Abstract - Drawing perspectives from science and technology studies, philosophy of science, and literature from ethics and social justice, this paper examines the promises and challenges in the development of self-driving vehicle (SDV) technology. We start with the premise that the combination of different computing technologies embedded in SDVs is a powerful tool for efficiency in communications, information gathering, processing, and storage. However, by focusing on efficiency, SDVs provide a new mode of industrialized transportation whose users can only choose between transportation services, but have little or no say about the broader social implications of the technology. We argue that perspectives from social justice and ethics show that SDVs have implications beyond transportation, with profound consequences for users and societies. In particular, values such as privacy, security, and responsibility may be changed for good or bad, in both the short and long-term. The examination of these changes, while the technology is still under foundational development, is as urgent as it is needed.

Keywords - Automated vehicles; self-driving vehicles; SDV; transportation; automobility; engineering ethics; social justice.

I. INTRODUCTION

Transportation of people and goods underpins our modern industrialized society. As Ruth Schwartz Cowan reminds us, “You cannot consume frozen TV dinners or acrylic knit sweaters or aspirin or a pediatrician’s services unless you can get to them, or unless someone is willing to deliver them to your door.” [1]. With over one billion vehicles in the world today [2], the motorized road vehicle is the pillar of modern transportation. Since the beginning of the 20th century, the automobile has revolutionized our spaces, practices, cultures, and identities through the complex matrix of technical, financial, economic, political, and social linkages [3][4] shaping its design, manufacture, production, and use.

Now, motorized road transport is set to revolutionize our society once again on a huge scale. Recent technological developments in propulsion, telecommunications, sensing, and in-vehicle computing technology are expanding the range of vehicles’ capabilities. A technological convergence is underway, moving towards self-driving vehicle (SDV) technology—vehicles that are intended to utilize computational algorithms, sensors, and communication devices to automatically navigate a variety of environments without human drivers. SDVs also promise increased safety, speed and convenience, as well as reduced energy consumption.

With their potential benefits, however, SDV technologies bring a new set of challenges at the vehicle and system levels. These include impacts on urban infrastructure, limitations on a vehicle’s driving range, equipment reliability, data privacy assurances, as well as concerns over infrastructure investment, and ethical implications for safety [5]. We argue that SDVs introduce new challenges in social justice, affecting the distribution of benefits and burdens in society. By privileging a positive and unproblematic future with SDV technology, current design practices and institutions revolving around these new vehicles have systematically ignored important ethical and social challenges.

II. OBJECTIVES

To address these challenges, this paper outlines a framework for a more ethical design practice for SDV systems. We aim to steer current engineering practices on SDV technologies away from a narrow, technical perspective. Ultimately, we hope to broaden the conversation about the design processes in general for technologies that seek to revolutionize society on a large scale. This approach is especially relevant to the new field of responsible innovation [6].

Our starting premise is that technology is neither benign nor neutral [7], carrying with it a raft of implications for social justice [8]. We have been aware for decades through the philosophy of technology [7][9] and Science and Technology Studies (STS) [10][11][12][13][8][14][15] that technology is a profoundly social phenomenon, with both intended and unintended consequences for our everyday lives. Theories of social justice provide a framework for assessing the distribution of advantages and disadvantages in a society, through a set of rules that distinguish between just and unjust actions or institutions [16][17]. With social justice as the inspiration of the framework, the technology designer can ask: what burdens and benefits are to be distributed by the new SDV technology, to whom and by whom, and using what means? This line of inquiry further leads to questioning how
institutional, economic, and political powers determine the benefits and burdens of a technology, and to identifying the agents of power involved, today and in the future.

While design processes have begun to move beyond the deterministic perspective of past theorizing by addressing individual driver concerns such as fashion and infrastructure [18][19], wider examination of the societal concerns for the distribution of benefits and burdens has thus far been rather limited. Earlier inquiries have focused on issues of safety and the decision-making of the vehicle or human. Ethical investigations usually point to fully automated SDVs and their version of the “trolley problem” [20] 1. The Trolley Problem is described as, when given the choice between killing an individual and killing a group, the most logical choice would be to kill the one to save the many (maximizing utility for the most people - the simple utilitarian perspective). In the case of SDVs, automated vehicles will programatically choose the most utilitarian course of action, thereby reducing safety to a matter of deliberately sacrificing a few to protect the whole [21].

