Received 22 June 2022, accepted 3 August 2022, date of publication 5 August 2022, date of current version 11 August 2022.
Digital Object Identifier 10.1109/ACCESS.2022.3196941

PERSPECTIVE

TrustFSDV: Framework for Building and Maintaining Trust in Self-Driving Vehicles

KAYA KURU
School of Engineering, University of Central Lancashire, Preston PR1 2HE, U.K.
e-mail: kkuru@uclan.ac.uk

ABSTRACT The deployment of fully self-sufficient, self-driving vehicles (SDVs) to city roads in a diverse range of benefit-driven concepts is on the horizon where the higher levels of automation with automated driving systems (ADSs) are being installed on vehicles. The acceptance of SDVs as safety-critical systems to successfully substitute for human drivers requires a high level of trustworthiness in numerous socio-technical aspects. By the time SDVs are available to the general public with challenges and opportunities, the trustworthiness of SDVs in rapidly changing, partially observable, multiagent, stochastic, sequential, dynamic, continuous and unknown environments shall be ensured from the perspective of the stakeholders. This paper analyses the dynamics of trust in SDVs from the perspectives of all stakeholders to result in how the "trustworthiness" of SDVs can be ensured with mitigated risks and dangers (i.e., non-maleficence) and maximised benefits (i.e., beneficence) as a holistic view of all the dimensions and dynamics of SDVs. Besides, a framework, namely, TrustFSDV, that designates a transitional roadmap with delegated responsibilities and liabilities along with the desired performance indicators, is proposed for the stakeholders to pursue in ensuring a sufficient level of trust within the concepts of human-vehicle integration and society-vehicle harmonisation — inducing the advancement of SDVs and acceptance of them by the society — while the future is being revolutionised by SDVs. To conclude the key findings in this research, i) although instant trust can be gained through hard work, it can be lost readily without forward-thinking during the life-cycle of SDVs, ii) a high level of trustworthiness cannot only be achieved by manufacturers, it requires the strict collaboration of all the stakeholders with distributed responsibilities and liabilities while moulding the requirements of the stakeholders in a system under the supervision of related disciplines, iii) the manufacturers along with the press shall not ignore the real technical limitations of the SDV technology when communicating about benefits to the potential customers, and iv) drastic steps are needed to be taken concerning ethical and legal perspectives, which is perceived as the primary concern beyond technology to build and maintain a high level of trustworthiness.

INDEX TERMS Autonomous vehicles, self-driving vehicles, autonomy in vehicles, trust in self-driving, vehicle teleportation.

I. INTRODUCTION

The automation levels in the automobile industry have been analysed in [1] with a categorical step-wise progression through the hierarchical levels. The higher the level, the more autonomy with decreasing human supervisory for extended periods of time. Level-5 fully autonomous ground vehicles (FAGVs) — i.e., self-sufficient self-driving vehicles (SDVs) — requiring a high level of trust with reliable and safe manoeuvres/operations by removing humans out of the loop with no steering wheel, no pedals, even no windshield, elevate manufacturers in a highly confident position compared to the preceding two-hybrid automated levels, namely level-3 and level-4 that represent a joint human-machine system with mixed-mode operations and shared responsibilities (i.e., human-in-the-loop (HITL)). The Level-5 autonomy as

The associate editor coordinating the review of this manuscript and approving it for publication was Hassan Omar.

representing the next-generation autonomous vehicles (AVs) is defined by the SAE J3016 standard as a system that can perform "under all driver-manageable on-road conditions" meaning that it is as perfect as a human driver. The deployment of commercial SDVs to city roads in a diverse range of benefit-driven concepts is on the horizon due to the accelerated advances in integrated electronics with high computation powers, sensors, actuators and intelligent reasoning equipped with cognitive computing. Cognitive computing, with a high level of reasoning and swarm-based solutions, in SDVs mimics human cognition to remove the human in the loop by creating a highly trustworthy ecosystem for all stakeholders while evolutionary approaches still have very limited abilities in cognitive learning when compared to a human. It aims to mould various data sources involving onboard sensor data, geo-distributed insights [2], ethics with advanced real-time analytics and actuation mechanisms within rapidly changing, partially observable, multiagent, stochastic, sequential, dynamic, continuous and unknown environments.[1]

SDVs with fully self-driving capabilities, lacking pedals and steering wheels, are aimed to be built for riding rather than driving. They may still have folded steering wheels and pedals for use by drivers under the conditions in which the automation is not capable of self-operating where they can be designed to be able to turn into level-4 autonomy by unfolding these components with strictly limited authorised keys. They, with the concept of connected and AVs (CAV), can communicate information to one another as they pave the way for swarm-based solutions through swarm engineering within the concepts of Automation of Everything (AoE) [4], Internet of Everything (IoE) and vehicle-to-everything (V2E) [1]. Building trust between the public and SDVs is crucial to their widespread acceptance and deployment. The intended benefits of SDVs can not be materialised unless a high level of trust in SDVs is achieved in the human mind through measurable, quantifiable and biological signals [5]. Despite the diversity in defining trust in the literature [6], trust in SDVs can be defined in a broader perspective as "confidence and reliance on the integrity, reliability, safety and ability of the fine-granular functionalities of SDVs in satisfying/achieving desired intended goals while performing the agreed-upon/expected tasks under particular kinds of regulated circumstances" where "confidence" is about feeling sure of the abilities of SDVs and "reliability" indicates the same expected results yielded through repeated experiments and trials, ensuring that the granular functionalities will consistently operate properly. It is noteworthy to point out that 'trust' — the perception of their capability — can be established and increased as each desired, beneficial and individual reliable functionality in SDVs is achieved in each stage of its lifecycle whereas 'trustworthiness' — the perception about how well the automation operates — corresponds to the excellent orchestration and harmonisation of all these highly related functionalities. All the requirements of the stakeholders need to be moulded in a system leading to achieving the desired goals through proper and appropriate actions and then leading to 'trustworthiness' that would realise the full potential of SDVs. Methodologies that determine the trust of SDVs are urgently needed to enable the certification of such systems. Within this context, this paper aims to close the literature gap about building and maintaining trust in SDVs by analysing trust from the perspectives of all stakeholders to result in how the "trustworthiness" in SDVs can be ensured as a holistic view of all the dimensions (Fig. 1) and dynamics (Fig. 2) of SDVs. The main goal of this analysis is to demonstrate a direction for alleviating the concerns during the penetration of this emerging technology into the mixed urban traffic environment with a well-deserved high-level trust. The contributions of this paper are outlined as follows.

1) The crucial dynamics of trust in SDVs are examined from the perspectives of all stakeholders within a high degree of autonomy.
2) How trust is gained and maintained during this multi-dimensional transformation — from vehicle ownership to vehicle usership — is delineated from ethical and socio-technical points of view.
3) A human-centric holistic trust model for SDVs — TrustFSDV — is conceptualised from human-oriented to vehicle-oriented concept through robust human-vehicle integration and society-vehicle harmonisation aimed at building trustworthy SDVs.

The rest of the paper is structured as follows. The related works are investigated in Section II. Trust from the perspective of the stakeholders is explored in Section III. Verification of trust with a proposed iterative approach is introduced in Section IV. Ethical aspects of trust are scrutinised in Section V. Discussion is provided in Section VI followed by the key findings and future directions in Section VII.

II. RELATED WORKS

Trust is the major factor that affects the human reaction to autonomous systems in uncertain situations [7] while actions are selected by those systems with little to no control from their users. Trust in the literature is analysed specific to technology in general [6], specific to intelligent agent-based systems [8], specific to AI [9], specific to automation in technology [10], specific to automation in vehicles [11], and specific to AVs [12]. Regarding the measurement of trust, humans tend to become more relaxed as they acquire trust under uncertainty. Several studies (e.g., [13]) on psychology and physiology suggest that body kinematics change with respect to the change in mental state and cognitive overload. From this point of view, Nahavandi [5] analyses trust for the autonomous systems by measuring changes in human biomarkers based on physiological signals (i.e., biomarker signals) such as heart rate variability (echocardiogram), brain activity (electroencephalogram), eye movement, skin temperature, muscle activation and patterns of body

[1]The author refers readers to [3] for the definitions of these terms.

movement while interconnecting with particular autonomous tasks. Again, Nahavandi [14] explores trust in autonomy between humans and robots. The human factors that induce trust in automation are explored by Hoff and Bashir [7] with a three-layered trust model — learned trust, situational trust and dispositional trust. The human-machine interface (HMI) design principles, that empower user trust in AV systems, have been investigated in [15]. Further studies that cover trust in SDVs are mentioned throughout this paper specific to the particular themes. Despite numerous comprehensive studies on trust between humans and robots, humans and machines or humans and automation, in this treatise, no comprehensive particular study that analyses trust between non-autonomous/autonomous entities and SDVs in a broader picture has been noted. To the best of my knowledge, this is the first comprehensive study that highlights a research gap in this particular subject which is analysed from a socio-technical point of view considering the perspectives of all stakeholders. With this analysis, a framework, namely, TrustFSDV is developed to conceptualise all the intrinsic multidimensional dynamics of trust in SDVs.

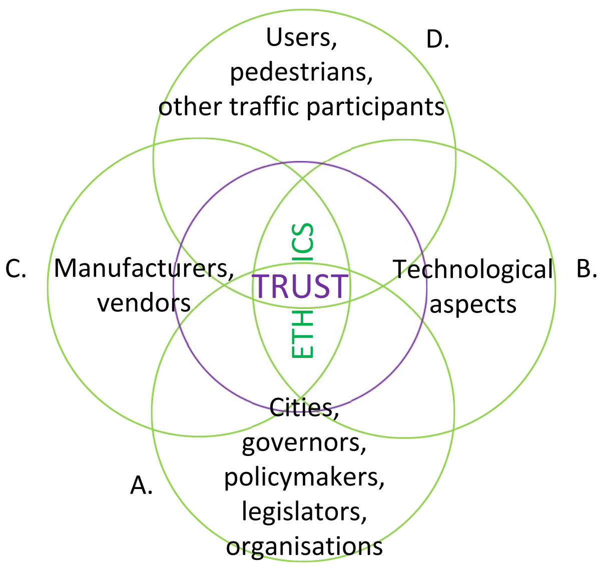

FIGURE 1. Primary dimensions of trustworthiness in SDVs.

III. TRUST IN SDVs FROM THE PERSPECTIVES OF STAKEHOLDERS

Trust, in a broader perspective, can be defined as the willingness of a party to be vulnerable to the actions of another party based on the expectation that the other will perform a particular action important to the trustor who doesn't have full control over the actions of the other party, [16]. It is widely accepted that the benefits of autonomous robotics outweigh the problems within the industrial revolution [5]. The trend of delegating our daily routines to intelligent machines is expected to accelerate in the future as trust in intelligent machines is gained and maintained [17]. In the case of automation, trust can be measured by considering the probability of reliable behaviours in the presence of externally induced uncertainty [17]. For SDVs to succeed, humans have to start trusting their behaviours in the presence of uncertainty with the reliable course of repeatable actions. One of the key factors influencing public acceptance of automated vehicle technologies is their level of trust [11]. Comprehensive efforts are required to achieve the desired level of trust in SDVs, in particular, with the ways of creating and maintaining confidence within a formation of positive thinking, feeling, beliefs, and attitudes in various key aspects.

Providing a sufficient amount of trust in SDVs requires a strict collaboration of stakeholders within a broader perception containing a multitude of dimensions as illustrated in Fig. 1 and as formalised in Fig. 2 with the key components and dynamics. Readers who would like to see these key trust components and dynamics (Fig. 2) within a table are referred to Table 1. It is worth mentioning that the components of trust presented in the figures cannot be distinctly separated from each other, rather, they are intertwined and complementary in the phases of building and maintaining a collective trustworthiness in which responsibilities and liabilities are distributed. In other words, the trustworthiness in SDVs can not be gained or maintained at the strategic level to satisfy all the stakeholders if all the proposed dynamics (Fig. 2) within specified four dimensions (Fig. 1) are not properly implemented in parallel with one another. Manufactures are those that design SDVs to meet the requirements of the stakeholders, but it is legislators, policy-makers and regulators that can set the rules and principles — satisfying all parties — to guide manufactures on what these general requirements are and what the measurement criteria and metrics are for realising them and finally enforce them. For instance, the trustworthiness can not be retained even if the manufacturer builds a safe SDV with many beneficial functionalities where the legal and regulatory framework is not formed by the legislative body and the government to specify the responsibilities, liabilities and required regulations which, in turn, would require the modifications of these functionalities in line with the legal and regulatory framework. Otherwise, a chaotic environment is created despite the tactical achievements, which leads to untrustworthiness in SDVs within unharmonised engagement situations. In this regard, the question "how can all the stakeholders (Fig. 1) be brought together and cooperate more efficiently at the strategic level to promote the development of self-driving?" should be addressed well to establish a sufficient level of trustworthiness. There is no doubt that all parties would benefit from effective cooperation and they should show a readiness to be involved in fruitful collaborative efforts in a way to move forward in determining ethical procedures and guidelines and in delivering trustworthy SDVs. The author strongly believes that the leading stakeholders, shown in Fig. 1A, have to be a driving force as an umbrella to help establish this synergistic collaboration among all stakeholders leading to the collective trustworthiness of SDVs. Furthermore, trust can be maintained through well-designed

