AI4People:

Ethical Guidelines for the Automotive Sector – Fundamental Requirements and Practical Recommendations

Christoph Lütge, Technical University of Munich, Munich, Germany

Franziska Poszler, Technical University of Munich, Munich, Germany

Aida Joaquin Acosta, Mobility and the Urban Agenda, Spain & Harvard University, USA

David Danks, Carnegie Mellon University, USA

Gail Gottehrer, Law Office of Gail Gottehrer LLC, USA

Lucian Mihet-Popa, Østfold University College, Halden, Norway

Aisha Naseer, Fujitsu Laboratories of Europe, UK

ABSTRACT

This paper presents the work of the AI4People-Automotive Committee established to advise more concretely on specific ethical issues that arise from autonomous vehicles (AVs). Practical recommendations for the automotive sector are provided across the topic areas: human agency and oversight, technical robustness and safety, privacy and data governance, transparency, diversity, non-discrimination and fairness, societal and environmental wellbeing, as well as accountability. By doing so, this paper distinguishes between policy recommendations that aim to assist policymakers in setting acceptable standards and industry recommendations that formulate guidelines for companies across their value chain. In the future, the automotive sector may rely on these recommendations to determine relevant next steps and to ensure that AVs comply with ethical principles.

KEYWORDS

AI4People, Autonomous Driving, Autonomous Vehicle Ethics, Ethics of AI, Fundamental Rights, Governance, Override, Regulation, Self-Driving Cars, Transparency

AIM AND SCOPE OF THIS PAPER

In the past decade, many policy documents have discussed ethical issues and potential future directions related to new emerging technologies such as artificial intelligence (AI) or autonomous systems. This paper presents the work of the AI4People-Automotive Committee[1] established to advise more concretely on specific ethical issues that arise from autonomous vehicles (AVs). The committee consisted of industry experts and researchers from the fields of ethics, law, philosophy, engineering,

technology and policy. The aim of this paper is to provide the automotive sector, including both companies and public entities such as regulators, with concrete and practical guidelines to comply with ethical principles within the AI systems of AVs. Therefore, this paper could serve as a checklist for policymakers and companies as well as a basis for developing a certification of ethics, an 'ecosystem of trust' (European Commission, 2020b) and ultimately a 'Good AI Society' (Floridi et al., 2018) in the automotive sector. These guidelines are intended to provide a clearer vision and moral compass on how to proceed and what to consider when developing AVs, rather than additional barriers to innovation. The automotive sector is defined here in the broadest terms possible to encompass a wide range of companies involved in the development of vehicles, including private cars, trucks, busses and shuttles. Sea, air and military-type applications have been excluded due to their functional and ethical specificity. This paper will focus on the ethics of the AI-based tools that are used in automotive technology, rather than on the ethics of vehicles in general.

This paper distinguishes between high-level guidelines for policymakers ('policy recommendations') and concrete actionable recommendations for companies ('industry recommendations'). However, the line between the two cannot always be drawn clearly which also highlights the importance of co-regulation (i.e. the interaction of legal regulation and self-regulation by companies) (Pagallo et al., 2019). The policy recommendations are designed to focus attention on pressing policy issues and assist in setting acceptable standards. Thus, the policy recommendations ultimately influence the industry recommendations. Responsible targets for the execution of the policy recommendations are: policymakers, legislators, ethics standards boards and commissions such as the United Nations Economic Commission for Europe (UNECE). The industry recommendations formulate guidelines for companies across their entire value chain (especially during research & development, production & operations and service). Therefore, original equipment manufacturer (OEM) and car manufacturers are the primary responsible targets for those recommendations.

Before turning to the principles and guidelines, we note three key points of consensus among the authors: (1) a responsible balancing of risks or estimated harm should be permitted at any time for AVs; (2) a large-scale introduction of full-mode AVs (level 4 and higher) onto streets is unlikely in the short run, so we must consider a more incremental, step-by-step approach; and (3) policymakers face significant challenges now, and so there are significant pressures to quickly develop a clear regulatory framework.

THE GUIDELINES

Fundamental Rights Underlying the Guidelines

Particular fundamental rights are the basis for the proposed seven requirements that were originally derived by the High-level Expert Group on Artificial Intelligence (2019) (i.e. human agency and oversight; technical robustness and safety; privacy and data governance; transparency; diversity, non-discrimination and fairness; societal and environmental wellbeing; accountability) and recommendations in this paper. In addition to general human dignity, key fundamental rights (United Nations, 1948) that policymakers and companies in the automotive sector should recognize are: *Right to self-determination and liberty* which draws attention to human agency in self-driving cars (i.e. importance of override options) (see Guideline 1). *Right to life and security* which entails ensuring technical robustness and safety of operating self-driving vehicles; on a broader level, this includes securing societal and environmental wellbeing (see Guideline 2 and 6). *Right to protection of personal data* drawing attention to data ownership, data governance and privacy of personal data that is generated during the operation of self-driving cars (see Guideline 3). *Right to equality and non-discrimination* requiring the avoidance of unfair bias in operating vehicles as well as the accessibility of benefits for every individual in society (see Guideline 5). *Right to explanation* which, in the field of autonomous driving, demands transparency and communication of the underlying functionality,

which can be achieved through accountability measures such as audits and logging mechanisms (see Guideline 3 and 7). Certainly, incompatibilities and trade-offs between fundamental rights can emerge; for example, life and security can be in tension with the right to self-determination. On the one hand, AVs are expected to improve traffic flow and decrease fatalities that are due to human error. On the other hand, automated driving systems reduce the driver's autonomy, perhaps to the point of being a mere passenger. In this regard, the Ethics Commission on Automated and Connected Driving (BMVI, 2017) formulated the following guideline: "In a free society, the way in which technology is statutorily fleshed out is such that a balance is struck between maximum personal freedom of choice in a general regime of development and the freedom of others and their safety" (Lütge, 2017, p. 550). In conflict situations, policymakers and legislators should decide which fundamental rights are to be prioritized.

Policy Recommendations

- Relevant fundamental rights to be considered in the field of autonomous driving are: human dignity, right to self-determination and liberty, right to life and security, right to protection of personal data, right to equality and non-discrimination as well as the right to explanation.
- It must be realized that there will be no technologies or policies that maximize all fundamental rights for everybody simultaneously. There will always be trade-offs. Therefore, policymakers and legislators should decide which fundamental rights are to be prioritized in particular situations.
- In doing so, policymakers and legislators should cooperate with multiple stakeholders to obtain necessary information for executing an evaluation and subsequent agreement on compromises and prioritization.

Human Agency and Oversight – Including Monitoring, Training, Human-Machine Interfaces and External Control of Vehicle Data

A few guidelines have already been developed that highlight the importance of maintaining personal autonomy in AVs, including possible requirements for a 'stop' or 'override' button (European Commission, 2020b; Lütge, 2017). At the same time, autonomy requires informed and deliberate control, and so overrides (and other measures) should not necessarily be universal. In particular, admissibility of human override should be conditional on two aspects:

1. The level of automation of the AV[2]:
   a. For levels up to and including 3, there should be an override function that can be executed at any time;
   b. For level 4, there should be an override function that can be executed only when not impacting or undermining the safety mechanisms of the AV (e.g., one helpful factor to satisfy this requirement might be to implement overrides with a time lag). The rationale for this is that, if individuals were allowed to intervene immediately at any point, the inherent logic and longer-term plan completion of the technically functional AV is disrupted which may lead to increased risks for all parties involved;
   c. For level 5, it is not necessary to include an override function, as it would take away many of the original advantages such as inclusive accessibility (e.g., by excluding elderly, disabled individuals, youth or individuals who do not hold a driving license), safety (e.g., humans taking control may be out of practice), trust (e.g., giving drivers the impression that the system could fail), and comfort (e.g., limiting opportunities for new and more comfortable mobility options and designs)[3];
2. The state and behavior of the driver:
   a. When the driver's abilities are impaired (e.g., due to alcohol consumption), the availability of an override function should be limited and preceded by a request for confirmation.

Nevertheless, recent examples of AVs involved in crashes draw attention to the failing assumption of responsibility by individuals. The underlying problems relate to overconfidence in, or overreliance on, the AI system as companies do not adequately warn drivers and/or drivers violate the guidelines provided by the companies.

Therefore, companies must clearly distinguish and make apparent whether a driverless system is being used or whether a driver remains accountable for driving (Lütge, 2017). In order to realize effective human agency and clarity over personal responsibility, our approach concerning AVs is threefold:

1. Companies should put in place technical safeguards to help drivers remain fully aware and ready to take over the driving when the AV expects them to. AVs should monitor drivers and help drivers remain awake and attentive. For example, current driving monitoring systems using camera-based facial recognition technology determine the driver's level of vigilance and trigger alerts to the driver when signs of distraction are detected (Research & Markets, 2019). Other monitoring systems are related to the amount of torque in the steering wheel. For example, Tesla (2020) locks the activation of the autopilot mode if the driver seems inattentive (e.g., insufficient torque is applied or warnings are repeatedly ignored). The upcoming regulation on automated lane keeping systems will obligate car manufacturers to introduce driver availability recognition systems and clarify the criteria that assess whether a driver is deemed to be unavailable (e.g., eye closure) (UNECE, 2020b). UNECE also considers that "[a]utomated/autonomous vehicles should include driver engagement monitoring in cases where drivers could be involved (e.g., take-over requests) " (UNECE, 2019, p. 3). It is important that handovers be aligned with the level of automation: As the level of automation increases, drivers engage more in other activities such as watching a video, which decreases human capability to take over control (Merat et al., 2014). Thus, handovers should conform to human capabilities by, for example, obviating "the need for an abrupt handover of control to the driver (emergency)" (Lütge, 2017, p. 556). There is currently no agreement on what constitutes a comfortable transition time, and so we do not propose a universal prescription on this point. In the meantime, companies should provide documentation that justifies their particular handover window. A possible starting point for determining a reasonable transition time might be that AVs, as they drive, could learn about the capabilities of drivers from aggregated traffic data and adjust the vehicle's parameters accordingly (respecting a safe minimum time response).