Making decisions algorithmically is much more difficult than has yet been discussed. What if there are two elderly pedestrians with a young child in a crosswalk? Might it not be better to kill the two old people than the child? This is partly at least within the capacity of an SDV which could judge the relative size of the people. What if the old people are shorter than the child? And should a machine be programmed to kill civilians?

Focusing solely on a utilitarian perspective runs the risk of reductionism, neglecting the complexity of sociotechnical interactions and their role in social justice.

In other words, blindly applying the “trolley problem” to SDVs reduces the debate to a choice between specific, narrowly conceived courses of action without discussing the wider context that gives rise to the choice and its implications. For example, an SDV may route from A to B by choosing from multiple possible paths, prioritizing a path based on a factor like time versus a passenger’s health. Adopting this simple ethical framework neglects consideration of the SDV as a complex artifact with a range of implications including privacy, freedom, security, and governance.

Our framework will attempt to address these limitations.

III. METHODOLOGY

Our methodological approach is inspired by the questions of social justice raised by automobility [17], and presents an approach to SDVs that goes beyond techno-centric, utilitarian perspectives [22][23]. Aside from automobility, we draw from social theories including technological transitions, ethics, STS, sociotechnical studies such as those pioneered by Rob Kling [24], foresight theory, and the tradition of participatory design, to examine the range of possible futures that SDVs might bring in terms of transport and society at large. This approach examines sociotechnical systems as composed of people and technology, including, for example, hardware, software, physical surroundings, different stakeholders, procedures, laws and regulations, data, and structures. The approach considers, as the subject of social practices, everyday routines and interactions that require the use of a vehicle for the enactment of sociality, such as household socialities and work socialities. Consequently, mobility provides a support structure to the social organization and activities of humans in households, workplaces, and so on, and is integral to the management of those places.

As engineers, we draw heavily on the intellectual tradition of “bearing witness to technology” including the work of Ilich [13], Mumford [25], Ellul [14], and Rochlin [15]. We also refer to the contemporary engineering literature on professional engineering ethics [26]. Our perspective is that the process by which a technological civilization like ours is developed is not about “evil statesmen” but a response to “laws of development” [14]. These laws are defined by what Ellul termed “La Technique”, namely “the totality of methods rationally arrived at and having absolute efficiency (for a given stage of development) in every field of human activity” [14]. Effectively, the laws of development are based on efficiency to produce a growing concentration of capital. In addition to the writings of Ilich and Mumford about the impact of the car on society, in designing our framework we are also inspired by Gene Rochlin’s theories on computers in society [15] to unpack the future impact of the growing implementation of new, computerized, SDV technologies. Our framework critically sees SDVs as disruptive because of the powerful computing technologies they combine, and their resulting power to collect voluminous data.

Based on this holistic understanding of technology in a social context, we propose methodological holism to help create better-informed design practices. This methodology will incorporate elements of action-orientation, openness to alternative futures, and participatory design [27][28][29]. Such a framework must envisage social effects from changes in urbanization to changes in computing, driving, and infrastructure, and therefore social evolution.

IV. TECHNOLOGY DESCRIPTION

A. What is an SDV?

In building a framework for improved SDV technology design, perhaps the first thing to ask is “What is an SDV?” We define SDVs in terms of a combination of technologies on the one hand and a level of automation on the other. SDVs are part of a new era of vehicle systems where part or all of the driver’s actions may be removed or limited, and where cars involve a combination of new technologies including sensors, computing power, and short-range communications (SCC for short), effectively creating a new human-automobile hybrid.

SDV technology, whether automated or not, is powerful. It comprises a range of possible technologies, including in-vehicle sensing technology that enables real-time gathering of data about the vehicle and its environment. Such data may include geographical coordinates and a car’s speed, direction, acceleration, and obstacles encountered [30]. Coupled with increased processing power and storage capacity, vehicles can store larger amounts of data than a typical personal computer [31][32]. Communications technology enables the SDV to detect and communicate with other vehicles and infrastructure by means of short-range vehicle-to-vehicle and vehicle-to-

---

1 The “trolley problem” is so named after its original formulation in terms of a train signal controller, who faces a choice of two tracks to route a train (a “trolley” is a kind of train carriage) onto—one where he will kill one person and another where he will kill a group of people—will choose to just kill the one man.
infrastructure communications technology by enabling the transfer of periodic messages to inform the surrounding vehicles and infrastructure, relaying data on the speed, position, and direction of the vehicle [32].