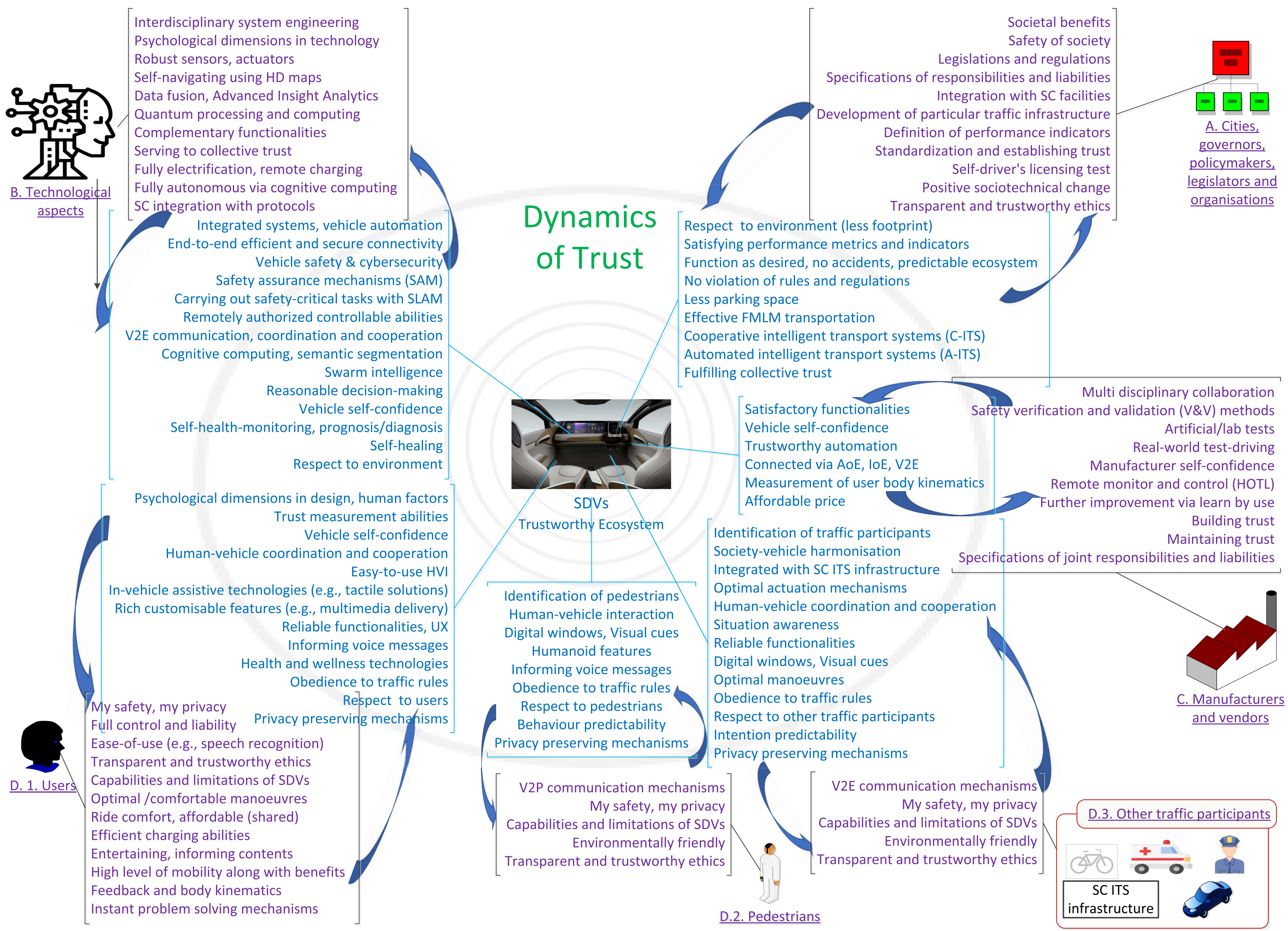

FIGURE 2. **TrustFSDV framework: i) Environmental trust dimensions: Distribution of responsibilities and liabilities and components of trust from the perspectives of all stakeholders within a synergistic collaboration (blue text) and ii) Intrinsic SDV trust dimensions shaped by the environmental trust dimensions: Interrelated components from the aspects of SDVs (brown text) leading to the collective trustworthiness of SDVs.**

TABLE 1. TrustFSDV framework: i) Environmental trust dimensions: Distribution of responsibilities and liabilities and components of trust from the perspectives of all stakeholders within a synergistic collaboration (blue text) and ii) Intrinsic SDV trust dimensions shaped by the environmental trust dimensions: Interrelated components from the aspects of SDVs (brown text) leading to the collective trustworthiness of SDVs.

A. Cities, governors, policymakers, legislators and organisations	B. Technological aspects	C. Manufacturers and vendors	D. 1. Users	D.2. Pedestrians	D.3. Other traffic participants
-Societal benefits -Safety of society -Legislations and regulations -Specifications of responsibilities and liabilities -Integration with SC facilities -Development of particular traffic infrastructure -Definition of performance indicators -Standardization and establishing trust -Self-driver's licensing test -Positive sociotechnical change -Transparent and trustworthy ethics	-Interdisciplinary system engineering -Psychological dimensions in technology -Robust sensors, actuators -Self-navigating using HD maps -Data fusion, Advanced Insight Analytics -Quantum processing and computing -Complementary functionalities -Serving to collective trust -Fully electrification, remote charging -Fully autonomous via cognitive computing -SC integration with protocols	-Multi disciplinary collaboration -Safety verification and validation (V&V) methods -Artificial/lab tests -Real-world test-driving -Manufacturer self-confidence -Remote monitor and control (HOTL) -Further improvement via learn by use -Building trust -Maintaining trust -Specifications of joint responsibilities and liabilities	-My safety, my privacy -Full control and liability -Ease-of-use (e.g., speech recognition) -Transparent and trustworthy ethics -Capabilities and limitations of SDVs -Optimal/comfortable manoeuvres -Ride comfort, affordable (shared) -Efficient charging abilities -Entertaining, informing contents -High level of mobility along with benefits -Feedback and body kinematics -Instant problem solving mechanisms	-V2P communication mechanisms -My safety, my privacy -Capabilities and limitations of SDVs -Environmentally friendly -Transparent and trustworthy ethics	-V2E communication mechanisms -My safety, my privacy -Capabilities and limitations of SDVs -Environmentally friendly -Transparent and trustworthy ethics
-Respect to environment (less footprint) -Satisfying performance metrics and indicators -Function as desired, no accidents, predictable ecosystem -No violation of laws, rules and regulations -Less parking space -Effective FMLM transportation -Cooperative intelligent transport systems (C-ITS) -Automated intelligent transport systems (A-ITS) -Fulfilling collective trust	-Integrated systems, vehicle automation -End-to-end efficient and secure connectivity -Vehicle safety & cybersecurity -Safety assurance mechanisms (SAM) -Carrying out safety-critical tasks with SLAM -Remotely authorized controllable abilities -V2E communication, coordination and cooperation -Cognitive computing, semantic segmentation -Swarm intelligence -Reasonable decision-making -Vehicle self-confidence -Self-health-monitoring, prognosis/diagnosis -Self-healing -Respect to environment	-Satisfactory functionalities -Vehicle self-confidence -Trustworthy vehicle automation -Connected via AoE, IoE, V2E -Measurement of user body kinematics -Affordable price	-Psychological dimensions in design, human factors -Trust measurement abilities -Vehicle self-confidence -Human-vehicle coordination and cooperation -Easy-to-use HVI -In-vehicle assistive technologies (e.g., tactile solutions) -Rich customisable features (e.g., multimedia delivery) -Reliable functionalities, UX -Informing voice messages -Health and wellness technologies -Obedience to traffic rules -Respect to users -Privacy preserving mechanisms	-Identification of pedestrians -Human-vehicle interaction -Digital windows, Visual cues -Humanoid features -Informing voice messages -Obedience to traffic rules -Respect to pedestrians -Behaviour predictability// - Privacy preserving mechanisms	-Identification of traffic participants -Society-vehicle harmonisation -Integrated with SC ITS infrastructure -Optimal actuation mechanisms -Human-vehicle coordination and cooperation -Situation awareness -Reliable functionalities -Digital windows, Visual cues -Optimal manoeuvres -Obedience to traffic rules -Respect to other traffic participants -Intention predictability -Privacy preserving mechanisms

iterative processes as illustrated in Figs. 3 and 4 by which the capabilities are immensely tested and any problematic space is urgently covered both during production and use. The main trust dimensions with their key components and dynamics (Figs. 1 and 2) are elaborated with those figures in the following subsections.

A. TRUST FROM THE PERSPECTIVE OF CITIES, GOVERNORS, POLICYMAKERS, LEGISLATORS AND ORGANISATIONS

(Figs. 1 A, 2 A). The results of a survey conducted by the World Economic Forum with 20 policymakers and transport authorities from several cities such as Helsinki, Dubai, New York, Singapore, Amsterdam and Toronto about what they expect from SDVs suggest that AV applications have the potential to provide useful solutions to public services [18]. A high level of trust, pertaining to SDVs from city policymakers and transport authorities, is evident based on their beliefs in SDVs. What might be the essential reasons behind this trust and how can it be quantified with tangible measurable pieces of evidence and reinforced with tangible assets using adequate legislation, rules, regulations and standards introduced under the supervision of related disciplines (e.g., philosophy, psychology, physiology, sociology, ergonomics, engineering)?

Building trust between the policymakers and the industry using effectively established communication platforms is critical during the development phases and, in particular, commercial use of SDVs to make this technology thrive and meet the expectations of all stakeholders with the agreed-upon legal and regulatory framework. From the point of the regulatory framework, i) the traffic infrastructure needs to be improved (e.g., digital communication of SDVs with traffic signs and signals, roadside units (RSUs) to connect SDVs to each other and smart city (SC) facilities for swarm intelligence and smart mobility with the concepts of the cooperative intelligent transportation system (C-ITS) and automated ITS (A-ITS)) and ii) traffic rules, regulations and legislation need to be amended by considering the specific features of SDVs and their interrelation with other traffic participants. It is worth noting that the construction of required rules, regulations and laws is well behind to keep up with the technology, in particular, in accommodating the new driver — AI, which poses a challenge in building the necessary required public trust in SDVs and in realising the objectives of this type of autonomy. Governors and legislators are responsible for establishing the proof of principles of all aspects of SDVs in coordination with all the other stakeholders involving the industry and organisations that represent a diverse range of citizens with different requirements and expectations. However, central governments and city governors are still hesitant regarding the deployment of SDVs. More specifically, our cities, city planners, policymakers, governors and legislators don't seem to have a distinct plan for this impending technology within a forward-thinking policy.

The conceptualisation of the main trust components between SDVs and cities, governors, policymakers, legislators and organisations is presented in Fig. 2 A. Generally speaking, the law considers the AV system as a responsible vehicle operator when it is engaged, even if there are commuters inside the vehicle, meaning that the company in charge of the vehicle and the manufacturer are taken accountable in the event of an accident [19]. This approach encourages the manufacturers to ensure that the technology they develop is safe and reliable. On the other hand, autonomous driving systems are composed of different components and many subsystems produced by different companies, hence, how responsibilities and liabilities should be distributed? More specifically, it is of prime importance to ensure that all the technological components embedded in SDVs behave predictably, securely, and reliably to materialise the objectives in a trustworthy ecosystem regarding joint liability. The law enforcement framework should be shaped together with all the stakeholders such as the Department of Transportation (DoT), city governors, industry, lawmakers, regulators, and the public involving other required expertises (e.g., sociology, psychology). An agreed-upon regulatory framework can help manufacturers release reliable autonomous models to the public [19] and can help materialise the societal expectations from SDVs. It is worth mentioning that this regulatory framework shall specify the responsibilities and liabilities of all other stakeholders such as i) local governments in providing the required support environments (e.g., electronic traffic signs and signals and RSUs in line with the industry) for the SDVs, ii) conventional vehicles and other traffic participants (e.g., pedestrians, traffic police, ambulances) in engagement with SDVs.

On one hand, there are no approved safety standards and protocols for automated driving, and it will be a long time before they exist. On the other hand, extensive tests are required to validate the particular computable metrics and capabilities of SDVs based on standardised performance indicators. Performance indicators, that are meant to show how an SDV is expected to perform its tasks accurately and efficiently, should be specified from the perspective of all stakeholders using the approved safety standards and protocols. Based on these performance indicators, SDVs have to be certified with self-driver's licensing test (SDLT) under the supervision of governors to ensure that the required safety has been met and goals have been achieved with an appropriate course of action. Technically speaking, decision-making capabilities of agent-based SDVs involving human factor objectives can be formally verified and licensed under the supervision of governors using SDLT through a series of tasks leading to a satisfactory self-confidence score to provide evidence for the required verified safety assurances and certification. Specifications of responsibilities and liabilities, with required legislation and regulations, in particular, safety regulations assured by standards for reducing negative impacts and keeping the risks at an acceptable level, have to be clearly defined by incorporating all the inputs from the stakeholders. Within this context, a federal AV policy was published by the U.S. DoT. This policy is composed of four themes, namely, i) regulatory duties of the federal and state governments, ii) a guided regulatory framework for the safe detection and use of automated driving technology, iii) SDVs, and finally, iv) NHTSA (National Highway Traffic Safety Administration) regulatory measures and new supervisory agencies and actions [20]. SDVs are required to be tested by qualified authorities both in quality test environments and in the real-world urban ecosystem using well-established standards and protocols, by which their competence, safety and reliability have to be ensured before being deployed into imperfect city traffic. In this manner, governments along with city planners in collaboration with manufacturers should develop innovative test environments (that is elaborated in Section III-C). In this context, SDLT within both augmented-reality-based artificial/virtual and real-world environments involving realistic scenarios is an emerging topic. It is urgent to specify the standards in these tests by considering SDVs' abilities for assuring that vehicles passing these tests can run on city roads safely with desired performance criteria (i.e., expected outcome) by coping with the rules and regulations. Government agencies such as NHTSA as a branch of the US DoT suggests an entirely voluntary approach for the industry with regulations and necessary required features to be embedded in SDVs. During SDLT with unexpected scenarios, the abilities of SDVs can be examined if they can handle unexpected conditions efficiently. However, we shouldn't expect AVs to cope with a predicament that can not even be figured out by the very best human driver [21]. The correct, efficient trade-off between "safety" and "productivity" has been a recurrent pain point to the safety-critical systems and stakeholders. While determining safety rules, it has to be taken into account that mandating excessive safety regulations would paralyse the city traffic and seriously hinder the acceptance of AVs [22] (e.g., 20 km/h speed limit). In this vein, the U.S. DoT released a series of procedures for the safe integration of AVs in the transportation system for industry [23]. AVs and their operators shall be registered in California and each operator shall attend a safety driving course [24]. The legal framework for SDVs to operate in Florida without a human operator has been identified [19].

