Press, New York (2020)
74. Boddington P: Towards a Code of Ethics for Artificial Intelligence. Springer, Cham (2017)
75. Coeckelbergh M: AI Ethics. MIT Press, Cambridge, MA (2020)
76. Bostrom N: Superintelligence: Paths, Dangers, Strategies. Oxford University Press, Oxford, UK (2014)
77. Office of Science and Technology Policy. Blueprint for an AI Bill of Rights, October 4, 2022. https://www.whitehouse.gov/ostp/ai-bill-of-rights/
78. Denton J: Is the trolley problem derailing the ethics of self-driving cars? Pacific Standard, November 29, 2018. https://psmag.com/economics/is-the-trolley-problem-derailing-the-ethics-of-self-driving-cars
79. Thomson JJ: The trolley problem. Yale Law J. 94(6), 1395–1415 (1985)
80. Goodall NJ: Away from trolley problems and toward risk management. Appl. Artif. Intell 30(8), 810–821 (2016)
81. Leben D: A Rawlsian algorithm for autonomous vehicles. Ethics Inf. Technol 19, 107–115 (2017)
82. Faulhaber AK, Dittmer A, Blind F, Wächter MA, Timm S, Sütfeld LR, Stephan A, Pipa G, König P: Human decisions in moral dilemmas are largely described by utilitarianism: Virtual car driving study provides guidelines for autonomous driving vehicles. Sci. Eng. Ethics 25(2), 399–418 (2019) [PubMed: 29357047]
83. Rhim J, Lee JH, Chen M, Lim A: A deeper look at autonomous vehicle ethics: An integrative ethical decision-making framework to explain moral pluralism. Front. Robot. AI 8, 632394 (2021)
84. Hanna MJ, Kimmel SC: Current US federal policy framework for self-driving vehicles: Opportunities and challenges. Computer 50(12), 32–40 (2017)
85. Borenstein J, Herkert JR, Miller KW: Self-driving cars and engineering ethics: The need for a system level analysis. Sci. Eng. Ethics 25(2), 383–398 (2019)
86. Martinho A, Herber N, Kroesen M, Chorus C: Ethical issues in focus by the autonomous vehicles industry. Transp. Rev 41(5), 556–577 (2021)
87. Pattinson JA, Chen H, Basu S: Legal issues in automated vehicles: Critically considering the potential role of consent and interactive digital interfaces. Human. Soc. Sci. Commun 7, 153 (2020)
88. Kuziemski M: A precautionary approach to artificial intelligence. Project Syndicate, May 1, 2018. https://www.project-syndicate.org/commentary/precautionary-principle-for-artificial-intelligence-by-maciej-kuziemski-2018-05
89. Castro D, McLaughlin M: Ten ways the precautionary principle undermines progress in artificial intelligence. Information Technology and Innovation Foundation, February 4, 2019. https://itif.org/publications/2019/02/04/ten-ways-precautionary-principle-undermines-progress-artificial-intelligence/
90. Thierer A: The proper governance default for AI. Medium, May 26, 2022. https://medium.com/@AdamThierer/the-proper-governance-default-for-ai-d05db6970924#:~:text=The%20Problem%20with%20the%20Precautionary,not%20cause%20any%20theoretical%20harms
91. Hemphill TA: The innovation governance dilemma: Alternatives to the precautionary principle. Technol. Soc 63, 101381 (2020)