2. Companies should train drivers on the capabilities and limitations of AVs (European Commission, 2020b), so that individuals can make informed decisions and do not over rely on the vehicle's capabilities (see also UNECE, 2019). This training should be tailored to different demographic groups, given recent studies that show demographic differences in interactions with AVs (Manser et al., 2019). Training programs should cover topics such as the "[system's] functional intent, operational parameters, system capabilities and limitations, engagement/disengagement methods, HMI, emergency fallback scenarios, operational design domain parameters (i.e., limitations), and mechanisms that could alter [the system's] behavior while in service" (NHTSA, 2017, p. 15). Drivers should also be trained on the purpose of using an AVs, the degree of automation, and conditions for potential system failures (Manser et al., 2019).

3. The importance of human autonomy applies not only to drivers but also to humans outside the vehicle such as pedestrians. Therefore, companies should ensure that these latter individuals can also exercise their autonomy. For example, AVs should have mechanisms to show pedestrians that they have been recognized and reveal the AV's motion intentions, perhaps with LED strips to convey perception information (e.g., displaying cool colors for far away obstacles and warm colors for near obstacles in the environment) (Florentine et al., 2016). These external human-machine interfaces facilitate human agency for pedestrians, as they enable them to feel less

anxious about the technology and have more information to move freely and safely. However, further research is required to determine the most useful interfaces (Rouchitsas & Alm, 2019).

Additionally, external oversight mechanisms need to be put in place to control for adequate human agency. Therefore, although internal overriding functions may not always or immediately be available for (drivers in) AVs, general oversight should be possible at all times. Live and total oversight is both impracticable and unwarranted (Lütge, 2017). However, under certain circumstances, such as following a fatal accident, and depending on the legal and regulatory framework in place in the country where the accident occurred, it may be appropriate to designate an organization in each jurisdiction that is permitted to retrospectively look at the code and data within the AV to determine the cause of the accident (for more information see Guideline 7).

Industry Recommendations

- There should be a conditional override option allowing the control to be handed back to the driver. The admissibility of an override function depends on the level of automation of the AV (up to level 3: at any time; level 4: corresponding to safety mechanisms of an AV; level 5: not required) as well as on the state and behavior of the driver (e.g., impaired ability).
- AVs should continuously assess and monitor the driver's attentiveness and ability to intervene. Before operation, the AV could pose control questions to the driver (e.g., did you ingest any drugs or alcohol?); during operation, the AV could use sensors and biometric technology to do so. The upcoming UN Regulation on Automated Lane Keeping Systems can serve as a baseline for car manufacturers to develop appropriate driver attentiveness recognition systems.
- Handover should correspond to the driver's capabilities. Therefore, AVs could learn about drivers' capabilities and response times during operation from aggregated data and adjust the vehicle's parameters accordingly (respecting a safe minimum time response).
- Companies should provide documentation that justifies their particular handover window.
- Training programs should be tailored to different demographic groups and exhibit minimum elements that should be regarded in a training curriculum (e.g., limitations and capabilities of AVs) based on findings of recent studies.
- AVs should offer a 'training mode', for the first kilometers to train drivers on the AV's functioning.
- External human-machine interfaces should clearly communicate about the vehicle's motion intention and awareness of other traffic participants to humans outside the vehicle.

Policy Recommendations

- Policymakers should finalize what constitutes acceptable and legitimate override functions and define applicable situations for activation.
- Policymakers should determine standards for drivers' monitoring, training requirements, handover routines and external human-machine interfaces. These standards should be as global as possible.
- Policymakers in each jurisdiction should consider designating an organization in each jurisdiction that is allowed to look at the code and data within the AV in the event of a fatal accident involving an AV or a corresponding legal proceeding.

Technical Robustness and Safety – Including Resilience to Attack and Security, Fallback Plan and General Safety, Accuracy and Reliability

A prime requirement of AVs should be safety, both in ordinary operations and if subject to adversarial attack (Lütge, 2017).

There are many differing forms of potential threats to AVs, and so governmental entities such as the ENISA (2019) or UNECE (2020a) have created holistic summaries and categorizations of relevant dangers and vulnerabilities. Firstly, there are threats that do not solely apply to AVs but also to conventional vehicles such as *technical malfunctions and outages* including sensor and other failures (ENISA, 2019). Secondly, there are threats that are particularly important for AVs and can be subsumed under the term 'cybersecurity'. Potential cybersecurity threats include the following:

- *Hijacking* such as unauthorized information disclosure or extraction of copyrighted or proprietary software from vehicle systems (product piracy) (UNECE, 2020a);
- *Abuse* such as attacks on back-end servers that stops the vehicle's functioning (e.g., disruptions of communication and external connectivity) or threats regarding the vehicle's update procedures (e.g., preventing the rollout of critical software updates) (UNECE, 2020a);
- *Passive behavioral attacks* such as individuals intentionally interfering with AVs. For example, human drivers might tend to drive more aggressively around AVs or jaywalking may increase because it is known that AVs respect the safety distance.

There are several categorizations of threats that relate to the data stored in vehicles on an associated server and to the information exchanged during communication between the vehicle and the server. These threats can impact the safe operation of the vehicle, alter the software operation, and generate data breaches, though many of these threats are not specific to AVs but also can be found in current vehicles.

It is essential to develop mechanisms to test an AV's cybersecurity management system before operation. The EU Cybersecurity Act aims to establish a general certification framework for ICT digital products, services, and processes that allows the "creation of tailored and risk-based EU certification schemes" (ECCG, 2020). Similarly, the UN is preparing a regulation on uniform provisions concerning the approval of vehicles with regard to cybersecurity and of their cybersecurity management systems. For example, the draft regulation (as of March 2020) proposes an international approval mark or the verification of a manufacturer's compliance by an approval authority (UNECE, 2020a). In the future, such clear regulations and standardized tests will be necessary so that all companies are informed about, and comply with, the universal requirements for cybersecurity management systems. Governments should "promote mutual recognition systems and certification schemes that are built upon international standards [...] to facilitate international harmonization on privacy and security" (Joaquin Acosta, 2019, p. 215). SAE J3061, a comprehensive cybersecurity implementation guideline for the automotive industry, can serve as a starting point (SAE International, 2016).

Additional to threats, measures need to be developed that assess the general functionality of an AV. The Ethics Commission on Automated and Connected Driving suggests that "[t]he public sector is responsible for guaranteeing the safety of the automated and connected systems introduced and licensed in the public street environment. Driving systems thus need official licensing and monitoring" (Lütge, 2017, p. 550). For example, a kind of TÜV, i.e. a technical inspection agency, for AVs could be developed. Relevant factors to be assessed here are accuracy, reliability and fallback options of AVs. In terms of accuracy and reliability, it could be tested to what extent the AV's underlying "AI meets, or exceeds, the performance of a competent & careful human driver", refrains from engaging in "careless, dangerous or reckless driving behavior" as well as to what extent it "remains aware, willing and able to avoid collisions at all times" (ADA, 2020). SAE International published a more detailed and elaborate list of driving safety performance assessment metrics such as minimum safe distance factors or proper responses (Wishart et al., 2020). Furthermore, safeguards against technical failures and outages need to be established. The IEEE P7009 standard for fail-safe design of autonomous and semi-autonomous systems could serve as a baseline for developers. The standard provides clear procedures for measuring, testing, and certifying a system's ability to fail safely as well as instructions for improvement in the case of unsatisfactory performance (IEEE, 2019).

In terms of general safety and fallback plans, "[i]n emergency situations, the vehicle must autonomously, i.e., without human assistance, enter into a 'safe condition'" (Lütge, 2017, p. 556). This condition has been specified by proposing the terms 'minimal risk condition' and 'minimum risk maneuver'. The Minimal risk condition is "[a] condition to which a user or an ADS may bring a vehicle after performing the DDT [dynamic driving task] fallback in order to reduce the risk of a crash when a given trip cannot or should not be completed" (SAE International, 2018, p. 11). The "Minimum risk maneuver means a procedure aimed at minimizing risks in traffic, which is automatically performed by the system" (Leonhardt, 2018, p. 12). Causes for the execution of such a maneuver could be detection that the driver is inactive and not reacting to transition demands, or reaching system failure / boundaries when the driver is not responding to transition demands. In such situations, potential maneuvers could entail "further lane keeping for a certain time, enlarging gap to other road users, […] slowing down to standstill" (BMVI, 2015, p. 4). What constitutes an appropriate maneuver depends on (1) the operation condition of the vehicle (e.g., technical failures that hinder the AV to perform a fallback), (2) the prevailing environmental conditions (e.g., density of traffic) and (3) regulatory boundary conditions (Leonhardt, 2018). Although SAE J3016 (Leonhardt, 2018) makes significant progress regarding the nature of a safe or minimal risk condition, the definition of such conditions as well as the particular circumstances in which such conditions should be activated (e.g., incidents that leave the driver incapacitated such as a stroke) need to be further determined and harmonized.