SDVs need to communicate with SC facilities, other vehicles and components of the street to increase their efficacy. SCs equipped with next-generation communication technologies and SC facilities will allow SDVs i) to be connected more than ever leading to swarm intelligence, and ii) to use real-time and near-real-time insights for excellent decision-making, e.g., selections of near-optimal paths with appropriate trajectories [1]. Furthermore, widespread use of SDVs will impact cities, urban planning, citizens and other traffic participants significantly, particularly, i) the sustainability of cities with less carbon footprint (60% reduction of CO_2 [25]), less congestion and more efficient fuel consumption, ii) road, parking, and pedestrian infrastructure, iii) travel behaviours, vehicle ownership and sharing, health and comfort of citizens with fewer worldwide traffic accidents, and finally, iv) interrelation with other AVs, conventional vehicles, police and pedestrians [1]. In this perspective, the accommodation of SDVs within mixed urban traffic should be facilitated with forward-thinking mechanisms not only to address the probable predicaments, but also to establish an integrated synergistic urban ecosystem that leads to a high level of collective trust in SDVs.

B. TRUST REGARDING TECHNOLOGICAL ASPECTS

(Figs. 1 B, 2 B). Reliable and trustworthy safety assurance mechanisms (SAM) are the key enabler of trust in SDVs. In-vehicle technology is designated for both improving the safety of SDVs and achieving the desired autonomous tasks that meet user expectations. The prominent players of the automotive industry are collaborating with leading technology companies (e.g., Intel, Samsung, Microsoft, Apple, Nvidia, Mobileye) i) to ensure a stringent level of trust from the perspective of technological aspects,[2] and ii) to expedite their objective of placing commercial driverless vehicles on city roads with large market penetration levels (e.g., 75% by 2050 [27]). The technology company, Perrone, can turn a conventional vehicle produced by any manufacturer into AV by embedding its technology such as sensors, AI and actuators [28]. The more penetration with AoE, the more benefits and efficacy leading to optimised mobility [1]. It is worth noting that the non-commercial AVs have already experimented millions of miles on city roads under the supervision of test drivers in the vehicles[3] whereas only a tiny fraction of driving has been performed without backup drivers in the vehicles with remote supervision and assistance.

Can SDVs be proven safe and reliable? SDVs are meant to eliminate the accidents caused by human drivers using the advanced abilities of various onboard sensors with data fusion approaches during every type of adverse weather condition. They are designed to comply with all the rules and regulations of the road. They, equipped with advanced sensor technologies, can observe their environment beyond human vision capabilities, especially, in low-light conditions. Moreover, SDVs can respond rapidly to avoid an imminent collision. Ensuring the security and safety of SDVs supported by advanced self-health-monitoring abilities for prognosis/diagnosis of failures and handling them with self-healing mechanisms become crucially important for making these systems trustworthy. SDVs make their decisions based on what is best in a specific condition where the scene changes constantly in a highly dynamic environment. SDVs can be monitored and controlled using cyber-physical devices — IoT, IIoT and AMSs — from a remote location in real-time in case of any emergency using high-security authentication protocols. The self-driving experience of AVs and security can be increased with instant or regular software updates

[2]Readers are referred to [26] for the particular SDV technologies.

[3]The author refers readers to the study [1] for examples of AVs.

using standardised protocols enabling a lifetime improvement within the aspects of AoE and IoE.

The conceptualisation of the main trust components between SDVs and technological aspects is presented in Fig. 2 B. It doesn‘t seem reasonable to place all the burden on the shoulders of SDVs in realising trust. The real benefits of SDVs can emerge when they are embedded in and able to work with whole systems, road users and physical and digital infrastructures [29]. SDVs may encounter GPS-denied manoeuvring within tunnels and with high rising buildings, where communication paths and data links can be lost [30]. Besides, the sensitivity of the in-vehicle sensors decreases due to extreme weather conditions (e.g., snow, fog, hail). Therefore, traffic infrastructure (e.g., RSU) has to be designed in such a way that SDVs can communicate with traffic signs, conventional vehicles, one another and other traffic participants for mitigating the aforementioned concerns. On one hand, advanced physical and digital road infrastructures can be highly beneficial in realising the objectives of autonomy (i.e., CAV, human-vehicle collaboration and interaction) by building desired communication links for swarm intelligence leading to fulfilling collective trust. On the other hand, the efficacy and reliability of a system where information is gathered and shared among autonomous entities raise concerns about data authenticity [31]. The more communication and interaction requirements with the environment, the more cybersecurity risks. A hacked SDV can cause disastrous consequences such as the shutdown of roads, damage to property and even loss of human lives. Potential daunting cyberattacks (e.g., spoofing attacks) on AVs and potential strategies for mitigating or overcoming these threats are analysed in numerous studies (e.g., [32], [33]). Effective cybersecurity techniques involving intelligent control of SDVs under unexpected manoeuvrers are required to avoid the malicious use of hijacked vehicles. Data sharing ends should be registered as trusted sources using effective authentication techniques and the contents of the information shared should be protected with high-security encryption techniques with high-performance abilities to concretise trust in SDVs.

To summarise, the optimisation of the cooperative actions of swarms of SDV agents within the concepts of CAV, IoE and AoE should be performed well to make SDVs do their jobs better, which in turn with more trusted functionalities, reinforces trustworthiness among stakeholders substantially. Fully electrification of SDVs using batteries with exceedingly long lifetimes, and near-zero emission (i.e., CO_2) increases the confidence of all stakeholders in this technology. Additionally, technologies for wireless power transfer such as inductive and magnetic resonant coupling offers an easily automated energy supply that can enable the fully autonomous operation of electric vehicles [34] and wireless power transfer stations distributed around the city will increase the autonomous abilities of SDVs further. The more the benefits with easier use of this autonomy, the more confidence it generates. According to a simulation study [1], the traffic flow can benefit significantly from SDVs without significant infrastructural investment where the penetration level of SDVs increases in the mixed traffic within C-ITS and A-ITS in which SDVs are integrated with SC facilities. It is worth emphasising that the perfection of SDV technology is only halfway or one dimension to building a high level of trust in SDVs. One of the other dimensions is already mentioned in Section III-A. Further dimensions of building trust are explored in Sections III-C and III-D.

C. TRUST FROM MANUFACTURES' PERSPECTIVE

(Figs. 1 C, 2 C). Google, Ford, Lyft, Uber, Volvo, Honda, Nissan, Toyota and Waymo are teaming up in the AV Coalition Group to work with legislators, regulators, and the public to realise the safety and societal benefits of SDVs.[4] The first version of safety verification and validation (V&V) methods of SAE level-3 and level-4 automated driving was published by a collaboration of main manufacturers in 2019 in a white paper titled ‘‘safety first for automated driving’’ [35] to establish a basis towards the industry-wide standardisation and legal framework. Algorithmic technologies require consultation with diverse disciplines. Automakers are immensely collaborating with multi-disciplines such as electronics and software companies (Section III-B), cognitive scientists, psychologists to ensure trust properly from the perspectives of all stakeholders. The conceptualisation of the main trust components between SDVs and manufacturers is presented in Fig. 2 C.

Who is to blame for an accident: manufacturer, SDV, subcontractor, policymakers, governors or laws? Volvo takes full responsibility for their cars in the autonomous mode, strictly speaking, they accept the liability of the accident for their SDVs [18], [36] which shows the self-confidence of Volvo. Manufacturers will likely be held liable for vehicles that react incorrectly, and so they would like to know how their vehicles respond in very complex and unpredictable environments in order not to face any dire consequences. For users, the most challenging question that needs to be answered is ‘‘how can they be assured with the safety of SDVs, especially, in uncertain conditions?’’. To be able to answer this question properly, first, the question, ‘‘how can the self-confidence of manufacturers be insured’’ needs to be revealed. The automobile industry must ensure that the SDV technologies are built, tested, and validated to the highest safety standards, first to boast vehicles' and manufacturers' self-confidence and second to help consumers start trusting them. Manufacturer self-confidence in SDVs — self-awareness in the competency of establishing robust SDVs is highly important to deploy the technology safely in real-world environments by ensuring compliance with laws, rules and regulations. To be able to achieve this, it is crucial to develop new effective approaches to test SDVs in complex scenarios before commercial use. In this direction, Ma *et al.* [20] propose a framework for designing, testing and evaluating SDVs in both experimental research and practical applications by considering different

[4]https://www.selfdrivingcoalition.org/

dimensions of intelligence grading. The manufacturers are investing heavily in test infrastructures to both increase the capabilities of their vehicles through mass testing and to gain self-confidence. For instance, Waymo invested $1.1 billion to develop their own simulation test environments and high definition (HD) mapping tools for gaining a high level of self-confidence in their SDVs [26]. SDVs, with their capabilities and limitations, are practising over and over again in very complex artificial and real-world environments, most of the time with safety drivers, to have the most skilled AI driver on urban roads. Manufacturer self-confidence can be concretised through iterative and extensive tests based on agreed-upon standards, policies and protocols: artificial and real-world tests.

1) ARTIFICIAL/LAB TESTS

Real-world SDV experimentation can be challenging due to cost and safety issues, constantly changing testing regulations and the difficulties in generating appropriate scenarios. Embedded AI packages (e.g., sensor data fusion) with Software-In-the-Loop (SIL) and in-vehicle hardware (e.g., sensors, actuators, electronics) with Hardware-In-the-Loop (HIL) can be tested rapidly in the simulation loop using near-to-real-world annotations by achieving high fidelity autonomous driving. Simulators, providing many scenarios, enable large-scale testing of SDVs practically possible up to millions of miles every day, which can help not only detect problems readily at early stages, but also improve the abilities of SDVs through self-learning. For this reason, an effective simulation test environment is an indispensable component of the research in testing the large-scale capabilities of SDVs with vast numbers of driving scenarios. Therefore, manufacturers are building virtual test environments with simulation engines to test and improve SDVs' capabilities before testing them in real-world environments. For instance, the Open Racing Car Simulator was developed to generate real-world-like scenarios to test the performance of AVs. The fidelity of the vehicle simulation engines, with an unlimited number of real-world-like scenarios, has advanced significantly [37] by creating synthetic training data using randomisation and generative adversarial networks where the collection of real-world data that represents self-driving conditions could be high cost. Some of the simulators that are capable of generating synthetic data, providing intensive testing with vehicle state and dynamics and simulating sensors with high fidelity are CARLA [38] for camera and LIDAR, Constellation [39] for camera, LIDAR and radar and DRIVE Sim [40] for camera, LIDAR, radar and ultrasonics. The comprehensive analysis of the AV simulators involving the ones enabling the test of real vehicles in the simulation circuit is conducted in [41]. Publicly available 37 driving datasets and 22 virtual testing environments by which the self-driving techniques can be readily tested are analysed in [42].

Simulation engines cannot completely replicate the complex interactions of AVs with the real world [37] in every possible scenario where the simulation to real-world transfer learning domain focuses on bridging the simulation-reality gap. In other words, testing all critical aspects of SDVs and transfer learning from simulation to reality may not be possible using simulators, in particular, testing the sensing and actuation capabilities. It should be noted that there is no adequate simulation technology to produce accurate and validated models of sensor phenomenology at the physical level (e.g., electromagnetic wave propagation effects that determine performance limitations), which is the main drawback of testing sensor systems, their critical performance limiters in simulation environments. Hence, in addition to virtual simulation tools, test sites (i.e., fake cities), that mimic real-world environments with challenging scenarios, are being built to prepare SDVs for real-world tests and use. In this vein, Mcity developed by the University of Michigan in collaboration with General Motors, Nissan, Ford, Toyota and the local government [43] is primarily used to measure the capabilities of AVs and their interaction with one another and pedestrians. The US DoT designated 10 test sites that can be used by manufacturers to test their AVs within designated policies [44]. Besides, several manufacturers such as Toyota [44], Uber [45], and Waymo [46] have also built their test sites. The test infrastructures should be designed in a way that gives SDVs the chance to respond to all possible untrustworthy circumstances. However, artificial tests using computer simulations or lab tests using fake cities cannot provide sufficient evidence in proving that the required safety assured by standards is reached. In this regard, the following subsection covers the real-world test environments to mitigate the aforementioned shortcomings.

2) REAL-WORLD TESTS

SDVs are being massively tested in very complex real-world environments encountering real-world challenges to ensure their efficacy and safe use in urban areas. Several countries such as France, China, the UK [24], and several U.S. states (e.g., California, Arizona, Nevada, Georgia, Washington, Texas, Michigan, Ohio and Pennsylvania) [47] established their legal framework to allow manufacturers to test their AVs in real-world environments with no backup drivers. The French PSA Group in Paris and Amsterdam, Nissan in the UK [48], BMW in collaboration with Baidu in Beijing and Shanghai, Google in California [24], PerceptIn in Japan and China [49], Waymo in Arizona [47] and Mercedes in Mannheim and Pforzheim [50] are testing their particular ADSs with no in-vehicle drivers.