The result of this technological shift is the potential for individual vehicles to navigate in their environment and perform driving actions without the need for a driver [33][34][35]. The level of automation—namely the second part of the definition of the SDV—is, however, independent of technology involved and only refers to the level of human involvement in the driving process. Industry defines five of these levels, from the most intense driver intervention (level 1) to the least (level 5).

Various combinations of SCC technologies do not translate readily into particular automation levels. Instead, the particular medley of SCC technologies found in each brand of SDVs is determined by the SDV vendors, and varies from SDV vendor to SDV vendor according to a number of business, economic, and cultural variables. The Google SDV, for example, combines sensing and computing technologies without employing communication technologies, enabling Google to sell maps that vehicles need. However, the Google SDV aims for the highest level of automation possible, even in the absence of communication. Other vendors, on the other hand, include communications technologies to reach full automation. For instance, the communications technology component will enable the formation of cooperative vehicle systems where there is no need for a human intervention in driving.

While vehicular automation has been around for a while (autopilot systems, automated trains, automated park rides), SDVs are revolutionary and will infiltrate society in a more pervasive, broader, and potentially disruptive sense than other automated vehicle systems. However, even with a full spectrum of technologies, if a city is not technologically-outfitted for SDVs, SDVs will not be able to drive properly.

B. SDVs and Social Justice

This section examines how the nature of SDV technology may determine social justice.

Just as Mumford [8] named the mechanical clock as the most important invention to shape the industrial revolution, describing it as “power-machinery” par excellence, we might consider data and automation as powerful a piece of power-machinery as the clock. SDV technology represents an unprecedented deeper level of embeddedness of computer technology in vehicles, whose “new degrees of interconnectivity” [15] will create new complexities in social and political environments [15]. Effectively, SDVs are a new cyberspace, “a universe defined and bounded by interactive electronics” [15].

In practice, through their data management power, SDVs constitute an industrial commodity that will allow greater standardization of times of travel, improved data collection, and powerful management of fleets of cars. As such, they can, under the “laws of development” described by Ellul, allow monopolies to more effectively demarcate target neighborhoods, and create different classes of highways and roads over more or less traveled paths. SDVs may impose on the traveller a certain scarcity of choice, time, and cost, making the traveller merely a client, not a citizen or a friend. Illich [13] refers to this client as a person who can only ask for better service in the form of more complex technological solutions, but not ask for social justice. For an similar example of this possibility, taking a look at the city of Oslo in Norway and another form of personal transport—electric cars—is revealing. Differences in uptake in EVs between comparatively wealthier, upper class neighbourhoods and comparatively less wealthy, working class neighbourhoods, for instance, can in part be explained by material geographies established in the time of Danish king Christian IV in 1624 [36].

SDVs could be instrumental in creating a “speed hierarchy” in which the extra speed gained from automation, including closer driving, shorter and less crowded travel distances, will command a premium. This may price out certain members of society in gaining access to SDVs, and to other means of transport in general, if few alternatives become available. Others may find themselves paying more for SDVs so as to create a balance between their leisure time and work time through transportation. We note that such a speed hierarchy is present in California with electric/gas hybrid cars being granted unlimited access to the faster carpool lanes—colloquially known as “Lexus Lanes”—by means of a sticker based on emissions [37][38][39].

In the case of SDVs used to transport goods, the power relationships involved in the industry will determine whether SDVs will impact jobs involving driving (for example long-haul truckers). The ability of SDVs to travel further in the same amount of time could benefit retailers by giving them a greater range of potential suppliers, driving down prices paid to farmers, and increasing profits for retailers.

This increased speed may result in greater distances travelled, as the social share of time society dedicates to travel does not change. And, since increased speed will require more expensive and sophisticated technologies, the cost of transportation will rise for the typical user, eating up a larger share of their income, and therefore work time, for the cost of transportation.

The framework must then consider the power relations that may give rise to such a speed hierarchy. The hierarchy can occur over geographical locations and across groups. Illich argued that radical monopoly is first established by a rearrangement of society for the benefit of those with the most money and power, typically pricing the most vulnerable out of transportation by means of expensive technology that they have no alternative but to use. Indeed, past transitions teach us, in the case of the automobile, that the rise of a wealthy, educated middle class of planners, judges, and entrepreneurs, had a huge impact on defining the automobile in the US, including deciding who could drive an automobile, and where, and who would ride public transport [4]. In the US, 300 million USD was spent in 1906 to build roads, redefining social geographies, in and outside of cities [4].