Overall, experimenting with new AVs and testing their technical robustness and safety should follow a stepwise approach: For example, "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" should be identified and standardized (European Commission, 2020a, p. 29). Recognizing the challenges of physical test strategies for AVs (length of time they take to complete, high number of hours of drive time required), ESTECO has developed a white-box / scenario-based verification system to investigate the performance of ADAS/AD functions across different sensors, algorithms, actuation and scenarios (ESTECO, 2020). Systems like these can act as helpful antecedents to actual testing on open roads.

Industry Recommendations

- The prime requirement of AVs should be safety.
- In addition to threats that relate to conventional vehicles, manufacturers of AVs should particularly focus on cybersecurity threats. In doing so, companies need to comply with regulations for cybersecurity management systems. SAE J3061 could serve as a guideline to design cybersecurity into AVs throughout the entire development life cycle process.
- In terms of general functionality and safety, vehicles need to pass an official test that assures the system's accuracy, reliability and adequacy of its fallback options. The SAE Driving Safety Performance Assessment Metrics and the IEEE P7009 standard could serve as a baseline to design fail-safe mechanisms of autonomous and semi-autonomous systems.

Policy Recommendations

- Regulations need to be developed that reflect consensus on the method by which to grant approval to a vehicle's cybersecurity management system.
- Policymakers need to work with industry experts to develop a standardized test for the general functionality and safety of AVs to assure the system's accuracy, reliability and adequacy of its fallback options. This test could serve as a basis for the approval of AVs for sale to consumers.
- Policymakers need to collaborate with industry experts to determine and harmonize the definition of a 'safe condition' / 'minimal risk condition', the corresponding 'minimum risk maneuvers',

and the circumstances in which such maneuvers should be executed. In doing so, SAE J3016 could serve as a baseline.

Privacy and Data Governance – Including Respect for Privacy, Transparency and Communication, and Access to Data

Conventional vehicles collect data through event data recorders (that record technical information about a vehicle's operation involved in crashes) and on-board diagnostic information (to access information about driver behavior, emission measures or diagnose performance issues). With new technological options, connected vehicles and AVs will make transportation safer and more convenient. However, many features depend on the collection and processing of ever more data in order to function effectively (Future of Privacy Forum, 2017).

Therefore, it is essential to specify the type and scope of data that AVs are permitted to collect. Three types of data can be distinguished that warrant special attention:

- *Geolocation data* (e.g., for activating route navigation), which could reveal the passenger's location and life habits of individuals (EDPB, 2020).
- *Biometric data* (e.g., for user recognition or tracking of driver's attention), which could be used to enable unauthorized access to a vehicle and enable access to a driver's profile settings and preferences (EDPB, 2020). The collection of this type of data applies not only to drivers but also to individuals outside the vehicle such as pedestrians.
- *Driver behavior data*, which could reveal unlawful behavior, including traffic violations such as speeding (EDPB, 2020).

Some of this data will be collected automatically, and some will require consent from the vehicle owner or driver in order to activate and use certain functions. Careful consideration needs to be given to the collection of data from inside the vehicle that relates to things other than the operation of the vehicle. Additionally, individual's rights should be considered at group level (e.g., drivers versus pedestrians) (European Commission, 2020a). For example, data (especially, biometric data) of external parties such as individuals walking on the street should warrant special protection. Although the European Commission (2020a) has additionally problematized data collection when AVs pass through particular locations such as private and non-public settings, we suggest that collecting data in private spaces should not in general be restricted in order to guarantee an AV's functionality. The focus should instead be on the mode of data collection and sharing.

Overall, services that collect and share data should comply with all applicable laws, and be accompanied by a strict privacy and data governance policy that includes, but is not limited to, the following (Future of Privacy Forum, 2017):

1. **Transparency and Communication:** Manufacturers need to provide clear and concise privacy policies to the vehicle owners that describe data collection and use. These policies must be readily understood by vehicle owners. These policies could, for example, be displayed in the purchase agreement, user manual or in the interface of an app. This is also in line with the General Data Protection Regulation (GDPR) stating that controllers must, before personal data is obtained, provide the data subjects with information necessary to ensure transparent processing about the existence of automated decision-making.
2. **Affirmative and Explicit Consent:** The driver's educated and affirmative consent is required before certain sensitive data is collected or used. This requirement is particularly critical for marketing uses, or if the data will be shared with unaffiliated third parties. This is in line with the guideline of the Ethics Commission on Automated and Connected Driving about the permissibility to use data that is generated by AVs for other business models, which states that lastly "[i]t is

the vehicle keepers and vehicle users who decide whether their vehicle data that are generated are to be forwarded and used" (Lütge, 2017, p. 555). Additionally, even in the absence of laws requiring it, users should always have the right to opt-out or request that particular data not be collected, unless those data are critical for the AV safe system's operation.

3. **Limited and Useful Sharing With Third Parties:** There should only be limited circumstances where manufacturers are allowed to share a vehicle's data with external parties. Under appropriate conditions and with the appropriate safeguards, data that guarantees safe operation of the vehicle and other traffic participants as well as data that provides benefits to overall society and is of public interest should be shared. For example, AVs could provide information to the local department of transportation about a pothole on the road, so that infrastructure inspections and maintenance resources can be better allocated (in consideration of a fair and unbiased distribution of resources) and traffic information can be shared to improve traffic flow and promote safety. Accordingly, the European Commission has issued a regulation requiring public or private road operators and service providers to share and exchange relevant road safety-related traffic data such as the observation of a temporary slippery road or exceptional weather conditions (European Commission, 2013). Personally identifiable information must always be given the highest levels of protection. If data must be shared with third parties due to the above mentioned reasons, they should be anonymized and deidentified before being transmitted (EDPB, 2020). For example, the EU Data Task Force partnered with TomTom to improve road safety by sharing anonymized vehicle and infrastructure data between countries and manufacturers. For example, this will allow the detection of dangerous road conditions such as slippery roads and issue warnings to other traffic participants. "The EU Data Task Force (DTF) will use a decentralised data collaboration architecture to share vehicle-generated data […]. The datasets will then be taken by TomTom, processed further, and delivered back to other vehicles and road authorities via its live Traffic services" (Europawire, 2019). In line with Article 3(c) of Directive 2010/40/EU, data and procedures for the provision of road safety-related traffic information should be free of charge to users (European Commission, 2013). However, past studies showed individuals can sometimes be identified using anonymized data (Archie et al., 2018; Techcrunch, 2006), and so companies must ensure that the shared data does not permit re-identification (e.g., by minimizing collected data or using differential privacy techniques).

4. **Compliance With Pertinent Data Protection Standards and Regulations:** All data collection and processing obviously must respect relevant regulations (EDPB, 2020), such as the GDPR that applies to the processing of personal individual data, as well as the ePrivacy directive for information access on the terminal equipment of a user (EDPB, 2020; European Commission, 2020b). The IEEE P7002 standard specifies how to manage privacy issues for systems that collect personal data, e.g., by providing a guideline for a privacy impact assessment (IEEE, 2019).

Industry Recommendations

●   Manufacturers should follow a strict privacy and data governance policy that includes transparency and communication to users, requesting affirmative consent and allowing limited sharing of data with third parties (including governments). In doing so, companies should comply with applicable standards and regulations such as the GDPR, the ePrivacy directive or the IEEE P7002.
●   Before transmitting personal information from an AV to third parties, steps must be taken to ensure that it cannot be traced back to an individual.
●   Manufacturers should implement data protocols defining who can have access to data under which conditions.

Policy Recommendations

- Before receiving AV data, policymakers need to make clear what types of AV data they are seeking and how that data will enable them to improve public safety or some other legitimate public purpose (e.g., improve infrastructure, traffic flow and law compliance).
- At the EU level, building on Article 3(c) of Directive 2010/40/EU, consideration should be given to expanding the list of events and relevant traffic information that should be communicated free of charge.

Transparency – As A Key Mechanism to Realize all Other Requirements

In the automotive sector, we contend that transparency is not a freestanding desideratum, but rather a key mechanism to realize the other principles or requirements. Transparency plays a major role for achieving the principle of privacy and data governance, requiring that manufacturers provide vehicle owners with information regarding data collection practices and intended uses (for more information see Guideline 3). Similarly, to satisfy the principle of accountability, the implementation of explicit transparency measures such as logging mechanisms or black boxes are essential (for more information see Guideline 7). The IEEE P7001 ("Transparency of Autonomous Systems") standard can serve as a baseline to address these issues.

Diversity, Non-Discrimination and Fairness – Including the Avoidance of Unfair Bias, Responsible Balancing of Risks and Accessibility

Generally and regarding the operations of AVs, no distinction between individuals should be allowed and fair treatment of all humans should be enacted. This is clearly stated in the Universal Declaration of Human Rights: "[e]veryone is entitled to all the rights […] without distinction of any kind, such as race, colour, sex, language, religion, political or other opinion, national or social origin, property, birth or other status" (United Nations, 1948, p. 2; Kriebitz & Lütge, 2020). In the field of AI (e.g., AVs), this obligation becomes ever more important as implicit biases and discrimination may unintentionally, and without transparency, be incorporated into algorithms. Studies show that systems can have differential performance for people of different ethnic groups, which consequently can result in them being harmed. For example, a study from the Georgia Institute of Technology illustrates how state-of-the-art AI object detection systems are less likely to detect pedestrians with darker skin color than those with lighter skin (Wilson, Hoffman & Morgenstern, 2019). Another study from the US National Institute for Transportation and Communities investigated the driving behavior through crosswalks that "revealed that Black pedestrians were passed by twice as many cars and experienced wait times that were 32% longer than White pedestrians" (Goddard, Kahn, & Adkins, 2015, p. 2). If such driving data is fed into a machine-learning algorithm, the system may discover this discriminatory pattern and adapt it into its functioning (Forbes, 2020).