Can the reliability of SDVs be measured through real-world test driving? The total number of accidents of 48 Google AVs is 12 in around 2 million miles [51]. The Baidu Apollo fleet of nearly 500 autonomous driving vehicles have driven more than 7 million kilometres with zero accidents and have safely carried more than 210,000 passengers as claimed by Baidu [52]. It has been estimated by RAND corporation [53] that SDVs need to be driven 275 million failure-free miles (11,000 times around the world) to assure a similar rate of reliability (i.e., 95% confidence level) as existing human-driven cars. However, adapting to the environment for SDVs is not a linear learning process. Swarm intelligence can be constructed by combining the learning abilities of all SDVs, which reduces the required learning time significantly regarding the number of SDVs in operation [54]. The more evolutionary SDVs (e.g., learning from their mistakes or mistakes already done by others) for improving their performance [54], the more trusted ADSs. A mistake done by an SDV is not repeated by other SDVs through continuous non-linear learning. In this regard, again, it was calculated by RAND that with a fleet of 1000 SDVs, this can be reduced to 12.5 years. Therefore, there is no point in worrying about these statistical big failure-free miles since there are already many ADSs on the roads involving level-4 and level-5 vehicles in operation (Table 1 in [1]), not to mention the data obtained and experience gained from thousands of level-3 vehicles on the roads used in real-life. Moreover, the rigorous lab and artificial tests performed in advanced simulated environments can reduce these large failure-free miles significantly by solving the problems before the real-world road tests. Furthermore, their interconnection with SC facilities (e.g., traffic signalling) as mentioned earlier in Section III-A is expected to reduce the error rates of SDVs substantially, which is elaborated in Section III-D3.

D. TRUST FROM THE PERSPECTIVE OF USERS, PEDESTRIANS, AND OTHER TRAFFIC PARTICIPANTS

(Figs. 1 D, 2 D). SDVs with great expectations are promising much more than self-driving. Concurrently, SDVs come with numerous benefits that could transform our daily life beyond self-driving and enhancement of road traffic flow. Human trust in SDVs can be defined as the level of human confidence in them in meeting humans' expectations and desired performance indicators reliably through observations and perceptions. Building trust between humans and SDVs is paramount in encouraging and facilitating this inevitable transition from the human-oriented to the vehicle–oriented concept seamlessly with a high level of negotiation between all parties, in particular, with increased human-vehicle cooperation through well-established formulated metrics by considering human factors. These metrics have to ensure that trust develops over time targeting the trustworthiness of SDVs as the interrelation between the SDVs and other entities evolves through growing interactions. In other words, each party can predict the imminent actions of the other party when they know each other further, which help increase the trust in SDVs gradually with desired predicted performances under uncertainty as time passes. Intelligent AI approaches, equipped with easy-to-use interfaces supporting effective communication and coordination with conventional vehicles, emergency vehicles, pedestrians, and all other traffic participants involving police and traffic signs, are required to orchestrate traffic flow with increased collective trust in a rapidly changing environment. Collective trust can be built readily if all parties feel that the autonomy is behaving as expected and taking their safety seriously. Building confidence is necessary for humans both inside and outside of vehicles, but, not sufficient to build a high level of trust in SDVs as explored in the following sections. SDVs and all traffic participants should understand one another to trust each other with advanced human-vehicle interfaces/interactions (HVI) leading to effective mutual communication mechanisms. Within this context, the purpose of the UK government investment in SDVs (£10m) is not only to encourage the development of new SDV technologies, but to measure their potential impact on citizens and all other traffic participants [51].

1) TRUST FROM THE PERSPECTIVE OF USERS

(Fig. 2 D.1). Research suggests that around 70% of drivers expressed interest in testing SDVs, and around 60% of them indicated that they may replace their vehicles with SDVs where younger citizens are much more receptive to SDVs [55]. Nearly 60% of citizens around the world wished to travel in SDVs in a survey conducted by the World Economic Forum [18]. More specifically, the level of trust in SDVs changes with different age groups, sexes and even countries. For instance, in the consumer survey among 5,500 citizens in ten countries, acceptance of SDVs was the highest in emerging markets, such as India, China and the United Arab Emirates; it was around 50% in the US and the UK; it was the lowest in Japan and Germany [18]. In another research, in Japan, 33% of respondents would not ride in SDVs, whereas it was only 3.1% in China [56]. In another survey, drivers ages 18–35 were three times as likely as those 55 and above whom wanted to replace their current vehicle with SDVs [55]. On the other hand, around 73% of citizens in the US didn't trust in SDVs, according to a 2018 survey conducted by the AAA [57]. Under the shadow of these conflicting survey results, based on the past experiences of human acceptance of the safety-critical technologies (e.g., automatic elevators replacing manned operators or cars replacing horses and carriages), it can be safely concluded that users may accept changes rapidly wherever there are significant benefits to them through trustworthy systems despite the strong cultural attachment to current technologies.

The SDV technology, with a highly increased human-vehicle relationship, promises a variety of benefits, such as i) the reduction of accidents and fatalities resulting from human errors (e.g., driving fatigue, distraction) and ii) comfort of passengers with the optimised and standardised manoeuvrers (e.g., most convenient turning, acceleration, deceleration). Moreover, SDVs are being designed to achieve a variety of tasks such as self-delivery, ride-sharing and individual and public transportation. They, capable of taking over our daily routines (e.g., delivering children to their schools and picking them up) can free us to focus on other things. Rich customisable features (e.g., multimedia delivery) instilled in SDVs would help develop trust against them substantially with further benefits. The most vulnerable people in the society (e.g., disabled, partially sighted, children, elder people) can be independent with SDVs supported by appropriate in-vehicle assistive technology. Brewer and Kameswaran [58], Brewer and Ellison [59], and Fink *et al.* [60] provide insights on the challenges and potential barriers to blind and low vision people's adoption of SDVs by designing in-vehicle assistive prototypes using voice-based and tactile solutions to address the perceived barriers. Wearable health technologies (e.g., body temperature, heart rate and respiration monitoring) embedded into SDVs and integrated with smart healthcare can transform SDVs into micro health centres with advanced autonomous abilities and preemptive treatments by monitoring passengers' vitals and alerting them and their healthcare providers autonomously where abnormalities are detected. Furthermore, the manufacturers are producing customisable SDVs benefiting the particular needs of users to increase trust among users with different gender and age groups from different cultures. For instance, their customisable cabin can turn into a place to conduct business, continue working via face-to-face calls, socialise or relax with customised content delivery via 5G/6G within millisecond latency; the colour of the Mini can be changed based on the user preferences with shared ownership; BMW's interior concept with separate private sections can allow each passenger to commute as they please [61].

Understanding and detecting change in trust during interactions with autonomous systems is crucial to improving their design [5]. Similar to in-vehicle wearable health technologies mentioned above, trust in SDVs can be measured by reading body kinematics (e.g., electroencephalogram (EEG), electrocardiography (ECG), electromyography (EMG), respiratory/heart rate) using gesture recognition and wearable sensors embedded in SDVs. These devices can help detect instant trust parameters per particular task performed by SDVs. For instance, Wu *et al.* [62] propose an emotion recognition approach via the cloud and edge platforms, which, along with similar approaches, may lead to the measurement of the trust level for SDVs. Similarly, heart rate variability using ECG was observed in the simulated AV driving to measure the confidence in this type of autonomy [63]. A methodology for validating user experience (UX) in AVs, based on the information acquired from physiological signals (i.e., galvanic skin response (GSR)), in order to customise various aspects of HVI while the user is immersed in a virtual reality-based driving simulation is proposed in [64].

Trust in the automated system increases with the introduction of knowledge about the true capabilities and limitations of the automated system [11] where the user expectations and high hopes can be calibrated. Manufacturers must see through the limitations of commercialised SDVs considering their competence while they are communicating their vehicles' benefits to customers. Consumer trust certainly cannot be bought, it can be gained using appropriate approaches. An EU-funded project, namely, TrustVehicle,[5] aims to provide solutions for both increasing the reliability and trustworthiness of AVs and contributing to end-user acceptance by following a user-centric approach with advanced technical solutions. The conceptualisation of the main trust components between SDVs and users is presented in Fig. 2 D.1. The building and maintaining trust in SDVs from the perspective of users involving the other main trust components extracted from Fig. 2 is demonstrated in Fig. 3. It is highly critical to gain trust to communicate information to users in ways they can understand and communicate feedback about whether their information is understood by users to developers. User trust can be primarily established by providing ''full control'' of SDVs with an advanced, easy-to-use HVI along with safe and comfortable riding. HVI, equipped with inner-vehicle and inter-vehicle interaction abilities, has to collaborate and cooperate with the user effectively using various communication modes (e.g., touch, gesture, speech) providing a high level of control consistent with the vehicle dynamics and high-quality UX. It is important to mention that user requests under ''full control'' ability can be only conducted if they are safe and legal by which users are shielded against liability in the event of disastrous consequences.

Any failure during the real-world use of SDVs can be mended to prevent it from happening again by determining what went wrong based on the huge amount of data acquired from such a high-tech vehicle. This ability leads to maintaining the already established trust through instant iterations and improvements. Furthermore, any imminent and instant non-serious failures can be fixed remotely and quickly using the human-on-the-loop (HOTL) approach [54] by considering the remote non-negligible security threats, which is elaborated in Section IV. Technically speaking, the HOTL abilities with remote operators through taking one course of optimal action safely over other less optimal actions by maximising the expected utility and expediting self-learning based on achieving certain goals can maintain and ensure trust, in particular, at the start of their commercial use.

2) TRUST FROM THE PERSPECTIVE OF PEDESTRIANS

(Fig. 2 D.2). Trust between humans and machines could be built in a manner similar to the development of trust between humans. If SDVs can communicate their understanding back to humans around them, that will cement the trust-building process [5]. Pedestrians with different types of crosswalk patterns are the most vulnerable traffic participants. Their interactions with AVs have been recently explored in [65] and [66] in various aspects. A vehicle-pedestrian negotiation model is proposed in [67] describing the exchange of negotiation cues from both parties to speed up the traffic flow. The conceptualisation of the main components between SDVs and pedestrians is presented in Fig. 2 D.2. Instant identification of pedestrians and better predicting their actions using advanced complementary sensors and vehicle-to-pedestrian (V2P) communication mechanisms are vital in building a high level of trust.

We have more of a liking toward robots that harbour humanoid features. With this in mind, manufacturers are adding humanoid visual cues to SDVs leading to the

[5] https://www.trustvehicle.eu/

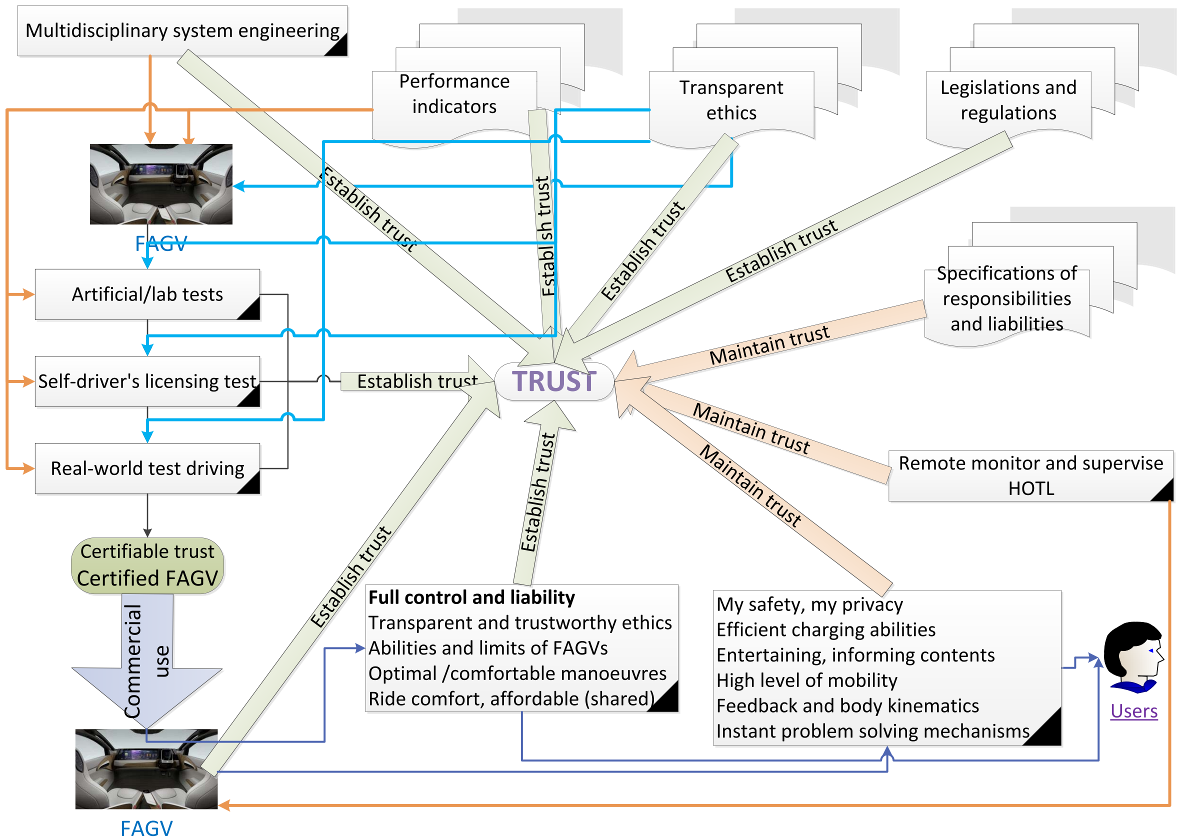

FIGURE 3. Building and maintaining trust in SDVs from the perspective of users.

predictability of the vehicle's intentions. For instance, the engineering team at Jaguar partnered with cognitive scientists to develop a trustworthy solution by putting huge googly eyes on the front of its prototype vehicle — a set of digital eyes that act like human eyes — and these faux-eyes follow the objects they see, which communicate with pedestrians [57]. This helps pedestrians trust SDVs more instinctively when they try to cross the street. As an efficient HVI option, SDVs, having digital windows, can inform pedestrians about their upcoming actions. They can speak to other traffic participants and pedestrians. By adding visual cues to the exterior of the car, such as LED bars, bumper displays (safe to cross), mirror screens, virtual drivers and projectors with zebra-crossings manufacturers attempt to fill up the information gap left by visible human drivers (e.g., eye contact, gestures) [68]. Similar innovative interfaces established between SDVs and pedestrians will help gain and maintain trust in SDVs substantially for all parties.