92. Resnik DB: Is the precautionary principle unscientific? Stud. Hist. Philos. Biol. Biomed. Sci 34(3), 329–344 (2003)
93. Sandin P: Better Safe than Sorry: Applying Philosophical Methods to the Debate on Risk and the Precautionary Principle. Theses in Philosophy from the Royal Institute of Technology, Stockholm (2004)
94. Munthe C: The Price of Precaution and the Ethics of Risks. Springer, Dordrecht (2011)
95. Hartzell-Nichols L: A Climate of Risk: Precautionary Principles, Catastrophes, and Climate Change. Routledge, New York (2017)
96. Steel D: Philosophy and the Precautionary Principle. Cambridge University Press, Cambridge, UK (2015)
97. United Nations Conference on Economic Development. The Rio Declaration on Environment and Development, 1992. https://www.un.org/en/development/desa/population/migration/generalassembly/docs/globalcompact/A_CONF.151_26_Vol.I_Declaration.pdf
98. Cranor C: Learning from law to address uncertainty in the precautionary principle. Sci. Eng. Ethics 7, 313–326 (2001) [PubMed: 11506420]
99. Elliott KC: Geoengineering and the precautionary principle. Int. J. Appl. Philos 24, 237–253 (2010)
100. Elliott KC. Nanomaterials and the precautionary principle. Environmental Health Perspectives 119, 6: A240 (2011). [PubMed: 21205582]
101. Kuhlau F, Höglund AT, Evers K, Eriksson S: A precautionary principle for dual use research in the life sciences. Bioethics 25(1), 1–8 (2011)
102. Kundi M, Hardell L, Sage C, Sobel E: Electromagnetic fields and the precautionary principle. Environ. Health Perspect 117(11), A484–485 (2009) [PubMed: 20049098]
103. Resnik DB: The precautionary principle and medical decision making. J. Med. Philos 29, 281–299 (2004) [PubMed: 15512973]
104. Resnik DB: H1N1 avian flu research and the ethics of knowledge. Hastings Cent. Rep 43(2), 22–33 (2013) [PubMed: 23390001]
105. Peterson M: The precautionary principle should not be used as a basis for decision-making. Talking point on the precautionary principle. EMBO Rep. 8(4), 305–308 (2007) [PubMed: 17401402]
106. Sandin P, Peterson M, Hansson SO, Rudén C, Juthe A: Five charges against the precautionary principle. J. Risk Res 5(4), 287–299 (2002)
107. Soule E: The precautionary principle and the regulation of U.S. food and drug safety. J. Med. Philos 29(3), 333–350 (2004) [PubMed: 15512976]
108. Nuffield Council on Bioethics. Public Health: Ethical Issues. London, UK: Nuffield Council, 2007.
109. Karliuk M: Proportionality principle for the ethics of artificial intelligence. AI Ethics (2022). 10.1007/s43681-022-00220-1
110. Omaar H: Principles to promote responsible use of AI for workforce decisions. Center for Data Innovation, August 9, 2021. https://datainnovation.org/2021/08/principles-to-promote-responsible-use-of-ai-for-workforce-decisions/#:~:text=To%20address%20these%20concerns%2C%20several,until%20it%20is%20proven%20safe
111. Rawls J: A Theory of Justice. Harvard University Press, Cambridge, MA (1971)
112. Dentons. Global Guide to Autonomous Vehicles 2022. https://www.dentons.com/en/insights/guides-reports-and-whitepapers/2022/may/6/global-guide-to-autonomous-vehicles-2022
113. Winton A: Get a horse! America's skepticism toward the first automobiles. Saturday Evening Post, January 9, 2017. https://www.saturdayeveningpost.com/2017/01/get-horse-americas-skepticism-toward-first-automobiles/
114. Floridi L: Enveloping the world: the constraining success of smart technologies. In: Mauger, J. (ed.) CEPE 2011: Ethics in Interdisciplinary and Intercultural Relations, pp. 111–116 (2011). https://coeckelbergh.files.wordpress.com/2015/03/48.pdf