In order to ensure non-discriminatory programming and functioning, the systems need to be trained and tested for unfair bias. Companies should test their algorithms for bias and discrimination and demonstrate that certain fairness standards are met (Vox, 2019). Laws could be enacted, for example, that mandate that facial recognition software used by public entities and companies must be independently tested for bias (Secretary of State Washington, 2020). The IEEE P7003 standard for algorithmic bias considerations sets out instructions for eliminating bias when developing algorithms: it provides developers of algorithms for autonomous systems with protocols and includes criteria for selecting validation data sets (IEEE, 2019). The training should be different depending on the location where the system operates: when a technology is launched into the market, companies could localize it using location specific data. Companies could ensure that their development teams are sufficiently diverse to guard against intentional and implicit bias being incorporated into their algorithms and technologies (Vox, 2019).

Past literature has extensively debated dilemma situations (e.g., unforeseen and unavoidable accidents) with reference to the famous trolley cases. The ideal is to avoid such situations in which accidents are unavoidable in the first place; for example, the lateral position of AVs on a lane can be adjusted to tune the risk posed to all other traffic participants (e.g., how much room should be given to a bicyclist?). Therefore, we argue to move away from the debate around dilemma situations. Instead, a responsible balancing of risk or estimated harm should be permitted for AVs at all times. This balancing decision should not be based on personal features of individuals such as age or gender (Lin, 2016; Lütge, 2017). Instead, as the severity of injury increases in proportion to the kinetic energy, estimated harm could be quantified and balanced by more objective factors such as the type or speed of particular traffic participants and the impact angle under which a collision would occur (Geißlinger et al., 2020). Taking into account the type of road users would grant vulnerable traffic participants (e.g., pedestrians or cyclists) the same level of protection as other road users (European Commission, 2020a). Overall, the consideration of these factors could help achieve a "[g]eneral programming to reduce the number of personal injuries" (Lütge, 2017, p. 552).

Besides the unbiased vehicle's internal functioning, AVs should be human-centric (European Commission, 2020a). In particular, AVs should be equally usable for and accessible to all individuals, which requires a non-discriminatory design. For example, age or the presence of a disability is not always considered by automotive companies, leading to potential issues of discrimination. Therefore, levels of differing abilities need to be acknowledged (e.g., a young individual may have quicker reflexes for executing requests than elderly people) and the systems need to be adapted accordingly for different users, so that everyone can benefit from this new technology (for more information see also Guideline 1).

Industry Recommendations

- Companies should test their vehicle's AI systems for unfair performance differences across skin tone, gender, age and other characteristics. The IEEE P7003 standard can serve as a baseline to address and eliminate issues of bias in the creation of algorithms.
- When a technology is launched into the market, companies should localize it using data and train the model using multiple diverse data sets that are location specific.
- The AI developing team should be as inclusive as possible to include the broadest group possible in terms of demographics such as ethnicity.
- A responsible balancing of risks and estimated harm to reduce the number of personal injuries should be permitted for AVs without discriminating against personal characteristics. Instead, factors underlying the balancing could include the type or speed of particular traffic participants and the impact angle under which a collision would occur.
- The personalization of AVs should be accessible by design and as inclusive as possible (e.g., disabilities included). Before an AV is released onto the streets, companies should demonstrate their plans and actions that ensure customizing-options to their vehicles (e.g., possibility to take away seats or include a ramp for entering the vehicle with a wheelchair).

Policy Recommendations

- Consideration should be given to having ethics standards boards test and assess that the systems for AVs are working properly, fairly and in an unbiased manner.
- Consideration should be given to requiring carmakers to explain the procedures they have put in place to make their designs accessible and avoid biases before granting them authorization to sell their vehicles to the public.

Societal and Environmental Wellbeing – Including Sustainability and Environmental Friendliness and Social Impact

In terms of societal and environmental wellbeing, the Sustainable Development Goals adopted by all United Nations Member States can serve as a reference point. Goal 3 (to "[e]nsure healthy lives and promote well-being for all at all ages") and goal 11 (aiming to "[m]ake cities and human settlements inclusive, safe, resilient and sustainable") are particularly relevant to this topic (United Nations, 2015, p. 14). Companies and policymakers in the automotive sector should focus on meeting the following objectives:

1. **Increased Public Health and Mobility:** AVs can improve society's health by avoiding fatalities that are due to human error (Bartneck et al., 2019). This is in line with Vision Zero, which states that eventually no one will and shall be killed or seriously injured within the road transport system (Ministry of Transport and Communications, 1997). The introduction of AVs could offer greater mobility solutions for a major part of society that is mobility-impaired, whether the elderly, young (without a driving license) or those who were otherwise unable to drive (BCG, 2017; WEF, 2018). This could positively affect mental health (e.g., due to feeling less dependent on others) and create a more inclusive society (Lim & Taeihagh, 2018). These benefits, however, can only be realized if safety and diversity standards are adhered to (for more information see Guideline 2 and 5);

2. **Better Traffic Flow:** AVs could reduce congestion and delays (e.g., during peak hours) and improve traffic flows and efficiency, especially when combined with shared mobility options. For example, using a traffic simulation model for Boston, it was found that the simulations that had included AV technology yielded less congestion, shorter travel times and more street space and (BCG, 2017; WEF, 2018). These benefits stem mostly from AVs' connectivity to external communication networks so that data can be managed and distributed in real time enabling methods such as platooning (Lim & Taeihagh, 2018). However, if not managed properly, it could also increase traffic flow and generate inefficiencies of uncoordinated traffic (Joaquin Acosta, 2018a). Proactive measures such as adopting a fitting physical and digital infrastructure, could improve the existing traffic situation by at least 15-20% (Inframix, 2020);

3. **Decreased Carbon Emissions:** Widespread adoption of AVs could reduce environmental degradation through reduced emissions and energy consumption (BCG, 2017). This is especially true if unnecessary acceleration and braking is reduced (Lim & Taeihagh, 2018). A concrete action point for companies would be to design AI systems that reduce vehicles' CO2 emissions. For example, companies could offer by default an eco-driving mode with a speed average that minimizes emissions and avoids unnecessary acceleration or braking. Many of the benefits relating to the reduction of carbon emission can be achieved by combining AVs with other disruptive technologies such as the electrification of vehicles (BCG, 2020). In addition, promoting AV shared mobility could "lessen the environmental impact of passenger vehicles by decreasing the number of vehicles on the road" (Joaquin Acosta, 2018a, p. 3). A concrete action point for policymakers would be to facilitate research and development for solutions to combine AVs with other disruptive technologies (e.g., electrification, shared mobility).

While these potential benefits are substantial, there is also significant uncertainty about the net impact of introducing AVs. Many measures of benefits focus on improvements per vehicle-mile traveled (VMT). However, the increased mobility and convenience benefits will potentially lead to significant increases in VMTs, potentially leading to increased total pollution, congestion, and so forth, despite the per-VMT gains (Geary & Danks, 2019). Thus, as technology continually develops, companies and policymakers in the automotive sector should follow a stepwise implementation process to ensure that introduction of AVs provides net benefits. Moreover, this implementation process must

be combined with a simultaneous adaption of infrastructure (physical and digital). "Needed structural improvements include dedicated lanes to separate AVs from other traffic, and sensors to enable self-driving cars to communicate with their environment" (BCG, 2020). Otherwise, if AVs enter traffic in an uncoordinated way and without a fitting infrastructure, traffic flow and other benefits may be degraded. Several projects of the EU Horizon 2020 program have been focusing on this challenge (e.g., CoEXist or Inframix) (European Commission, 2019).

City planners, road operators and local authorities should use the findings of such projects to make informed decisions on where to roll out new mobility models and how to update their road network accordingly. Collaboration with multiple private-sector leaders and national agencies is key to fostering innovation and progress: "the success of AMoD [autonomous mobility on demand] will depend to a large extent on establishing close partnerships among mobility providers, infrastructure companies, and city authorities" (BCG, 2020).

Industry Recommendations

- When developing their products, automotive companies should consider integrating and providing benefits of increased public health and mobility, better traffic flow and decreased carbon emission.
- Manufacturers should offer by default an eco-driving mode with a speed average that avoids unnecessary acceleration or braking and thus reduces carbon emissions.
- When developing AVs, car manufacturers should try to integrate other disruptive technologies such as electrification and shared mobility.

Policy Recommendations

- Policymakers should follow a stepwise implementation process and concentrate on mixed traffic scenarios. Policymakers should promote the integration of AVs in existing transport systems instead of competition between them, for example, by prioritizing research and development of AV solutions for public and shared mobility.
- A simultaneous adaption of physical and digital infrastructure is essential (e.g., lanes that separate AVs from other traffic).
- In doing so, collaboration with multiple actors such as private-sector leaders and national agencies is key to fostering innovation and progress (e.g., make use of projects investigating differing mobility models).