3) TRUST FROM THE PERSPECTIVE OF OTHER TRAFFIC PARTICIPANTS

(Fig. 2 D.3). In addition to establishing effective communication and behaviour predictability as explained in Section III-D2, the acceptance of SDVs by all traffic participants is strictly dependent on addressing safety concerns with this new technology. There is some degree of negotiation between SDVs and other traffic participants, and false assumptions in those negotiations, particularly, inexplicable movements from the point of SDVs' view may result in crashes [1]. Because of this challenge that reduces trust in this autonomy, SDVs are thought to be deployed in restricted regions or on a geofencing basis rather than letting them mix with daily traffic. The designers of SDVs must consider how their choices impact human drivers as well as their own vehicles' passengers [69] to reduce the conflicts and to provide a required level of trust between these two entities where the risks can be kept at an acceptable level by predicting intentions properly. Without well-established intelligent coordination tools and techniques between SDVs and other traffic participants, the collective trust can not be gained at a high level and maintained in the long run.

The conceptualisation of the main trust components between SDVs and other traffic participants is presented in Fig. 2 D.3. Understanding the intricate behaviour of other road participants (e.g., human drivers, ambulances, cyclists, police, SC ITS infrastructure) using scene prediction techniques through semantic segmentation and motion planning is a major challenge where their subsequent behaviours are unknown and modelling of this uncertainty for reasoning via sensor fusion is a non-trivial task. This can be alleviated by analysing their instant actions (e.g., velocities, direction, signals) and by predicting their intentions or at

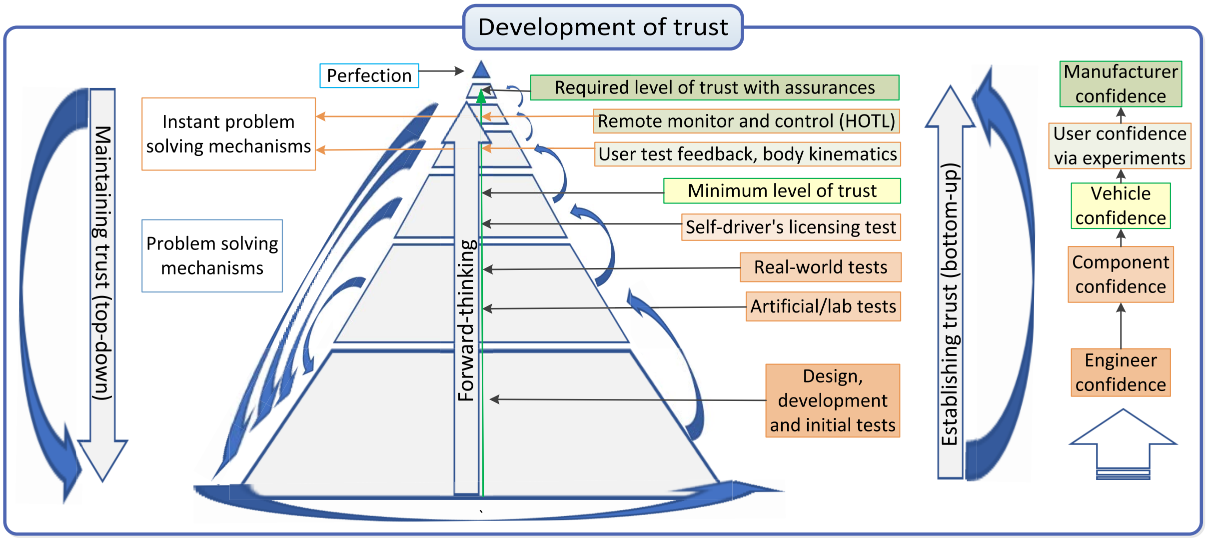

FIGURE 4. Structured block diagram of reaching the desired trust level with a trust loop. The magnitude of each trapezoid within the pyramid corresponds to the responsibility of increasing trust in SDVs.

least their imminent acceptable behaviours/motions using advanced analytics.

SC ITS infrastructure along with advanced 3D HD mapping involving intelligent digital traffic sign messaging is aimed to make SDVs behave as desired with optimal actuation mechanisms in every type of weather condition (e.g., snow, rain, fog, hail) day and night [1], which will help build trust between SDVs and other entities. Taking proper actions regarding ambulance waypoints and police directions can be managed through this infrastructure as well. This area still requires extensive research. The required unambiguous collaboration and cooperation between SDVs can be established using the concepts of CAV and IoE through V2E with agreed-upon standards, protocols and communication links to avoid accidents and provide optimal actuation mechanisms. Computer drivers can have "telepathy": A computer driver could let another computer driver know that it is considering changing lanes before deciding to do so; it could communicate to traffic lights to minimise wait times at intersections and optimise traffic flow [70]. The more desired actions with proper manoeuvres, the more trust can be gained.

IV. VERIFICATION OF TRUST

First and foremost, the consolidation of trust can be provided by ensuring the desired objectives of SDVs with the distributed liabilities and responsibilities of all the stakeholders (Fig. 2) as mentioned earlier. Second, well-devised verification mechanisms with simulation, and real-world interactions are the key to building reliable and trustworthy SDVs. It is worth emphasising that building trust can be achieved in the long run with repetitive realisable system behaviours using standardised assurances, but, it can be lost easily with a simple disappointing action or a negative outcome in an unpredictable and uncertain environment. What's worse, trust is hard to regain, once lost. It has to be earned based on the ability to demonstrate safe desired operations. Therefore, the development of trust is a dynamic process and needs to be calibrated to the correct levels for safe deployment to ensure the appropriate use of such systems [11]. Maintaining trust, as an ongoing process, can be secured by instilling the permanent improvement abilities with standardised procedures in every phase in a timely manner from the development of the technology to its use. In this sense, the proposed structured block diagram of the trust loop by which the desired trust level can be reached step by step and maintained is shown in Fig. 4 with a high level of reassurances within a series of hierarchy verifications using a forward-thinking mechanism. The quality of services needs to be increased using the backward problem-solving mechanisms starting from the first phase, "design, development and initial tests" if the desired level of quality standards is not reached in the specified phases. With these iterative phases, the main objective is to ensure the first minimum level of vehicle and manufacturer self-confidence (e.g., the minimum level of trust) by computing the formulated task-specific metrics for safety and functional assessments within the initial design, artificial/lab, real-world rigorous tests and SDLT that are elaborated in Section III. Then, the desired level of trust can be reinforced by encouraging user interactions (their feedback and measurement of their body kinematics) and most importantly, with the ability to fix any unexpected problem remotely with instant problem-solving mechanisms using the HOTL vehicle teleoperation (remote driving).

Vehicle teleoperation with instant problem-solving mechanisms using remotely authorized controllable abilities seems to be a viable solution to enable further improvements [54] and to maintain trust in SDVs during their real-world use, i.e., trust by use. It, with increasing momentum in the industry, allows SDVs to navigate through unorthodox situations that they cannot deal with on their own [71]. Major manufacturers, together with telecommunication companies, are developing vehicle teleoperation systems [71]. Waymo replaced in-vehicle human operators with human teleoperators [71]. Baidu integrated 5G-enabled teleoperation into its vehicles to ensure public safety in extreme road conditions. All the remote human operators of Baidu complete more than 1,000 hours of cloud-based driving training without any accidents to ensure the safety of users and pedestrians when the non-autonomous driving mode is engaged [52]. Drive.ai initiated its pilot SDVs in Texas by keeping humans-in-the-loop (HITL) [72] with remote interaction and intervention abilities (e,g., vehicle performance monitoring, stepping in emergency situations). Ericsson and Einride built a remote HOTL teleoperation system with SDVs in the delivery sector in Sweden [73]. The results of the trials in real-world environments suggest that it is highly imperative to keep HOTL to step in when the new driver, AI, experiences a complex situation that can not be handled by autonomy [54]. Despite decades of prior research and a renewed interest from technology companies and the research community, many gaps remain in the capabilities of AVs [37], in particular, performing evolutionary cognitive computing in the event of uncertain outcomes; this is what makes the HOTL concept a reasonable intervention to maintain trust in SDVs. Therefore, human remote problem-solving abilities should be incorporated into SDVs during the design and development phases of SDVs to help calibrate user trust, and retain the already gained trust during the use of the technology by taking the ethical and most importantly privacy issues into consideration which are elaborated in Section V.

V. ETHICAL ASPECTS OF TRUST IN SDVs

AI is perceived as trustworthy when it is developed, deployed, and used in ways that not only ensure its compliance with all relevant laws and its robustness but especially its adherence to general ethical principles [9]. Ethics as the other major pillar of trust (Fig. 1) enabling ethical SDVs to help reinforce their trustworthiness. Major technology companies (e.g., Microsoft, Google, IBM, Sony) are building their own ethical frameworks to meet the social challenges and increase the trust in their products by targeting long-term company-based trustworthiness while the responsible governmental bodies lag behind in specifying the ethical criteria in the development of particular AI products in cooperation with the stakeholders and enforcing them at the strategic level. Despite an apparent agreement that AI should be 'ethical', there is debate about both what constitutes 'ethical AI' and which ethical requirements, technical standards and best practices are needed for its realisation [74]. Ethics has been evolved and built on experiences over many years and the social governance of AI instilled in SDVs can be implemented by incorporating them into SDVs leading to the establishment of effective communication and relationship between human-centred SDVs and society, and then trustworthy SDVs with socially acceptable norms. Should the use of SDVs with a fully autonomous mode be allowed legally and ethically? Up until the present, the ethics for vehicles were embedded in humans and the rules of the road, not in software and not in the vehicle and this will profoundly change with SDVs, and as such, the opportunity for people to embed their ethics in their own SDVs (or any other form of automated algorithmic assistance) may not be considered or may be considered and discarded [75]. Currently, there is no technical evidence that AI in SDVs would be forced to make such ethical decisions, whether inherently based on software architecture or learned via training methods [25]. For both engineers and the general public, a car's decision-making system should weigh the ethical implications of its actions [76]. Main inputs in the ethical transition to autonomous driving, while building trust from the perspective of all stakeholders, are introduced in Fig. 5.

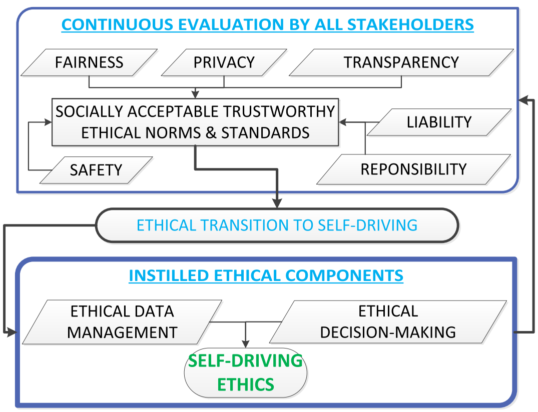

FIGURE 5. Evolving ethical transition to self-driving through a dynamic structure.

Ethics in autonomous robotic applications has been immensely researched for several decades. IEEE started a global project with the P7000 standard series to address ethical concerns during the system design and development of intelligent autonomous systems [77]. More concretely, P7003 algorithmic bias considerations, P7007 ontological standard for ethically driven robots and automation systems, and P7010 Wellbeing Metrics Standard for Ethical Artificial Intelligence (AI) and Autonomous Systems. Besides, the IEEE AI Ethics initiative has developed Ethically Aligned Design (EAD) aiming to set global standards for ethics in autonomous and intelligent systems (A/IS).[6] Similar initiatives would help develop agreed-upon ethical standards and protocols for SDVs. In the same direction, ethics specific to SDVs has been recently focused. The ethics in SDVs are analysed in [78] from a philosophical point of view based on the concepts of utilitarianism, deontology, relativism, absolutism (monism), and pluralism. Recommendations on ethics of CAV, particularly, road safety, privacy, fairness, explainability and responsibility raised by driverless mobility have

[6]https://ethicsinaction.ieee.org/

been recently presented by the Horizon 2020 European Commission Expert Group [79]. The report (E03659), with a set of 20 ethical recommendations concerning the future development and use of CAV, aims to promote a safe and responsible transition to CAV by supporting stakeholders in the systematic inclusion of ethical considerations that are grounded in the fundamental ethical and legal principles laid down in the EU Treaties and Charter of Fundamental Rights.In a broader perspective, ethical guidelines for trustworthy AI[7] were published by the European Union aiming at shaping Europe's digital future by bridging the ethical gaps in the implementation of trustworthy AI. The Guidelines put forward a set of 7 key requirements that AI systems should meet in order to be deemed trustworthy — human agency and oversight, technical robustness and safety, privacy and data governance, transparency, diversity and non-discrimination and fairness, societal and environmental well-being, and accountability. According to the guidelines, trustworthy AI from a general point of view should be: i) lawful — respecting all applicable laws and regulations, ii) ethical — respecting ethical principles and values, and iii) robust — from a technical perspective while taking into account its social environment.