With this history in mind, we could ask whether SDVs will have a disproportionate impact in different countries. We could ask whether SDVs in Palo Alto will have the same impact as SDVs in a setting with a different economy. Careful consideration must be given to the impact of displacing road users—especially those using cheaper, less capital-intensive transport such as bicycles and mopeds—as well as the tax revenue that could be used on other sectors such as health or education. Careful consideration must be given to displacing jobs, including delivery jobs but also farming and production jobs that go hand in hand with transport in economies in which
those jobs are important. Indeed, there are arguments that SDV technology will make freight cheaper, and enable greater distances to be covered from increased speed and volume in a
set time. We know from studies of the transition to the automobile in the US that local residents objected to the automobile because they saw it as making the street a transport
artery rather than a meeting place. As a result, local trade and socialities lost out [4]. Traffic flow was prioritized over the safety of pedestrians, resulting in the reeducation of pedestrians
and beginning the trend of a shockingly high rate of pedestrian deaths.

As computers become more embedded in our socio-
technical systems, we become more reliant on them, and this
forces us to accept, or at least live with, the societal transformations that follow from them. Careful thought should be
given to a technological future involving automation as it
involves far-reaching consequences.

C. SDVs and Ethics

The theory of technique has deeper, ethical implications for
social justice and SDVs that go beyond transportation as a
system. The World Health Organization’s Global Status Report
on Road Safety 2013 estimates that there are 1.24 million
traffic deaths a year; traffic is one of the top ten causes of death
worldwide (nearly triple the UN’s estimate of annual murders,
and twenty times the estimated annual total of deaths in wars
[40], [41]). Yet while we have sophisticated theory to examine
the ethics of war, we have no such framework for casualties of
the road.

While there are conventions in war directed at defining who
is a combatant and who is not, and who may be killed [42],
traffic has no such conventions. Traffic deaths are not
distributed equally across physical and social geographies—a
disproportionate number of victims are vulnerable road users
such as cyclists (as the Critical Mass movement demonstrates)
and pedestrians. Given a fatal crash between two like vehicles,
a woman driver is more likely to die than a man. [43]

Perhaps herein lies the real Trolley Problem: having
become accustomed to traffic deaths, perhaps we are willing to
accept them as an unfortunate and unavoidable cost of the
freedoms the car gives us. Perhaps we should be asking
questions about the current Trolley Problem approach to safety
ethics: programming automated vehicles to choose the most
utilitarian course of action, and thereby reducing safety to a
matter of deliberately sacrificing a few to protect the whole
[21].

There is an even darker side to the social construction of
SDV technology with bigger implications. Safety technology
choices are often determined by financial flows and lobbying
practices [43][44]. Evans gives an example of the early airbags
on US cars being used despite scientific evidence that their
design increased deaths in accidents [43].

In the mass media, car advertising glamorizes speed and
cohabits with alcohol advertising—two deadly risk factors
mutually reinforcing each other for the sake of selling products.
Evans argues that these strong commercial links between the
automobile, alcohol industries, and mass media, through
advertising revenue, influence safety but also responsibility
[43]. For instance, beer consumption is implicated in most
drunk-driving accidents, and is at the same time the most
advertised form of alcohol on TV [43].

Considering the flows of money involved, one can understand why automation, especially fully- or highly-
automated SDVs, raises issues of who holds power and who
holds responsibility in, for instance, accidents. A broader ethics
perspective raises the question of individual and collective
responsibility—the need for “responsibility by design, based on
distribution of responsibilities at all levels” [45].

The implications for social justice stemming from SDVs do
not stop at SDV implementation and safety considerations.
While SDVs are a radical technology, they are also a
combination of existing technologies that have converged into
a new one, generating a new technological context, which, in
turn, is affecting our conception and values regarding security.
There is evidence that new technological contexts created by
the increasing convergence between technological fields is
changing the very nature of security values now [46], and one
could therefore envisage such an impact in the future. Many
security and computing technologies, including some SDV
technologies, are technologies used in military applications,
such as military autonomous vehicles [26]. Therefore, the
framework must consider the impact on society and security of
these technologies. There is evidence in other fields that it is
the changing security and changing science landscapes (and
their convergence) that are shaping the regulation of lethal
technology, with self-regulation in the form of non-lethal
technology leading to a grey area [47]. In other words, a
society capable of designing and building technology-
advanced SDVs may also be a militarist one with the capability
to build lethal automated weapons whilst achieving economies
of scale and improving its balance of payments—an economic
justification in other technology fields.