Accountability – Including Auditability, Measures of Transparency, Reporting of Negative Impact, and Redress

The attribution of liability and responsibilities for AVs is a challenging issue. "The first step towards the creation of a culture of responsibility is the study, deliberation and agreement on the different responsibilities of different stakeholders" (European Commission, 2020a, p. 56). In case of accidents, the AV itself cannot be held morally accountable since it is not considered a moral agent (Gogoll & Müller, 2017). Responsible parties will instead be manufacturers, component suppliers, technology companies, infrastructure providers or car holders / drivers. Therefore, policymakers should clarify the concept of a producer as well as review regulations on product liability (e.g., European Commission, 2018). This will, of course, vary depending on the motor vehicle laws in place in different countries. When adapting existing regulations to AVs, regulators may have to choose between different liability regimes for different situations and levels of automation. For example, strict liability concepts may mean that for AVs, the manufacturer can be held liable if the automated mode is switched on, whereas, if not, the driver is considered liable. On the other hand, one could argue that liability should move gradually from one actor to the next (e.g., from the car manufacturer to the driver) depending on the

driver's level of autonomy and solo action. For guidance, regulators could look to Law Labs (Joaquin Acosta, 2018b). Law Labs are a proposed concept to experiment with different regulatory approaches for a given innovation (e.g., AVs), similar to how regulatory sandboxes experiment with innovations in controlled environments (operating under temporary regulatory exemptions). For example, traffic rules could be revised and it could be investigated under which circumstances AVs are allowed to not to comply with a traffic rule (European Commission, 2020a).

In order to provide clarity about the causes of accidents, companies in the automotive sector may want to consider the following issues.

Regularly Conduct Internal and External Audits

In terms of internal audits, manufacturers should execute *continuous optimization and tests*. This is in line with the guidelines of the Ethics Commission on Automated and Connected Driving, which state that "[l]iability for damage caused by activated automated driving systems is governed by the same principles as in other product liability [...:] manufacturers or operators are obliged to continuously optimize their systems and also to observe systems they have already delivered" (Lütge, 2017, p. 553). In doing so, companies should conduct a *risk assessment* by listing factors that may increase risk and uncertainty regarding a vehicle's operation and by proactively implementing appropriate countermeasures. Risks may stem from the vehicle's technology (e.g., technical failure to transmit sensor data), the actions of other traffic participants (e.g., disobeying traffic law such as jaywalking), external circumstances (e.g., the state of the road, weather conditions) or the vehicle's driving behavior (e.g., speed).

In addition to adhering to internal standards and audit requirements, external test centers could perform *conformity assessments* and grant *certifications* since "independent assessment will increase trust and ensure objectivity" (European Commission, 2020b, p. 25). Common audit areas for a certification system are – similar to the AI4People requirements – autonomy and control, fairness, transparency, reliability, security and data protection. Expected benefits of an AI certification would be trust in the application, orientation for customers and developers, fulfillment of norms such as cybersecurity and data security and comparable market equality (IAIS, 2019).

Implement Explicit Measures of Transparency

These transparency measures should pertain to the development as well as the operation of AVs. Before operation, during the development phase, companies should retain records and data including data sets used for training (e.g., selection process) and document the programming and training methodologies. This is particularly important if authorities seek to review the underlying logic of a system and inspect relevant documentation (European Commission, 2020b). Also, during operations, relevant information should be recorded through logging mechanisms and black boxes integrated into AVs (Lütge, 2017). Regarding the black box, companies could consider an event data recorder and data storage system for AVs that record data of "the system status, occurrence of malfunctions, degradations or failures in a way that can be used to establish the cause of any crash and to identify the status of the automated/autonomous driving system and the status of the driver" (UNECE, 2019, pp. 3-4). These measures will ensure that the functioning and actions of AVs are explainable in retrospect. Overall, "[i]nternational standardization of the [...] documentation (logging) is to be sought in order to ensure the compatibility of the logging or documentation obligations as automotive and digital technologies increasingly cross national borders" (Lütge, 2017, p. 555). For example, the upcoming regulation on automated lane keeping systems will determine what events are recorded by data storage systems for AVs (e.g., emergency maneuvers, failures) (UNECE, 2020b). IEEE P7001 provides such a standard for the transparency and accountability of autonomous systems so that the reasons why a technology makes certain decisions can be determined (IEEE, 2019). Similarly, SAE J3197 aims to govern data element definition, to provide a minimum data element set and common data output formats for an automated driving system data logger (SAE International, 2020).

Finally, external communication and reporting of performance and negative impact should be regularly required for companies in the automotive sector. Manufacturers and regulators should anticipate that individuals will want explanations when an AV's system did not perform as expected and intended. This is similarly stated in the ethical guidelines for trustworthy AI: "Whenever an AI system has a significant impact on people's lives, it should be possible to demand a suitable explanation of the AI system's decision-making process" (High-level Expert Group on Artificial Intelligence, 2019, p. 18). In California, for example, Transportation Network Companies (TNCs) such as Lyft have to provide the California Public Utilities Commission with reports regarding zero tolerance complaints, violations and collisions of their vehicles on an annual basis (California Public Utilities Commission, 2020). Further information to disclose may be the tradeoffs made within algorithms, the number of past accidents put into context (e.g., relative number of accidents during test drives compared to total number of test drives) as well as safety measures initiated to counteract these accidents. Thereby data, algorithmic and AI literacy is improved (European Commission, 2020a).

Industry Recommendations

- Manufacturers should continuously conduct internal audits (e.g., assessing potential risks to the safe operation of AVs) and subsequently optimize their systems.
- Manufacturers should be transparent about the scope and process of their internal audits and risk assessments (e.g., space of conditions that are checked for).
- The internal audits should be complemented with regular external audits by independent test centers.
- Manufacturers should develop specific measures of transparency. This includes storing records and data of the underlying system logic (e.g., used training data sets) as well as logging mechanisms and black boxes (e.g., an event data recorder and data storage system) that document the actions of / in AVs during operation. The upcoming UN Regulation on Automated Lane Keeping Systems can serve as a baseline for vehicle manufacturers to develop appropriate data storage systems for AVs. SAE J3197 and the IEEE P7001 standard can serve as a baseline to address requirements for transparency and accountability of autonomous systems.
- Companies should transparently communicate and report performance and negative impacts of AVs (e.g., number of collisions, tradeoffs within algorithms).

Policy Recommendations

- Regulators should adapt laws and regulations concerning AVs and liability as the technology continues to develop. Regulators should clarify where responsibility lies in certain situations and ensure that privacy and cybersecurity damages are taken into account.
- Policymakers should consider establishing test centers that regularly request that companies perform conformity assessments and provide certifications.

CONCLUSION

This paper provides practical recommendations for the automotive sector to deal with ethical issues regarding: human agency and oversight, technical robustness and safety, privacy and data governance, transparency, diversity, non-discrimination and fairness, societal and environmental wellbeing as well as accountability. By doing so, this paper distinguishes between policy and industry recommendations in suggesting first steps to be taken by both policymakers and companies. The following list summarizes all recommendations. In the future, we encourage stakeholders in the automotive sector to rely on these recommendations to determine relevant actions and to ensure that AVs comply with ethical principles.

AI4PEOPLE PRACTICAL RECOMMENDATIONS FOR THE AUTOMOTIVE SECTOR

Underlying Fundamental Rights

Policy Recommendations

- Relevant fundamental rights to be considered in the field of autonomous driving are: human dignity, right to self-determination and liberty, right to life and security, right to protection of personal data, right to equality and non-discrimination as well as the right to explanation.
- It must be realized that there will be no technologies or policies that maximize all fundamental rights for everybody simultaneously. There will always be trade-offs. Therefore, policymakers and legislators should decide which fundamental rights are to be prioritized in particular situations.
- In doing so, policymakers and legislators should cooperate with multiple stakeholders to obtain necessary information for executing an evaluation and subsequent agreement on compromises and prioritization.

Human Agency and Oversight

Industry Recommendations

- There should be a conditional override option allowing the control to be handed back to the driver. The admissibility of an override function depends on the level of automation of the AV (up to level 3: at any time; level 4: corresponding to safety mechanisms of an AV; level 5: not required) as well as on the state and behavior of the driver (e.g., impaired ability).
- AVs should continuously assess and monitor the driver's attentiveness and ability to intervene. Before operation, the AV could pose control questions to the driver (e.g., did you ingest any drugs or alcohol?); during operation, the AV could use sensors and biometric technology to do so. The upcoming UN Regulation on Automated Lane Keeping Systems can serve as a baseline for car manufacturers to develop appropriate driver attentiveness recognition systems.
- Handover should correspond to the driver's capabilities. Therefore, AVs could learn about drivers' capabilities and response times during operation from aggregated data and adjust the vehicle's parameters accordingly (respecting a safe minimum time response).
- Companies should provide documentation that justifies their particular handover window.
- Training programs should be tailored to different demographic groups and exhibit minimum elements that should be regarded in a training curriculum (e.g., limitations and capabilities of AVs) based on findings of recent studies.
- AVs should offer a 'training mode', for the first kilometers to train drivers on the AV's functioning.
- External human-machine interfaces should clearly communicate about the vehicle's motion intention and awareness of other traffic participants to humans outside the vehicle.

Policy Recommendations

- Policymakers should finalize what constitutes acceptable and legitimate override functions and define applicable situations for activation.
- Policymakers should determine standards for drivers' monitoring, training requirements, handover routines and external human-machine interfaces. These standards should be as global as possible.
- Policymakers in each jurisdiction should consider designating an organization in each jurisdiction that is allowed to look at the code and data within the AV in the event of a fatal accident involving an AV or a corresponding legal proceeding.