The lesson learned during an intercontinental driverless trial from Italy to China (13,000 km) is that SDVs not only have to comply with traffic rules and regulations, but also need to determine when to violate them [80] for the purpose of protecting themselves and other traffic participants. Sometimes, instantaneous judgement assisted by wisdom can induce us to violate the laws, rules and regulations for the benefit of the society and should an SDV be allowed in the same way [81]? Google has acknowledged allowing its SDVs to violate speed limits to keep up with the traffic flow where lower speeds can jeopardise the traffic flow, which may lead to accidents [76]. In a similar decision-making context, ethically driven principles must be built into AV algorithms that determine how scenarios would be ranked and retrieved in various critical driving situations [82]. There are various moral judgment concerns along with several challenges to consider under a utilitarian framework before taking our hands off the wheel and before targeting a global SDV market. How SDVs decide the best course of action under a moral dilemma framework or how they dictate their behaviour under emergencies needs to be programmed based on ethical and legal aspects established with the involvement of all the stakeholders where there is still ample time until the first commercialised SDVs hit the public roads. Substantial literature can be found on the famous ''trolley problem'' and an agreed-upon solution is yet to be found even though it seems rather simpler than many other dilemmas. Whom or what to prioritise within moral dilemma — e.g., passengers in the vehicle or pedestrians; choosing to kill one person instead of 3 people; which vehicle to strike first if the accident is inevitable (the big guilty truck or the small innocent car). Mercedes-Benz always aims to prioritise protecting its passengers under any dilemma in decision-making [25]. According to the study [83], AVs must make reasonable decisions to reduce causalities and property damage, where, of course, life is the priority; more concretely, AVs must make the decisions to find the best outcome possible, without doing an assessment of the life value of each actor because any human life has the same value. Thus, the goal is to minimise the quantity of casualties [83] and the number of entities collided with. In this vein, according to E03659 [79], fair distribution of risk and the protection of basic rights, including those of vulnerable users should apply to dilemma scenarios. How transparent are manufacturers supposed to be about the ethics embedded in their AI driver, particularly in highly critical and paradoxical traffic conditions [84]? Ethics in building trust against SDVs will be immensely analysed in the years to come, mainly from a moral and cultural perspective, suggesting many definitive outcomes, particularly, on ethical decision-making in the cases of unavoidable accidents where life and death matters.

Who is going to be responsible in the case of an undesired outcome? What happens if any unexpected failure happens out of the agreed-upon context? The social norms and distribution of responsibilities among all stakeholders under any dire circumstances (e.g., accidents, violation of privacy rights) should be outlined well. On one hand, if the software is the main controller of the vehicle behaviour and AI algorithms are owned by manufacturers to protect their intellectual property, then liability may also be in their custody as well [82]. On the other hand, neither the manufacturer nor the passengers could be held accountable for behaviours if conflict-solving algorithms and risk preferences are performed in advance by societal consensus, leaving no room for individual intervention [22]. In other words, since the vehicle's decisions and reactions follow socially established norms, manufacturers can no longer be held responsible for damages that occur as a consequence of these norms [22]. In accordance with E03659 [79], considering who should be liable for paying compensation following a collision is not sufficient; it is also important to make different stakeholders willing, able and motivated to take responsibility for preventing undesirable outcomes and promoting societally beneficial outcomes of CAV.

On one hand, the more data collected from SDVs, the better training and learning and the better the decision-making abilities of these vehicles. The vehicle itself would be a repository of personal information about everywhere its user had travelled, how the vehicle had travelled, and everything encountered along the way [85]. On the other hand, more data collecting and sharing means more unauthorised access and probable violations of privacy and sensitive information (e.g., passengers' private data, camera recordings). The privacy of the user has to be protected using effective privacy-preserving applications equipped with encryption or anonymisation mechanisms directed by the legal and regulatory framework (e.g., European General Data Protection Regulation, Data Protection Act) by which the collected

[7]https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai

vast amount of data is processed appropriately. Additionally, crowdsourcing applications with SDVs in a context-aware fashion with mobile sensing should be available by mitigating the privacy and data protection concerns of other road participants. In this context, the good governance of data acquired from many onboard sensors, in particular, ethical privacy must be incorporated into the rules and regulations mentioned in Section III-A. Strictly speaking, privacy within the ethical concept of vehicle intelligence is the primary concern beyond technology to be addressed properly while exploiting SDVs effectively. The author refers readers to [85] for the legal evaluation of privacy analysed specific to AVs. According to E03659 [79], the acquisition and processing of static and dynamic data by CAV should safeguard basic privacy rights, should not create discrimination between users, and should happen via processes that are accessible and understandable to the subjects involved.

Deploying an autonomous platform that has proven to be safe and reliable under rigorous testing in stringent conditions would, on the whole, constitute an acceptable, even a morally responsible, action [86]. Evolutionary intuitive abilities instilled in SDVs should not be based on individual moral concepts and diversity of ethical approaches. They have to use standard ethical transparent rules in a socially acceptable way embraced by all stakeholders since SDVs cannot be perceived only from a technological point of view, but also, from a socio-technical point of view. This can only be realised by moulding the requirements of the stakeholders in the development of SDVs around agreed-upon concepts from an ethical point of view as an integral part of the technology development where ethical and legal perspectives are perceived as the primary concerns beyond technology, but can not be separated from the design of the technology in ensuring that SDVs are being developed ethically at every stage of the delivery and deployed morally and responsibly at each stage of their implementation lifecycle. It is worth emphasising that this is not an easy process, but it has to start from a point in which all stakeholders should show willingness to be forward-thinking through the ethical transition to autonomous driving. One ethical condition, however, should be crucial: in no case should the ethical algorithms be put into practice as non-transparent black-boxes [22]. To conclude, ethics embedded in driverless mobility applications should be determined in agreement by all parties for avoiding poor ethical design and should be declared explicitly.

VI. DISCUSSION

Despite the concerns and anxieties caused by the introduction of new digital technology into the automotive industry [87], SDVs are perceived as a ‘‘somewhat low risk’’ form of transport and there is not much opposition to the prospect of their use on public roads [88]. It is also commonly accepted that the first SDVs should be demonstrably safer than a vehicle driven by the median human driver [21]. It is envisioned that the investment in SDVs will accelerate in the years to come until human intervention is eliminated from the system within the conflicting opinions: i) ‘‘SDVs will never be perfect as human drivers’’, ii) ‘‘SDVs will be operating better than human drivers using well advanced in-vehicle technologies, in particular, in challenging conditions (e.g., difficult weather conditions, dark environment, tiredness)’’. It is critical to state that trust can be established based not only on the improved safety and functionalities, but also on the increasing benefits specific to the needs of the individual. Aside from economic rationality, the availability of ubiquitous, affordable, on-demand, and automated personal transportation along with other benefits dealing with fueling, parking, registrations, insurance, asset depreciation, maintenance, breakdowns, traffic citations, and road stress is likely to inspire some consumers not to own a vehicle by simplifying their lives [89] with the well-established trust.

It was boldly envisioned that by 2040, all vehicles will be driverless by eliminating the human out of the loop [90]. The level-5 concept SDVs are yet to be deployed to real-world commercial use. The manufacturers along with the press shall not ignore the real technical limitations of the SDV technology when communicating about benefits to their customers while aiming at establishing trust between the stakeholders where still there are serious limitations even with the commonly used ‘‘autopilot’’ feature. Today, the public became more realistic about the potential of SDVs despite the high expectations and speculative future hopes during the last decade. Manually steered vehicles will never be replaced with SDVs if trustworthiness among the stakeholders is not established sufficiently where trust plays a pivotal role in deciding to buy or use a product. Explainable AI (i.e., XAI) approaches leading to reasoning with interpretability for AVs are in high demand among the stakeholders to build trust further (e.g., achieving public approval by regulators) with transparency in real-time decisions where DNN approaches are known to be black-box systems. AI of AVs is expected to explain how decisions are constructed. Furthermore, correcting the mistakes that are made by AV systems based on ML is far more difficult than visualised, because it is not obvious what needs to be changed inside the black box to fix the problem, which is a very serious challenge that has not mastered yet. Trained networks for specific locations may not be as accurate at the same rate in other locations, which requires location-specific data collection and retraining. SDVs will never be perfect. We can get closer to perfection with proper actions, e.g. if the data sets and knowledge on SDVs gained by various manufacturers so far are shared within joint venture corporations based on the best practices. This will reduce the cost of manufacturers substantially as well. Various experiences and know-how along with the large data sets obtained by various manufacturers can be merged to help improve the decision-making abilities of SDVs leading to the establishment of the desired trust in the commercial use of SDVs. For instance, Ford and Waymo opened a tiny, carefully curated, a subset of their vast data sets collected during autonomous-vehicle tests to the general public and challenged developers to use them to come up with faster and smarter self-driving algorithms [91].

In practice, autonomous driving (AD) consists of multiple tasks with conflicting multiple objectives and an optimal decision is supposed to be predicted at each instant while driving in this high-dimensional space with sequential decision-making. AI, neuroscience, and psychology synergistically interact in cognitive science [92]. Recent revolutionary advances in AI using the learning principles of biological brains and human cognition has fuelled the development and use of Deep Reinforcement Learning (DRL) in numerous fields by both leveraging the powerful generalisation ability of Deep Neural Networks (DNN) and the self-learning ability of Reinforcement Learning (RL). The success of several AI applications such as i) games: 49 classic Atari [93], self-taught AlphaGo [94], maze [95], ii) solving real-world problems: large-scale traffic grid signal control [96], pedestrian regulation [97], nonlinear discrete-time systems [98], developed by DRL demonstrates that the machine can not only be close the intelligence of humans, but can also be smarter than humans, and these breakthrough works give DRL a high level of confidence. However, the complexity dimensions of these applications, mainly using a single reward function, are fully observable, single/double-agent, deterministic, sequential, semi-static/semi-dynamic, discrete and known environments. Most importantly, an explicit single-objective reward function (e.g., the distance between the monster and the hero) can be readily given to the agent easily which makes the problem space extremely less complex in those applications compared to self-driving (partially observable, multiagent, stochastic, sequential, dynamic, continuous and unknown environment) where the determination of explicit multi-objective (related or independent that can be combined; conflicting requiring trade-off) and multi-agent (compete or cooperate) reward functions is a non-trivial task regarding i) the learning of the desired behaviour requiring the simultaneous satisfaction of multiple objectives with a reward vector rather than a scalar reward signal, which may lead to a set of policies ii) the scaling up the determined policies to larger and more complex SDV dynamic environments with the increasing computation cost.

DRL is expected to make behaviour learning and collaborative learning easier and faster in SDVs with further improvements, in particular, with Multi-Objective DRL (MODRL) where SDVs perform inherently with multi-objective tasks in nature. On the other hand, the learning process in DRL requires too large data samples for leading to a reasonable and generalisable policy for varying environments. Moreover, the large, unbalanced distribution of observations in data samples is the major drawback in building a generalisable policy with DRL. Despite the huge improvements in AI, still, intelligent agents have very limited abilities in cognitive learning when compared to humans by whom specific tasks can be learnt in a matter of minutes by processing the prior gained knowledge and experiences that are generalisable to new scenarios. Evolutionary self-learning techniques with swarm intelligence containing the optimisation of the cooperative and collaborative actions of swarms of SDV agents via sensing, learning, reasoning, communicating, and acting within the concepts of CAV, IoE and AoE leading to better decision-making will help develop trust in the capabilities of SDVs further. Within this context, learning with supervisory interaction is highly useful where there are no prior datasets and it may not be feasible to establish sufficient training. Collaborative learning during interaction using cognitive computing can help SDVs to gain the ability to rapidly master the particular fine-granular tasks during the collaboration modes of the HOTL teleoperation instantly. In this direction, learning from demonstration (LfD) [99] (i.e., programming by demonstration [100]) lets intelligent agents acquire new skills to imitate a skilled human without needing to be programmed by engineers, i.e., by facilitating non-expert agent programming whereas programming by cooperation (PbC) [100] focuses on making intelligent agents learn tasks from humans by interacting and cooperating with them by alleviating the labour-intensive task of exploration. LfD and PbC can be chosen over other agent learning methods, particularly where ideal behaviours can be neither scripted (as is done in traditional robot programming) nor easily defined as optimising multi-objective reward functions (as is done in RRL), but can be demonstrated [99]. It is important to mention that learning from these methods is very restrictive in generalisation for varying scenarios.

Besides, customisable approaches embedded in SDVs that behave in an ethically responsible manner by solving moral dilemmas will increase trust in this type of autonomy, which accelerates the acceptance in many ways from logistics to private and public transportation. When SDVs meet the market, there will still be a lot of people who do not trust in SDVs. The question that needs to be answered is if this number would decrease or if new people would join this group. Vehicle teleoperation is expected to help consolidate trust between SDVs and other entities with the integration of 5G and beyond along with the SC fog/edge and cloud computing, in particular, in the initial phases of using them commercially. Humans with past experiences are accustomed to using some form of autonomous tasks in their everyday lives for a long time. Humans now depend on the safety, reliability, and security of intelligent systems. Therefore, I envision that society will accept the high level of vehicle autonomy rapidly as long as the trust building phases mentioned in this paper are properly taken into consideration by all stakeholders. To summarise, safety challenges and concerns are the prime issues for the public acceptance of autonomous services and these concerns, while bringing SDVs into the market, can be overcome considerably by incorporating the priorities of all stakeholders into the SDV development and commercialisation phases using comprehensive frameworks with formulated metrics and indicators to which a direction is provided in this paper with the proposed framework — TrustFSDV.