115. Robbins S: AI and the path to envelopment: knowledge as a first step towards the responsible regulation and use of AI-powered machines. AI Soc. 35, 391–400 (2020)
116. Elliott KC, Resnik DB: Making open science work for science and society. Environ. Health Perspect 127(7), 75002 (2019) [PubMed: 31353949]
117. Centers for Disease Control and Prevention. About CDC. 2023. https://www.cdc.gov/about/
118. Wiseman Y, Grinberg I: Circumspectly crash of autonomous vehicles. Proceedings of IEEE Conference on Electro Information Technology, pp. 387–392. https://u.cs.biu.ac.il/~wisemay/eit2016a.pdf (2016)
119. Shamoo AE, Resnik DB: Responsible Conduct of Research, 4th edn. Oxford University Press, New York (2022)
120. Tengilimoglu O, Carsten O, Wadud Z: Infrastructure requirements for the safe operation of automated vehicles: Opinions from experts and stakeholders. Transp. Policy 133, 209–222 (2023)
121. Levassor. Enhancing public transport using autonomous vehicle technology. Intelligent Transport, 2022. https://www.intelligenttransport.com/transport-articles/137016/public-transport-autonomous-vehicle/
122. de Lara A: Autonomous vehicles in public transit today…and in the future. Metro Magazine, January 31, 2022. https://www.metro-magazine.com/10154922/autonomous-vehicles-in-public-transit-today-and-in-the-future
123. National Association for State Community Services Programs. The Stranded Poor: Recognizing the Importance of Public Transportation for Low-Income Households, 2008. https://nascsp.org/wp-content/uploads/2018/02/issuebrief-benefitsofruralpublictransportation.pdf
124. US Library of Congress. Autonomous trucking. 2023. https://guides.loc.gov/trucking-industry/autonomous-trucking
125. Metz C: The long road to driverless trucks. New York Times, September 28, 2022. https://www.nytimes.com/2022/09/28/business/driverless-trucks-highways.html
126. National Academies of Sciences, Engineering, and Medicine: Gene Drives on the Horizon: Advancing Science, Navigating Uncertainty, and Aligning Research with Public Values. National Academies Press, Washington, DC (2016)
127. Ventura P: Autonomous vehicles need a manual override. The Telegram, December 16, 2018. https://www.telegram.com/story/opinion/columns/guest/2018/12/16/as-i-see-it-autonomous-vehicles-need-manual-override-to-reduce-automobile-fatalities/6632734007/
128. Boyce C, Patterson T, Arguello JC, Vena M: Manual override of self-driving vehicle. Paper prepared for Delphi Automotive, December 8, 2017. https://cpb-us-e1.wpmucdn.com/sites.psu.edu/dist/9/78438/files/2017/09/PROJECT-02-REPORT-209t8a6.pdf
129. Sucharski IL, Fabinger P: Privacy in the age of autonomous vehicles. Wash. Lee Law Rev 73, 724 (2017)
130. Menon C, Alexander R: A safety-case approach to the ethics of autonomous vehicles. Saf. Reliab 39(1), 34–58 (2020)
131. Collingwood L: Privacy implications and liability issues of autonomous vehicles. Inf. Commun. Technol. Law 26(1), 32–45 (2017)
132. Hevelke A, Nida-Rümelin J: Responsibility for crashes of autonomous vehicles: an ethical analysis. Sci. Eng. Ethics 21(3), 619–630 (2015) [PubMed: 25027859]
133. Evans K, de Moura N, Chauvier S, Chatila R, Dogan E: Ethical decision making in autonomous vehicles: The AV ethics project. Sci. Eng. Ethics 26, 3285–3312 (2020) [PubMed: 33048325]
134. Floridi L, Cowls J: A unified framework of five principles for AI in society. Harvard Data Sci. Rev 1, 1 (2019)

Table 1

Levels of automation (based on National Highway Transportation Safety Administration) [36]

Automation Level	Description	Examples
0 Momentary Driver Assistance	System provides momentary driving assistance, like warnings and alerts or emergency safety interventions, while driver remains fully engaged and attentive	Automatic emergency braking Forward collision warning Lane departure warning
1 Driver Assistance	System provides continuous assistance with either acceleration/braking OR steering, while driver remains fully engaged and attentive	Adaptive cruise control Lane keeping assistance
2 Additional Assistance	System provides continuous assistance with both acceleration/braking AND steering, while driver remains fully engaged and attentive	Highway pilot Automated parallel parking
3 Conditional Automation	System actively performs driving tasks while driver remains available to take over	Tesla's Autopilot Vehicles available for purchase in some countries
4 High Automation	System is fully responsible for driving tasks within limited service areas, while occupants act only as passengers and do not need to be engaged	Vehicles not available for purchase
5 Full Automation	System is fully responsible for driving tasks, while occupants act only as passengers and do not need to be engaged	Vehicles not available for purchase

Table 2

Benefits and risks of AVs

Type	Potential benefits	Risks
Public Health	May significantly improve vehicle safety and reduce mortality and morbidity related to driving	May significantly decrease vehicle safety and increase mortality and morbidity related to driving
Psychosocial	May improve access to transportation for people who are disabled, elderly, or socioeconomically disadvantaged and thereby improve their opportunities and quality of life	May be available only to people with high incomes or wealth May encourage deskilling and “dumbing down” of human mental capacities by increasing reliance on machines to perform tasks that require cognitive work
Environmental	Electric autonomous vehicles may reduce production of gases that contribute to air pollution and global warming	May increase energy consumption and pollution by increasing utilization of automated transportation
Economic	May create jobs related to AV development, manufacture, servicing, and operation	May eliminate jobs, especially among professional drivers, such as truck drivers and taxi drivers

Table 3

Some AV policy recommendations supported by the precautionary principle

Recommendation	Rationale
1. Develop a publicly accessible AV database	Proportionality, fairness, consistency, epistemic responsibility
2. Conduct AV research and data collection	Proportionality, fairness, consistency, epistemic responsibility
3. Pursue incremental development and deployment of AVs	Proportionality, fairness
4. Include manual overrides in AVs	Proportionality, fairness
5. Develop systems of legal and ethical oversight of AVs	Proportionality, fairness, consistency, epistemic responsibility
6. Educate and inform the public about AV operations and safe use	Proportionality, fairness