Technical Robustness and Safety

Industry Recommendations

- The prime requirement of AVs should be safety.
- In addition to threats that relate to conventional vehicles, manufacturers of AVs should particularly focus on cybersecurity threats. In doing so, companies need to comply with regulations for cybersecurity management systems. SAE J3061 could serve as a guideline to design cybersecurity into AVs throughout the entire development life cycle process.
- In terms of general functionality and safety, vehicles need to pass an official test that assures the system's accuracy, reliability and adequacy of its fallback options. The SAE Driving Safety Performance Assessment Metrics and the IEEE P7009 standard could serve as a baseline to design fail-safe mechanisms of autonomous and semi-autonomous systems.

Policy Recommendations

- Regulations need to be developed that reflect consensus on the method by which to grant approval to a vehicle's cybersecurity management system.
- Policymakers need to work with industry experts to develop a standardized test for the general functionality and safety of AVs to assure the system's accuracy, reliability and adequacy of its fallback options. This test could serve as a basis for the approval of AVs for sale to consumers.
- Policymakers need to collaborate with industry experts to determine and harmonize the definition of a 'safe condition' / 'minimal risk condition', the corresponding 'minimum risk maneuvers', and the circumstances in which such maneuvers should be executed. In doing so, SAE J3016 could serve as a baseline.

Privacy and Data Governance

Industry Recommendations

- Manufacturers should follow a strict privacy and data governance policy that includes transparency and communication to users, requesting affirmative consent and allowing limited sharing of data with third parties (including governments). In doing so, companies should comply with applicable standards and regulations such as the GDPR, the ePrivacy directive or the IEEE P7002.
- Before transmitting personal information from an AV to third parties, steps must be taken to ensure that it cannot be traced back to an individual.
- Manufacturers should implement data protocols defining who can have access to data under which conditions.

Policy Recommendations

- Before receiving AV data, policymakers need to make clear what types of AV data they are seeking and how that data will enable them to improve public safety or some other legitimate public purpose (e.g., improve infrastructure, traffic flow and law compliance).
- At the EU level, building on Article 3(c) of Directive 2010/40/EU, consideration should be given to expanding the list of events and relevant traffic information that should be communicated free of charge.

Diversity, Non-Discrimination and Fairness

Industry Recommendations

- Companies should test their vehicle's AI systems for unfair performance differences across skin tone, gender, age and other characteristics. The IEEE P7003 standard can serve as a baseline to address and eliminate issues of bias in the creation of algorithms.
- When a technology is launched into the market, companies should localize it using data and train the model using multiple diverse data sets that are location specific.
- The AI developing team should be as inclusive as possible to include the broadest group possible in terms of demographics such as ethnicity.
- A responsible balancing of risks and estimated harm to reduce the number of personal injuries should be permitted for AVs without discriminating against personal characteristics. Instead, factors underlying the balancing could include the type or speed of particular traffic participants and the impact angle under which a collision would occur.
- The personalization of AVs should be accessible by design and as inclusive as possible (e.g., disabilities included). Before an AV is released onto the streets, companies should demonstrate their plans and actions that ensure customizing-options to their vehicles (e.g., possibility to take away seats or include a ramp for entering the vehicle with a wheelchair).

Policy Recommendations

- Consideration should be given to having ethics standards boards test and assess that the systems for AVs are working properly, fairly and in an unbiased manner.
- Consideration should be given to requiring carmakers to explain the procedures they have put in place to make their designs accessible and avoid biases before granting them authorization to sell their vehicles to the public.

Societal and Environmental Wellbeing

Industry Recommendations

- When developing their products, automotive companies should consider integrating and providing benefits of increased public health and mobility, better traffic flow and decreased carbon emission.
- Manufacturers should offer by default an eco-driving mode with a speed average that avoids unnecessary acceleration or braking and thus reduces carbon emissions.
- When developing AVs, car manufacturers should try to integrate other disruptive technologies such as electrification and shared mobility.

Policy Recommendations

- Policymakers should follow a stepwise implementation process and concentrate on mixed traffic scenarios. Policymakers should promote the integration of AVs in existing transport systems instead of competition between them, for example, by prioritizing research and development of AV solutions for public and shared mobility.
- A simultaneous adaption of physical and digital infrastructure is essential (e.g., lanes that separate AVs from other traffic).

- In doing so, collaboration with multiple actors such as private-sector leaders and national agencies is key to fostering innovation and progress (e.g., make use of projects investigating differing mobility models).

Accountability

Industry Recommendations

- Manufacturers should continuously conduct internal audits (e.g., assessing potential risks to the safe operation of AVs) and subsequently optimize their systems.
- Manufacturers should be transparent about the scope and process of their internal audits and risk assessments (e.g., space of conditions that are checked for).
- The internal audits should be complemented with regular external audits by independent test centers.
- Manufacturers should develop specific measures of transparency. This includes storing records and data of the underlying system logic (e.g., used training data sets) as well as logging mechanisms and black boxes (e.g., an event data recorder and data storage system) that document the actions of / in AVs during operation. The upcoming UN Regulation on Automated Lane Keeping Systems can serve as a baseline for vehicle manufacturers to develop appropriate data storage systems for AVs. SAE J3197 and the IEEE P7001 standard can serve as a baseline to address requirements for transparency and accountability of autonomous systems.
- Companies should transparently communicate and report performance and negative impacts of AVs (e.g., number of collisions, tradeoffs within algorithms).

Policy Recommendations

- Regulators should adapt laws and regulations concerning AVs and liability as the technology continues to develop. Regulators should clarify where responsibility lies in certain situations and ensure that privacy and cybersecurity damages are taken into account.
- Policymakers should consider establishing test centers that regularly request that companies perform conformity assessments and provide certifications.

REFERENCES

Archie, M., Gershon, S., Katcoff, A., & Zeng, A. (2018). *De-anonymization of Netflix reviews using Amazon reviews*. Retrieved from https://courses.csail.mit.edu/6.857/2018/project/Archie-Gershon-Katchoff-Zeng-Netflix.pdf

Autonomous Drivers Alliance (ADA). (2020). *ADA Turing test*. Retrieved from: https://ada.ngo/ada-turing-test

Awad, E., Dsouza, S., Kim, R., Schulz, J., Henrich, J., Shariff, A., Bonnefon, J.-F., & Rahwan, I. (2018). The moral machine experiment. *Nature*, *563*(7729), 59–64. doi:10.1038/s41586-018-0637-6 PMID:30356211

Bartneck, C., Lütge, C., Wagner, A., & Welsh, S. (2019). *An introduction to ethics in robotics and AI*. Springer.

Boston Consulting Group (BCG). (2017). *Making autonomous vehicles a reality: Lessons from Boston and beyond*. Retrieved from https://www.bcg.com/de-de/publications/2017/automotive-making-autonomous-vehicles-a-reality.aspx

Boston Consulting Group (BCG). (2020). *Can self-driving cars stop the urban mobility meltdown?* Retrieved from https://www.bcg.com/de-de/publications/2020/how-autonomous-vehicles-can-benefit-urban-mobility

California Public Utilities Commission. (2020). *Required reports TNCs must provide the CPUC*. Retrieved from https://www.cpuc.ca.gov/General.aspx?id=3989

ESTECO. (2020). *Driving change for autonomous vehicles*. Retrieved from https://mcusercontent.com/e18919a10879a5f50c06081a5/files/fb6d6b93-86c0-4bec-9bb0-0161a0629e09/WhitePaper_ADAS.pdf?utm_source=mailchimp&utm_campaign=0300efc2e1f0&utm_medium=page

Eurocybcar. (2019). *Cybersecurity test for cars*. Retrieved from https://eurocybcar.com/en/

Europawire. (2019). *TomTom part of EU Data Task Force's proof of concept to make roads in EU safer*. Retrieved from https://news.europawire.eu/tomtom-part-of-eu-data-task-forces-proof-of-concept-to-make-roads-in-eu-safer-20943847/eu-press-release/2019/06/05/11/18/39/73674/

European Commission. (2013). REGULATION (EU) No 886/2013. *Official Journal of the European Union*. Retrieved from https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32013R0886&from=EN

European Commission. (2018). *Liability for emerging digital technologies*. Retrieved from https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52018SC0137&from=en

European Commission. (2019). *Automated road transport – On the way to connected and automated mobility*. Retrieved from https://ec.europa.eu/inea/sites/inea/files/art_brochure-2019.pdf

European Commission. (2020a). *Ethics of connected and automated vehicles*. Retrieved from https://ec.europa.eu/info/sites/info/files/research_and_innovation/ethics_of_connected_and_automated_vehicles_report.pdf

European Commission. (2020b). *On artificial intelligence – A European approach to excellence and trust*. Retrieved from https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf

European Cybersecurity Certification Group (ECCG). (2020). *The EU cybersecurity certification framework*. Retrieved from https://ec.europa.eu/digital-single-market/en/eu-cybersecurity-certification-framework

European Data Protection Board (EDPB). (2020). *Guidelines 1/2020 on processing personal data in the context of connected vehicles and mobility related applications*. Retrieved from https://edpb.europa.eu/sites/edpb/files/consultation/edpb_guidelines_202001_connectedvehicles.pdf

European Union Agency for Cybersecurity (ENISA). (2019). *ENISA good practices for security of mart cars*. Retrieved from https://www.enisa.europa.eu/publications/smart-cars/at_download/fullReport