VII. CONCLUSION AND FUTURE DIRECTIONS

The trust deficit between non-autonomous/autonomous entities and SDVs should be decreased substantially for a smoother transformation to a future driverless society through robust human-vehicle integration and society-vehicle harmonisation. A high level of trust in SDVs cannot only be achieved by manufacturers, it requires the strict collaboration of all the stakeholders with distributed responsibilities and liabilities. Against this background, in this paper, the dynamics of trust in SDVs are formalised with a multi-dimensional framework, namely, TrustFSDV (Fig. 2). This framework provides the basis for building and maintaining trust between users involving all traffic participants and SDVs (Fig. 3). Besides, it shows the way of ensuring trust through well-designed iterative procedures (Fig. 4). It is noteworthy to mention that trust can be gained through hard work as explored in this paper with the proper directions of the stakeholders, but can be lost easily without forward-thinking during the life-cycle of SDVs (Fig. 4). In this direction, this paper shows the directions in building trust and maintaining it within a holistic structured framework — TrustFSDV — inducing the advancement of SDVs and acceptance of them by society.

The key findings in this research can be summarised as i) the fundamental enabler to build trust in SDVs is building reliable AV systems based on underlying human factors influencing people's trust in SDVs, ii) the manufacturers along with the press should not ignore the real technical limitations of the SDV technology when communicating about benefits to their customers while aiming at establishing trust between the stakeholders, iii) the manufacturers of SDVs should collaborate with cross-disciplinary fields (e.g., psychology, sociology, ergonomics, engineering, cognitive scientists) to develop trustworthy entities to meet the stringent requirements of all the stakeholders in designing and developing SDVs, iv) they should team up in numerous areas to expedite the pace of acceptance with increasing trust based on best practices, which will reduce their cost significantly as well, v) the more penetration of SDVs as a high-level rational decision-maker into mixed traffic and the more integrated with SC facilities and other traffic participants within the concepts of IoE, AoE and CAV, the more trusted SDVs with increased decision-making abilities and error-free actuation mechanisms, vi) drastic steps are needed to be taken concerning ethical and legal perspectives, which is perceived as the primary concerns beyond technology to build and maintain a high level of trust in SDVs, vii) the good governance of data acquired from the many onboard sensors, in particular, ethical privacy must be addressed properly and incorporated not only into the rules and regulations, but also into vehicle intelligence to exploit SDVs appropriately, viii) there is certainly a long way to go in building a high level of trust in SDVs regarding the strong cultural attachment to conventional vehicles and ix) sometime in the future, humans may forget how to drive physically, but will learn very well how to drive virtually/teleport/teleoperate (i.e., HOTL supervisory role) SDVs to easily step in when the new driver, AI, experiences an unorthodox situation that can not be handled by autonomy [54].

ACKNOWLEDGMENT

The author would like to thank the anonymous reviewers for their constructive comments which doubtlessly improved the quality of the paper significantly.

REFERENCES

[1] K. Kuru and W. Khan, "A framework for the synergistic integration of fully autonomous ground vehicles with smart city," *IEEE Access*, vol. 9, pp. 923–948, 2021.
[2] K. Kuru, "Management of geo-distributed intelligence: Deep insight as a service (DINSaaS) on forged cloud platforms (FCP)," *J. Parallel Distrib. Comput.*, vol. 149, pp. 103–118, Mar. 2021.
[3] S. Russell and P. Norvig, *Artificial Intelligence: A Modern Approach*, 4th ed. Upper Saddle River, NJ, USA: Prentice-Hall, 2020.
[4] K. Kuru and H. Yetgin, "Transformation to advanced mechatronics systems within new industrial revolution: A novel framework in automation of everything (AoE)," *IEEE Access*, vol. 7, pp. 41395–41415, 2019.
[5] S. Nahavandi, "Trust in autonomous systems-iTrust lab: Future directions for analysis of trust with autonomous systems," *IEEE Syst., Man, Cybern. Mag.*, vol. 5, no. 3, pp. 52–59, Jul. 2019.
[6] J. Xu, K. Le, A. Deitermann, and E. Montague, "How different types of users develop trust in technology: A qualitative analysis of the antecedents of active and passive user trust in a shared technology," *Appl. Ergonom.*, vol. 45, no. 6, pp. 1495–1503, Nov. 2014.
[7] K. A. Hoff and M. Bashir, "Trust in automation," *Hum. Factors, J. Hum. Factors Ergonom. Soc.*, vol. 57, no. 3, pp. 407–434, Sep. 2014.
[8] A. J. Jones, "On the concept of trust," *Decis. Support Syst.*, vol. 33, no. 3, pp. 225–232, 2002.
[9] S. Thiebes, S. Lins, and A. Sunyaev, "Trustworthy artificial intelligence," *Electron. Markets*, vol. 31, no. 2, pp. 447–464, Oct. 2020.
[10] A. Chavaillaz, D. Wastell, and J. Sauer, "System reliability, performance and trust in adaptable automation," *Appl. Ergonom.*, vol. 52, pp. 333–342, Jan. 2016.
[11] S. Khastgir, S. Birrell, G. Dhadyalla, and P. Jennings, "Calibrating trust through knowledge: Introducing the concept of informed safety for automation in vehicles," *Transp. Res. C, Emerg. Technol.*, vol. 96, pp. 290–303, Nov. 2018.
[12] F. Ekman, M. Johansson, L.-O. Bligård, M. Karlsson, and H. Strömberg, "Exploring automated vehicle driving styles as a source of trust information," *Transp. Res. F, Traffic Psychol. Behav.*, vol. 65, pp. 268–279, Aug. 2019.
[13] O. A. Kannape, A. Barré, K. Aminian, and O. Blanke, "Cognitive loading affects motor awareness and movement kinematics but not locomotor trajectories during goal-directed walking in a virtual reality environment," *PLoS ONE*, vol. 9, no. 1, Jan. 2014, Art. no. e85560.
[14] S. Nahavandi, "Trusted autonomy between humans and robots: Toward human-on-the-loop in robotics and autonomous systems," *IEEE Syst., Man, Cybern. Mag.*, vol. 3, no. 1, pp. 10–17, Jan. 2017.
[15] F. Ekman, M. Johansson, and J. Sochor, "Creating appropriate trust in automated vehicle systems: A framework for HMI design," *IEEE Trans. Human-Mach. Syst.*, vol. 48, no. 1, pp. 95–101, Feb. 2018.
[16] R. C. Mayer, J. H. Davis, and F. D. Schoorman, "An integrative model of organizational trust," *Acad. Manage. Rev.*, vol. 20, no. 3, p. 709, Jul. 1995.
[17] P. Andras, L. Esterle, M. Guckert, T. A. Han, P. R. Lewis, K. Milanovic, T. Payne, C. Perret, J. Pitt, S. T. Powers, N. Urquhart, and S. Wells, "Trusting intelligent machines: Deepening trust within socio-technical systems," *IEEE Technol. Soc. Mag.*, vol. 37, no. 4, pp. 76–83, Dec. 2018.
[18] E. Uhlemann, "Connected-vehicles applications are emerging [connected vehicles]," *IEEE Veh. Technol. Mag.*, vol. 11, no. 1, pp. 25–96, Mar. 2016.
[19] J. P. Trovao, "Trends in automotive electronics [automotive electronics]," *IEEE Veh. Technol. Mag.*, vol. 14, no. 4, pp. 100–109, Dec. 2019.
[20] Y. Ma, Z. Li, and M. A. Sotelo, "Testing and evaluating driverless vehicles' intelligence: The Tsinghua lion case study," *IEEE Intell. Transp. Syst. Mag.*, vol. 12, no. 4, pp. 10–22, Winter 2020.
[21] E. Coelingh, J. Nilsson, and J. Buffum, "Driving tests for self-driving cars," *IEEE Spectr.*, vol. 55, no. 3, pp. 40–45, Mar. 2018.

[22] D. Birnbacher and W. Birnbacher, "Fully autonomous driving: Where technology and ethics meet," *IEEE Intell. Syst.*, vol. 32, no. 5, pp. 3–4, Sep. 2017.
[23] USDOT. (2018). *Preparing for the Future of Transportation: Automated Vehicles 3.0*. [Online]. Available: https://www.transportation.gov/av/3/preparing-future-transportation-automated-vehicles-3
[24] M. Harris, "State of automation [news]," *IEEE Spectr.*, vol. 52, no. 9, p. 18, Sep. 2015.
[25] A. Takacs, I. Rudas, D. Bösl, and T. Haidegger, "Highly automated vehicles and self-driving cars [industry tutorial]," *IEEE Robot. Autom. Mag.*, vol. 25, no. 4, pp. 106–112, Dec. 2018.
[26] M. Daily, S. Medasani, R. Behringer, and M. Trivedi, "Self-driving cars," *Computer*, vol. 50, no. 12, pp. 18–23, 2017.
[27] J. Nieuwenhuijsen, "Diffusion of automated vehicles: A quantitative method to model the diffusion of automated vehicles with system dynamics," M.S. thesis, Dept. Transp. Planning, Delft Univ., Delft, The Netherlands, 2015.
[28] P. E. Ross, "Self-driving tech for non-cars: From mall shuttles to monster trucks, Perrone Robotics is ready to debut its self-driving tech—[News]," *IEEE Spectr.*, vol. 56, no. 3, pp. 9–10, Mar. 2019.
[29] E. Stayton and J. Stilgoe, "It's time to rethink levels of automation for self-driving vehicles [opinion]," *IEEE Technol. Soc. Mag.*, vol. 39, no. 3, pp. 13–19, Sep. 2020.
[30] M. Elbanhawi, M. Simic, and R. Jazar, "In the passenger seat: Investigating ride comfort measures in autonomous cars," *IEEE Intell. Transp. Syst. Mag.*, vol. 7, no. 3, pp. 4–17, Fall 2015.
[31] H. Hartenstein and L. P. Laberteaux, "A tutorial survey on vehicular ad hoc networks," *IEEE Commun. Mag.*, vol. 46, no. 6, pp. 164–171, Jun. 2008.
[32] A. Chowdhury, G. Karmakar, J. Kamruzzaman, A. Jolfaei, and R. Das, "Attacks on self-driving cars and their countermeasures: A survey," *IEEE Access*, vol. 8, pp. 207308–207342, 2020.
[33] Q. Luo, Y. Cao, J. Liu, and A. Benslimane, "Localization and navigation in autonomous driving: Threats and countermeasures," *IEEE Wireless Commun.*, vol. 26, no. 4, pp. 38–45, Aug. 2019.
[34] G. Guidi, A. M. Lekkas, J. E. Stranden, and J. A. Suul, "Dynamic wireless charging of autonomous vehicles: Small-scale demonstration of inductive power transfer as an enabling technology for self-sufficient energy supply," *IEEE Electrific. Mag.*, vol. 8, no. 1, pp. 37–48, Mar. 2020.
[35] Apollo. (2021). *A Whitepaper on Automated Driving Safety*. [Online]. Available: https://apollo.auto/platform/whitepaper.html
[36] J. Gorzelany. (2015). *Volvo Will Accept Liability for its Self-Driving Cars*. [Online]. Available: https://www.forbes.com/sites/jimgorzelany/2015/10/09/volvo-will-accept-liability-for-its-self-driving-cars/#7e0e8e9d72c5
[37] B. Goldfain, P. Drews, C. You, M. Barulic, O. Velev, P. Tsiotras, and J. M. Rehg, "AutoRally: An open platform for aggressive autonomous driving," *IEEE Control Syst.*, vol. 39, no. 1, pp. 26–55, Feb. 2019.
[38] A. Dosovitskiy, G. Ros, F. Codevilla, A. Lopez, and V. Koltun, "CARLA: An open urban driving simulator," in *Proc. 1st Annu. Conf. Robot Learn.*, Nov. 2017, pp. 1–16.
[39] NVIDIA. (2019). *NVIDIA Drive Constellation Now Available—Virtual Proving Ground for Validating Autonomous Vehicles*. [Online]. Available: https://nvidianews.nvidia.com/news/nvidia-drive-constellation-now-available-virtual-proving-ground-for-validating-autonomous-vehicles
[40] (2021). *Simulator for Self-Driving Cars*. [Online]. Available: https://nvidianews.nvidia.com/news/nvidia-drive-constellation-now-available-virtual-proving-ground-for-validating-autonomous-vehicles
[41] F. Rosique, P. J. Navarro, C. Fernández, and A. Padilla, "A systematic review of perception system and simulators for autonomous vehicles research," *Sensors*, vol. 19, no. 3, p. 648, Feb. 2019.
[42] Y. Kang, H. Yin, and C. Berger, "Test your self-driving algorithm: An overview of publicly available driving datasets and virtual testing environments," *IEEE Trans. Intell. Vehicles*, vol. 4, no. 2, pp. 171–185, Jun. 2019.
[43] P. Sisson. (2015). *A Fake City's Real Purpose: Test Tomorrow's Driverless Cars*. [Online]. Available: https://www.curbed.com/2015/7/22/9937998/university-of-michigan-fake-city-selfdriving-cars
[44] A. J. Hawkins. (2018). *Tesla Says Autopilot was Engaged During Fatal Model X Crash*. [Online]. Available: https://www.theverge.com/2018/3/30/17182824/tesla-model-x-crash-autopilot-statement
[45] D. Muoio. (2017). *Uber Built a Fake City in Pittsburgh With Roaming Mannequins to Test its Self-Driving Cars*. [Online]. Available: https://www.businessinsider.com/ubers-fake-city-pittsburgh-self-driving-cars-2017-10?r=US&IR=T
[46] A. Davies. (2017). *Inside the Ersatz City Where Waymo Trains its Self-Driving Cars*. [Online]. Available: https://www.wired.com/story/google-waymo-self-driving-car-castle-testing/
[47] A. Walker. (2019). *Are Self-Driving Cars Safe for Our Cities?* [Online]. Available: https://www.curbed.com/2016/9/21/12991696/driverless-cars-safety-pros-cons
[48] J. Loughran, "Only four vehicles complete driverless car race [news briefing]," *Eng. Technol.*, vol. 12, no. 4, p. 16, May 2017.
[49] S. Liu and J.-L. Gaudiot, "Autonomous vehicles lite self-driving technologies should start small, go slow," *IEEE Spectr.*, vol. 57, no. 3, pp. 36–49, Mar. 2020.
[50] S. Zang, M. Ding, D. Smith, P. Tyler, T. Rakotoarivelo, and M. A. Kaafar, "The impact of adverse weather conditions on autonomous vehicles: How rain, snow, fog, and hail affect the performance of a self-driving car," *IEEE Veh. Technol. Mag.*, vol. 14, no. 2, pp. 103–111, Jun. 2019.
[51] L. Jones, "Are we ready to the steering [transport automotive]," *Eng. Technol.*, vol. 10, no. 9, pp. 32–36, Oct. 2015.
[52] Baidu. (2020). *Building a Self-Driving Car That People Can Trust*. [Online]. Available: https://www.technologyreview.com/2020/12/16/1014672/building-a-self-driving-car-that-people-can-trust/
[53] N. Lalra and S. M. Paddock. (2016). *How Many Miles of Driving Would it Take to Demosnstrate Autonomous Vehicle Reliability?* [Online]. Available: https://www.rand.org/content/dam/rand/pubs/research_reports/RR1400/RR1478/RAND_RR1478.pdf
[54] K. Kuru, "Conceptualisation of human-on-the-loop haptic teleoperation with fully autonomous self-driving vehicles in the urban environment," *IEEE Open J. Intell. Transp. Syst.*, vol. 2, pp. 448–469, 2021.
[55] B. Markwalter, "The path to driverless cars [CTA insights]," *IEEE Consum. Electron. Mag.*, vol. 6, no. 2, pp. 125–126, Apr. 2017.
[56] M. Baker and J. Manweiler, "Sensing, privacy, and things we don't discuss," *IEEE Pervasive Comput.*, vol. 16, no. 3, pp. 7–11, Jul. 2017.
[57] J. Diaz. (2019). *People Don't Trust Autonomous Vehicles, So Jaguar Added Googly Eyes*. [Online]. Available: https://www.fastcompany.com/90231563/people-dont-trust-autonomous-vehicles-so-jaguar-is-adding-googly-eyes
[58] R. N. Brewer and V. Kameswaran, "Understanding the power of control in autonomous vehicles for people with vision impairment," in *Proc. 20th Int. ACM SIGACCESS Conf. Comput. Accessibility (ASSETS)*. New York, NY, USA: Association for Computing Machinery, Oct. 2018, pp. 185–197.
[59] R. Brewer and N. Ellison, "Supporting people with vision impairments in automated vehicles: Challenge and opportunities," USDOT CCAT Project, Washington, DC, USA, Final Rep., 2020.
[60] P. D. S. Fink, J. A. Holz, and N. A. Giudice, "Fully autonomous vehicles for people with visual impairment: Policy, accessibility, and future directions," *ACM Trans. Accessible Comput.*, vol. 14, no. 3, pp. 1–17, Sep. 2021.
[61] J. Fell, "Cars of the future [concept cars]," *Eng. Technol.*, vol. 12, no. 2, pp. 48–53, Mar. 2017.
[62] D. Wu, X. Han, Z. Yang, and R. Wang, "Exploiting transfer learning for emotion recognition under cloud-edge-client collaborations," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 2, pp. 479–490, Feb. 2021.
[63] A. Waytz, J. Heafner, and N. Epley, "The mind in the machine: Anthropomorphism increases trust in an autonomous vehicle," *J. Exp. Social Psychol.*, vol. 52, pp. 113–117, May 2014.
[64] L. Morra, F. Lamberti, F. G. Prattico, S. L. Rosa, and P. Montuschi, "Building trust in autonomous vehicles: Role of virtual reality driving simulators in HMI design," *IEEE Trans. Veh. Technol.*, vol. 68, no. 10, pp. 9438–9450, Oct. 2019.
[65] A. Pillai, "Virtual reality based study to analyse pedestrian attitude towards autonomous vehicles," M.S. thesis, School Comput. Sci. Commun., Stockholm, Sweden, 2017.
[66] S. Deb, M. M. Rahman, L. J. Strawderman, and T. M. Garrison, "Pedestrians' receptivity toward fully automated vehicles: Research review and roadmap for future research," *IEEE Trans. Human-Mach. Syst.*, vol. 48, no. 3, pp. 279–290, Jun. 2018.
[67] S. Gupta, M. Vasardani, and S. Winter, "Negotiation between vehicles and pedestrians for the right of way at intersections," *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 3, pp. 888–899, Mar. 2019.
[68] K. Roorda, "The effects of driving behavior consistency in a car on pedestrian crossing decisions," M.S. thesis, Dept. Ind. Eng. Innov. Sci., Eindhoven Univ. Technol., Eindhoven, Netherlands, 2018.
[69] B. Brown, "The social life of autonomous cars," *Computer*, vol. 50, no. 2, pp. 92–96, Feb. 2017.