Federal Ministry of Transport and Digital Infrastructure (BMVI). (2015). *Minimum risk manoeuvres*. Retrieved from https://wiki.unece.org/download/attachments/27459841/ACSF-04-07%20%20%28D%29%20-%20ACSF-MRM.pdf?api=v2

Federal Ministry of Transport and Digital Infrastructure (BMVI). (2017). *Ethics Commission automated and connected driving*. Retrieved from https://www.bmvi.de/SharedDocs/EN/publications/report-ethics-commission.pdf?__blob=publicationFile

Florentine, E., Ang, M. A., Pendleton, S. D., Andersen, H., & Ang, M. H. Jr. (2016). Pedestrian notification methods in autonomous vehicles for multi-class mobility-on-demand service. In *Proceedings of the Fourth International Conference on Human Agent Interaction* (pp. 387-392). New York, NY: Association for Computing Machinery. doi:10.1145/2974804.2974833

Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Lütge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P., & Vayena, E. (2018). AI4People – An ethical framework for a good AI society: Opportunities, risks, principles, and recommendations. *Minds and Machines*, *28*(4), 689–707. doi:10.1007/s11023-018-9482-5 PMID:30930541

Forbes. (2020). *Overcoming racial bias in AI systems and startlingly even in AI self-driving cars*. Retrieved from https://www.forbes.com/sites/lanceeliot/2020/01/04/overcoming-racial-bias-in-ai-systems-and-startlingly-even-in-ai-self-driving-cars/#2b1cc433723b

Fraunhofer Institute for Intelligent Analysis and Information Systems (IAIS). (2019). *Trustworthy use of artificial intelligence: Priorities from a philosophical, ethical, legal, and technological viewpoint as a basis for certification of artificial intelligence*. Retrieved from https://www.iais.fraunhofer.de/content/dam/iais/KINRW/Whitepaper_Thrustworthy_AI.pdf

Future of Privacy Forum. (2017). *Data and the connected car*. Retrieved from https://fpf.org/wp-content/uploads/2017/06/2017_0627-FPF-Connected-Car-Infographic-Version-1.0.pdf

Geary, T., & Danks, D. (2019). Balancing the benefits of autonomous vehicles. In *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society* (pp. 181-186). New York, NY: Association for Computing Machinery. doi:10.1145/3306618.3314237

Geißlinger, M., Poszler, F., Betz, J., Lütge, C., & Lienkamp, M. (2020). *Autonomous driving ethics: From Trolley problem to ethics of risk*. Working paper.

Goddard, T., Kahn, K. B., & Adkins, A. (2015). Racial bias in driver yielding behavior at crosswalks. *Transportation Research Part F: Traffic Psychology and Behaviour*, *33*, 1–6. doi:10.1016/j.trf.2015.06.002

Gogoll, J., & Müller, J. F. (2017). Autonomous cars: In favor of a mandatory ethics setting. *Science and Engineering Ethics*, *23*(3), 681–700. doi:10.1007/s11948-016-9806-x PMID:27417644

High-level Expert Group on Artificial Intelligence. (2019). *Ethics guidelines for Trustworthy AI*. Retrieved from https://ec.europa.eu/futurium/en/ai-alliance-consultation/guidelines#Top

IEEE. (2019). *Ethically aligned design – A vision for prioritizing human well-being with autonomous and intelligent systems*. Retrieved from https://standards.ieee.org/content/dam/ieee-standards/standards/web/documents/other/ead1e.pdf

Inframix. (2020). *Expected impact – a step by step introduction of automation*. Retrieved from https://www.inframix.eu/expected-impact/

International, S. A. E. (2018). *J3016 – Taxonomy and definitions for terms related to driving automation systems for on-road motor vehicles*. Retrieved from https://saemobilus.sae.org/content/J3016_201806

International, S. A. E. (2020). *J3197 – Automated driving system data logger*. Retrieved from https://www.sae.org/standards/content/j3197_202004/

Joaquin Acosta, A. (2018a). Autonomous vehicles: A smart move? 24 essentials of a SWOT analysis policymakers need to consider. *Berkman Klein Center for Internet and Society at Harvard University*. Retrieved from https://cyber.harvard.edu/sites/default/files/2018-07/2018-07_AVs02_0.pdf

Joaquin Acosta, A. (2018b). Autonomous vehicles: 3 practical tools to help regulators develop better laws and policies. *Berkman Klein Center for Internet and Society at Harvard University*. Retrieved from https://cyber.harvard.edu/sites/default/files/2018-07/2018-07_AVs04_1.pdf

Joaquin Acosta, A. (2019). IoT international regulator challenges: The European approach. In C. H. Cwik, C. A. Suarez, & L. L. Thomson (Eds.), *The internet of things (Iot): Legal issues, policy, and practical strategies* (pp. 191–215). American Bar Association.

Kochupillai, M., Lütge, C., & Poszler, F. (forthcoming). Programming away human rights and responsibilities? The Moral Machine Experiment and the need for a more 'Humane' AVs Future. *NanoEthics*.

Kriebitz, A., & Lütge, C. (2020). Artificial intelligence and human rights: A business ethical assessment. *Business and Human Rights Journal*, *5*(1), 84–104. doi:10.1017/bhj.2019.28

Leonhardt, T. (2018). *Minimal risk maneuver*. Retrieved from https://www.ko-haf.de/fileadmin/user_upload/media/abschlusspraesentation/12_Ko-HAF_Minimal-Risk-Maneuver.pdf

Lim, H. S. M., & Taeihagh, A. (2018). Autonomous vehicles for smart and sustainable cities: An in-depth exploration of privacy and cybersecurity implications. *Energies*, *11*(5), 1062. doi:10.3390/en11051062

Lin, P. (2016). Why ethics matters for autonomous cars. In M. Maurer, J. C. Gerdes, B. Lenz, & H. Winner (Eds.), *Autonomous driving* (pp. 69–85). Springer.

Lütge, C. (2017). The German ethics code for automated and connected driving. *Philosophy & Technology*, *30*(4), 547–558. doi:10.1007/s13347-017-0284-0

Manser, M. P., Noble, A. M., Machiani, S. G., Shortz, A., Klauer, S. G., Higgins, L., & Ahmadi, A. (2019). *Driver training research and guidelines for automated vehicle technology*. Retrieved from https://vtechworks.lib.vt.edu/bitstream/handle/10919/95178/01-004_Final%20Research%20Report_Final.pdf?sequence=1

Merat, N., Jamson, A. H., Lai, F. C., Daly, M., & Carsten, O. M. (2014). Transition to manual: Driver behaviour when resuming control from a highly automated vehicle. *Transportation research part F: traffic psychology and behaviour*, 27(Part B), 274-282.

Ministry of Transport and Communications. (1997). *En route to a society with safe road traffic*. Retrieved from https://trid.trb.org/View/512093

National Transportation Safety Board. (2019). *Vehicle automation report, Tempe, AZ, HWY18MH010*. Retrieved from https://www.documentcloud.org/documents/6540547-629713.html

National Transportation Safety Board. (2020). *Collision between a sport utility vehicle operating with partial driving automation and a crash attenuator*. Retrieved from https://www.ntsb.gov/investigations/accidentreports/pages/har2001.aspx

NHTSA. (2017). *Automated driving systems 2.0*. Retrieved from https://www.nhtsa.gov/sites/nhtsa.dot.gov/files/documents/13069a-ads2.0_090617_v9a_tag.pdf

Pagallo, U., Aurucci, P., Casanovas, P., Chatila, R., Chazerand, P., Dignum, V., Lütge, C., Madelin, R., Schafer, B., & Valcke, P. (2019). *AI4People on good AI governance: 14 priority actions, a SMART model of governance, and a regulatory toolbox*. Retrieved from: https://ssrn.com/abstract=3486508

Research and Markets. (2019). *Analysis of driver monitoring systems, 2020 Edition*. Retrieved from https://www.researchandmarkets.com/reports/4893877/analysis-of-driver-monitoring-systems-2020?utm_source=dynamic&utm_medium=GNOM&utm_code=c968tk&utm_campaign=1338270+-+2020+Analysis+of+the+Global+Driver+Monitoring+Systems+Market&utm_exec=cari18gnomd

Rouchitsas, A., & Alm, H. (2019). External human–machine interfaces for autonomous vehicle-to-pedestrian communication: A review of empirical work. *Frontiers in Psychology*, *10*(2757), 1–12. doi:10.3389/fpsyg.2019.02757 PMID:31920810

SAE International – Vehicle Cybersecurity Systems Engineering Committee. (2016). *SAE J3061: Cybersecurity guidebook for cyber-physical vehicle systems*. Retrieved from https://www.sae.org/standards/content/j3061_201601/

Secretary of State Washington. (2020). *Certification of enrollment – engrossed substitute senate bill 6280*. Retrieved from http://lawfilesext.leg.wa.gov/biennium/2019-20/Pdf/Bills/Senate%20Passed%20Legislature/6280-S.PL.pdf?q=20200409103455

Techcrunch. (2006). *AOL proudly releases massive amounts of private data*. Retrieved from https://techcrunch.com/2006/08/06/aol-proudly-releases-massive-amounts-of-user-search-data/

Tesla. (2020). *Using autopilot and full self-driving capability*. Retrieved from https://www.tesla.com/support/autopilot

United Nations. (1948). *Universal Declaration of Human Rights*. Retrieved from https://www.ohchr.org/EN/UDHR/Documents/UDHR_Translations/eng.pdf

United Nations. (2015). *Transforming our world: The 2030 agenda for sustainable development*. Retrieved from https://www.un.org/ga/search/view_doc.asp?symbol=A/RES/70/1&Lang=E

United Nations Economic Commission for Europe (UNECE). (2019). *Revised framework document on automated/autonomous vehicles*. Retrieved from https://www.unece.org/fileadmin/DAM/trans/doc/2020/wp29/ECE-TRANS-WP29-2019-34-Rev2e.pdf

United Nations Economic Commission for Europe (UNECE). (2020a). *Draft new UN regulation on uniform provisions concerning the approval of vehicles with regard to cyber security and of their cybersecurity management systems*. Retrieved from https://www.unece.org/fileadmin/DAM/trans/doc/2020/wp29grva/GRVA-06-19r1e.pdf

United Nations Economic Commission for Europe (UNECE). (2020b). *UN regulation on automated lane keeping systems is milestone for safe introduction of automated vehicles in traffic*. Retrieved from https://www.unece.org/info/media/presscurrent-press-h/transport/2020/un-regulation-on-automated-lane-keeping-systems-is-milestone-for-safe-introduction-of-automated-vehicles-in-traffic/doc.html

Vox. (2019). *A new study finds a potential risk with self-driving cars: Failure to detect dark-skinned pedestrians*. Retrieved from https://www.vox.com/future-perfect/2019/3/5/18251924/self-driving-car-racial-bias-study-autonomous-vehicle-dark-skin

Wilson, B., Hoffman, J., & Morgenstern, J. (2019). *Predictive inequity in object detection*. arXiv preprint arXiv:1902.11097

Wishart, J., Como, S., Elli, M., Russo, B., Weast, J., Altekar, N., James, E., & Chen, Y. (2020). Driving safety performance assessment metrics for ADS-equipped vehicles. *SAE International*. Retrieved from https://www.researchgate.net/profile/Jeffrey_Wishart/publication/340652968_Driving_Safety_Performance_Assessment_Metrics_for_ADS-Equipped_Vehicles/links/5eb0a39e92851cb2677403ba/Driving-Safety-Performance-Assessment-Metrics-for-ADS-Equipped-Vehicles.pdf

World Economic Forum (WEF). (2018). *Reshaping urban mobility with autonomous vehicles – Lessons from the City of Boston*. Retrieved from http://www3.weforum.org/docs/WEF_Reshaping_Urban_Mobility_with_Autonomous_Vehicles_2018.pdf

ENDNOTES

[1]   All co-authors of this paper constitute the AI4People-Automotive Committee.

[2]   The levels refer to the taxonomy developed by the SAE International (2018) for six levels of driving automation, ranging from no driving automation (level 0) to full driving automation (level 5).

[3]   The override function does not need to be similar to the way we are driving today such as taking over using a steering wheel or a paddle. On the contrary, the control can be a function provided through some interfaces that do not take away the original advantages of AVs such as inclusive accessibility.

Christoph Lütge is the Director of the IEAI and has held the Chair of Business Ethics at TUM since 2010. He has a background in business informatics and philosophy, having received his doctorate in philosophy from TU Braunschweig in 1999 and his habilitation from Ludwig Maximilians University Munich in 2005. He has held visiting positions at Harvard, Taipei, Kyoto and Venice, and was awarded a Heisenberg Fellowship from the German Research Foundation. In 2020, he was appointed Distinguished Visiting Professor of the University of Tokyo. Professor Lütge has been a member of the Ethics Commission on Automated and Connected Driving of the German Federal Ministry of Transport and Digital Infrastructure, as well as of the European AI Ethics initiative, AI4People.

Franziska Poszler is a research associate and doctoral student at the Peter Löscher Chair of Business Ethics at the Technical University of Munich (TUM), Germany. Before she worked as a research associate at the Chair for Strategy and Organization at TUM. She holds a Master's degree in Human Resources and Organisations from the London School of Economics (LSE) and a Bachelor's degree in Business Administration from the Ludwig Maximilian University of Munich (LMU).

Aida Joaquin Acosta researches the ethics and governance of artificial intelligence (AI) and the impact of emerging technologies, such autonomous vehicles (AVs) and the Internet of Things (IoT), in law, public policy, and society. She has developed a series of policy papers for regulators and policymakers to help them govern AVs and IoT. Aida is Head of the International Relations Department of the Ministry of Transport, Mobility, and the Urban Agenda in Spain; and an Affiliate of the Berkman Klein Center (BKC) at Harvard University, where she participated in the Ethics and Governance of AI Initiative of the BKC and the MIT Media Lab. Her educational background includes a Fulbright Scholarship at the University of Washington, where she obtained an LL.M. focusing on robotics law and received an Academic Excellence Award; a Master's Degree in Computer Engineering focused in AI and robotics; and a Master's Degree in Law.

David Danks is L.L. Thurstone Professor of Philosophy & Psychology, and Head of the Department of Philosophy, at Carnegie Mellon University. He is also the Chief Ethicist of CMU's Block Center for Technology & Society; and co-director of CMU's Center for Informed Democracy and Social Cybersecurity (IDeaS). His research interests are at the intersection of philosophy, cognitive science, and machine learning, using ideas, methods, and frameworks from each to advance our understanding of complex, interdisciplinary problems. Danks has examined the ethical, psychological, and policy issues around AI and robotics in transportation, healthcare, privacy, and security. He has also done significant research in computational cognitive science, culminating in his Unifying the Mind: Cognitive Representations as Graphical Models (2014, The MIT Press). Danks is the recipient of a James S. McDonnell Foundation Scholar Award, as well as an Andrew Carnegie Fellowship.

Gail Gottehrer is the Founder of the Law Office of Gail Gottehrer LLC. Her practice focuses on emerging technologies, including autonomous vehicles, artificial intelligence, and the Internet of Things, and the privacy laws, cybersecurity requirements and ethical issues associated with the data these technologies collect and use. Gail is one of the few defense lawyers to have been involved in the trial of a class action to verdict before a jury. She is the Co-Chair of the New York State Bar Association's Technology and the Legal Profession Committee, the Chair of the American Bar Association's TIPS Automobile Litigation Committee, and the Co-Chair of the National Association of Women Lawyers' IP & Technology Affinity Group. She is a Member of the State of Connecticut's Task Force to Study Fully Autonomous Vehicles and Co-Chaired the Regulatory, Safety, Law and Policy Subcommittee of the New York State Bar Association's Task Force on Autonomous Vehicles and the Law. Gail's recent publications include Mobilizing "Digital First Responders": Level 4 Autonomous Vehicles, Can States Steer Clear of Liability for Accidents Involving Autonomous Vehicle Technology?, and The Intersection of the Fourth Amendment and Level 5 Vehicle Autonomy.

Lucian Mihet (M'12, SM'15) was born in 1969. He received his Habilitation (2015) and Ph.D. degree (2002) in Electrical Engineering, a Master's degree (2000) in Electric Drives and Power Electronics and a Bachelor's degree (1999) in Electrical Engineering, from the Politehnica University of Timisoara-Romania. Since 2016, he is working as Full Professor in Energy Technology at Oestfold University College in Norway. From 1999 to 2016, Professor Mihet has been with the Politehnica University of Timisoara, Romania. He has also worked as a Research Scientist with the Danish Technical University (2011–2014) and Aalborg University (2000–2002) in Denmark and as a Post Doc with Siegen University in Germany in 2004. Dr. Lucian Mihet has published more than 140 papers in national and international journals and conference proceedings, and 10 books. Since 2017, he has been a Guest Editor for five Special Issues for the MDPI Energies and Applied Sciences Journals, for Maljesi and for the Advances in Meteorology Journal. He has served as Scientific and Technical Programme Committee Member for many IEEE Conferences. Professor Mihet-Popa has participated in more than 15 international grants/projects, such as FP7, EEA and Horizon 2020, and has been awarded more than 10 national research grants. He is also the Head of the Research Lab "Intelligent Control of Energy Conversion and Storage Systems" and one of the Coordinators of the Master Program in "Green Energy Technology" at the Faculty of Engineering. His research interest include modelling, simulation, control and testing of Energy Conversion Systems, Distributed Energy Resources (DER) components and systems, including battery storage systems-BSS (for electric vehicles and hybrid cars and vanadium redox batteries (VRB)), energy efficiency in smart buildings and smart grids. Professor Mihet was invited to join the Energy and Automotive committees as an Expert by the President and the Honorary President of Atomium European Institute, working in close cooperation with—and under the umbrella—of the EC and EU Parliament, and was Chairing Energy Section of AI4People 2020.

Aisha Naseer is Research Manager at Fujitsu Laboratories of Europe Limited (FLE) UK, where she is leading the AI Ethics research activity through representing Fujitsu's involvement at the AI4People forum, the EU AI Alliance, and the IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems. She is a GDPR-certified professional and participates in debates around AI Ethics, privacy, and digital trust including bias mitigation and AI fairness. In previous research projects, she has been providing strategic direction with leadership and expert advice in various industry sectors such as Healthcare, Finance, IoT & Cyber Security. She has extensive experience in successfully leading projects internationally, negotiated complex and strategic contracts. Dr Naseer is an innovative scientist with several patents, and has been an author of numerous papers in journals, international conferences, and book chapters. She has also chaired various ACM and IEEE conference sessions, and pioneered multiple special tracks and workshops. Dr Naseer is member of the founding Editorial Board of the Springer Journal of AI and Ethics.