[70] N. A. Greenblatt, ''Self-driving cars and the law,'' *IEEE Spectr.*, vol. 53, no. 2, pp. 46–51, Feb. 2016.
[71] T. Zhang, ''Toward automated vehicle teleoperation: Vision, opportunities, and challenges,'' *IEEE Internet Things J.*, vol. 7, no. 12, pp. 11347–11354, Dec. 2020.
[72] E. Ackerman, ''Drive.ai launches robot car pilot in Texas with a focus on humans,'' *IEEE Spectr.*, May 2018. [Online]. Available: https://spectrum.ieee.org/driveai-launches-dallas-robot-car-pilot-with-a-focus-on-humans
[73] S. Higginbotham, ''Autonomous trucks need people [opinion],'' *IEEE Spectr.*, vol. 56, no. 3, p. 21, Mar. 2019.
[74] A. Jobin, M. Ienca, and E. Vayena, ''The global landscape of AI ethics guidelines,'' *Nature Mach. Intell.*, vol. 1, no. 9, pp. 389–399, Sep. 2019.
[75] S. Applin, ''Autonomous vehicle ethics: Stock or custom?'' *IEEE Consum. Electron. Mag.*, vol. 6, no. 3, pp. 108–110, Jul. 2017.
[76] N. J. Goodall, ''Can you program ethics into a self-driving car?'' *IEEE Spectr.*, vol. 53, no. 6, pp. 28–58, Jun. 2016.
[77] *IEEE Draft Model Process for Addressing Ethical Concerns During System Design*, Standard IEEE P7000, 2016.
[78] S. Karnouskos, ''Self-driving car acceptance and the role of ethics,'' *IEEE Trans. Eng. Manag.*, vol. 67, no. 2, pp. 252–265, May 2020.
[79] J.-F. Bonnefon, D. Cerny, J. Danaher, N. Devillier, V. Johansson, T. Kovacikova, M. Martens, M. Mladenovic, P. Palade, N. Reed, F. S. de Sio, S. Tsinorema, S. Wachter, and K. Zawieska, ''Ethics of connected and automated vehicles: Recommendations on road safety, privacy, fairness, explainability and responsibility,'' Horizon Commission Expert Group, Luxembourg, Tech. Rep., Sep. 2020.
[80] M. Bertozzi, A. Broggi, A. Coati, and R. I. Fedriga, ''A 13,000 km intercontinental trip with driverless vehicles: The VIAC experiment,'' *IEEE Intell. Transp. Syst. Mag.*, vol. 5, no. 1, pp. 28–41, Spring 2013.
[81] P. Liin. (2013). *The Ethics of Autonomous Cars*. [Online]. Available: https://www.theatlantic.com/technology/archive/2013/10/the-ethics-of-autonomous-cars/280360/
[82] S. A. Applin, A. Riener, and M. D. Fischer, ''Extending driver-vehicle interface research into the mobile device commons : Transitioning to (nondriving) passengers and their vehicles,'' *IEEE Consum. Electron. Mag.*, vol. 4, no. 4, pp. 101–106, Oct. 2015.
[83] A. M. Nascimento, L. F. Vismari, A. C. M. Queiroz, P. S. Cugnasca, J. B. Camargo, and J. R. de Almeida, ''The moral machine: Is it moral?'' in *Computer Safety, Reliability, and Security*, A. Romanovsky, E. Troubitsyna, I. Gashi, E. Schoitsch, and F. Bitsch, Eds. Cham, Switzerland: Springer, 2019, pp. 405–410.
[84] J. Borenstein, J. Herkert, and K. Miller, ''Self-driving cars: Ethical responsibilities of design engineers,'' *IEEE Technol. Soc. Mag.*, vol. 36, no. 2, pp. 67–75, Jun. 2017.
[85] D. J. Glancy. (2012). *Privacy in Autonomous Vehicles*. [Online]. Available: https://digitalcommons.law.scu.edu/lawreview/vol52/iss4/3/
[86] G. R. Lucas, *Ethics and UAVs*. Dordrecht, The Netherlands: Springer, 2015, pp. 2865–2878.
[87] S. Kuutti, S. Fallah, R. Bowden, P. Barber, and A. Khajepour, *Deep Learning for Autonomous Vehicle Control: Algorithms, State-of-the-Art, and Future Prospects*. Williston, ND, USA: Morgan & Claypool, 2019.
[88] L. M. Hulse, H. Xie, and E. R. Galea, ''Perceptions of autonomous vehicles: Relationships with road users, risk, gender and age,'' *Saf. Sci.*, vol. 102, pp. 1–13, Feb. 2018.
[89] T. Fournier, ''Will my next car be a libertarian or a utilitarian? Who will decide?'' *IEEE Technol. Soc. Mag.*, vol. 35, no. 2, pp. 40–45, Jun. 2016.
[90] S. Liu, J. Peng, and J.-L. Gaudiot, ''Computer, drive my car!'' *Computer*, vol. 50, no. 1, p. 8, Jan. 2017.
[91] M. Anderson, ''The road ahead for self-driving cars: The AV industry has had to reset expectations, as it shifts its focus to level 4 autonomy—[News],'' *IEEE Spectr.*, vol. 57, no. 5, pp. 8–9, May 2020.
[92] M. Botvinick, S. Ritter, J. X. Wang, Z. Kurth-Nelson, C. Blundell, and D. Hassabis, ''Reinforcement learning, fast and slow,'' *Trends Cogn. Sci.*, vol. 23, no. 5, pp. 408–422, May 2019.
[93] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski, S. Petersen, C. Beattie, A. Sadik, I. Antonoglou, H. King, D. Kumaran, D. Wierstra, S. Legg, and D. Hassabis, ''Human-level control through deep reinforcement learning,'' *Nature*, vol. 518, no. 7540, pp. 529–533, 2015.
[94] D. Silver, A. Huang, C. J. Maddison, A. Guez, L. Sifre, G. van den Driessche, J. Schrittwieser, I. Antonoglou, V. Panneershelvam, M. Lanctot, S. Dieleman, D. Grewe, J. Nham, N. Kalchbrenner, I. Sutskever, T. Lillicrap, M. Leach, K. Kavukcuoglu, T. Graepel, and D. Hassabis, ''Mastering the game of go with deep neural networks and tree search,'' *Nature*, vol. 529, no. 7587, pp. 484–489, Jan. 2016.
[95] C. Beattie *et al.*, ''DeepMind lab,'' 2016, *arXiv:1612.03801v2*.
[96] T. Tan, F. Bao, Y. Deng, A. Jin, Q. Dai, and J. Wang, ''Cooperative deep reinforcement learning for large-scale traffic grid signal control,'' *IEEE Trans. Cybern.*, vol. 50, no. 6, pp. 2687–2700, Jun. 2020.
[97] Z. Wan, C. Jiang, M. Fahad, Z. Ni, Y. Guo, and H. He, ''Robot-assisted pedestrian regulation based on deep reinforcement learning,'' *IEEE Trans. Cybern.*, vol. 50, no. 4, pp. 1669–1682, Apr. 2020.
[98] B. Luo, D. Liu, and H. Wu, ''Adaptive constrained optimal control design for data-based nonlinear discrete-time systems with critic-only structure,'' *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 6, pp. 2099–2111, Jun. 2018.
[99] H. Ravichandar, A. S. Polydoros, S. Chernova, and A. Billard, ''Recent advances in robot learning from demonstration,'' *Annu. Rev. Control, Robot., Auton. Syst.*, vol. 3, no. 1, pp. 297–330, May 2020.
[100] Y. Sung and K. Cho, ''Collaborative programming by demonstration in a virtual environment,'' *IEEE Intell. Syst.*, vol. 27, no. 2, pp. 14–17, Mar./Apr. 2012.

KAYA KURU received the B.Sc. degree from the National Defence University, the major/A.D.P. degree in computer engineering from Middle East Technical University (METU), the M.B.A. degree from Selçuk University, and the M.Sc. and Ph.D. degrees in computer science from METU. He has completed the Postdoctoral Studies with the School of Electronics and Computer Science, University of Southampton, U.K. He worked as a DBA, a SW Developer, a SW Development Manager, and an IT Manager with the IT Department at a university, for a period of 15 years. He is currently an Associate Professor in computer-information systems engineering. He has recently engaged with the implementation of numerous AI-based real-world systems within various funded projects. His research interest includes the development of autonomous intelligent systems, using FL, DL, and DRL with CPSs.