As a consequence, the nature and scale of the technology
behind SDVs may redefine the norms for some of our cultural
values that are linked to security such as privacy and
responsibility. Privacy involves power relationships (who is
allowed privacy and how much?), questions about legitimacy
(who has access to data and can be trusted with it under what
conditions?), ownership (does the SDV client have as many or
fewer rights to data about him or her than a citizen?) and
empowerment (how easily can the client act as a citizen?) [48].
These concerns have implications for civic society, harmony,
and mutual respect [49]. Furthermore, because privacy is also a
cultural practice that varies between countries [49], we can
imagine different futures in different cultures and countries.

Thought must be given to the sustainability of SDVs,
especially as unsustainable and unjust transport policies are
often articulated and framed in terms of individual consumer
choice rather than sustainable alternatives [50]. Combustion-
engined SDVs driving around in cities between rides, thereby
eliminating the need for parking, may result in increased CO2
emissions and increased pollution, as the driving between rides
may exceed the emissions saved by not looking for parking and
not being parked. Secondly, the increase in speed that SDVs
might offer people and freight may mean an increase in travel
distance and frequency, resulting in higher CO2 emissions.
Nothing is known about how travel practices might change and
how this might affect emissions. Self-driving technology is
drive-train independent, and while battery-electric vehicles use
30% of the energy of an ICE vehicle, their uptake, along with
the sustainability of the energy system which produces
electricity, depends on broader economic, institutional and
cultural factors.
D. A Framework for Expanding the Technological Design Horizon

Considering the range of issues that SDVs might introduce, it becomes apparent that if we ask the ethical questions on the micro scale of one situation, there is a danger of neglecting the complexity of sociotechnical systems, and consequently neglecting interdependencies between social justice considerations.

This approach requires an evolution of current design practices. Instead of focusing on SDVs as primarily technical objects, and designing by a focused observation of the object of the design (SDV), an improved approach would require engineers and others to look into the future they are creating through the object of their design. This view into the future should involve imagining a range of possible and desired futures while considering a range of values.

Here, an additional layer of complexity arises when designers are faced with accommodating conflicting values. An example would be developing a traffic control mechanism for SDVs that balances questions of safety, privacy, mobility, and environmental sustainability [23]. In sum, a framework for the ethical design of SDVs must consider a range of implications of automation on technology and the future. These include: greater degrees of embeddedness and interconnectivity that generate social and political complexities, privacy and responsibility, changes of cultural values, along with the subtle ethical decisions in driving. However, the first step in dealing with this complexity in technology design would be accepting that complexity exists. By looking away or shrugging at this complexity, engineers and designers avoid the responsibility for the distribution of benefits and burdens that SDV technology will have on current and future generations. Instead, applying an expanded design horizon will require them to avoid dissecting technical and social questions, just as, to put it in plain words, one cannot produce a magnificent color by using only green, blue, or red.

V. CONCLUSIONS AND SUMMARY RECOMMENDATIONS

We find that conventional design practice for the development of vehicular technology focuses primarily on technical challenges and can be considered unethical. We have argued that better understanding of technology in a broader social context, along with the distribution of burdens and benefits in a current and future society, enables the development of improved design practice, i.e., an expanded design horizon. The best way to express a range of ethical concerns in a framework is by means of a foresight-based, social construction of technology methodology, which should include at least:

- An account of the changing role of the human inside and outside the vehicle, and consequent impacts on sociotechnical structures and practices; for example, how the SDV radically reconfigures the experience and practices of driving, mobility planning, household and work socialities.
- An account of the roles, objectives and design approaches of the agents of technological development, including corporations, government, and researchers; affecting the developmental trajectories of SDV technology.
- The development of a range of possible and desired futures, and assessment of their system-level impacts, including environmental and social justice perspectives on the distribution of burdens and benefits across society, now and in the future.
- Engagement of the general public in the design process; including using institutional mechanisms for evaluating a wide range of general and localized values to consider in the development of SDV technology.

An important point to emphasize from this framework is that the critical concerns related to the improvement of design practices must be addressed now, while the technology is still under foundational development. A complexity- and foresight-based methodology enables a perspective in which the future is something that can be shaped, rather than being already decided or “inevitable”.

Clearly, in the context of culture, ethics, and the knowledge society, SDVs are a powerful technology that will have profound implications for social justice, in the transport system and beyond, sometimes in obvious ways, other times more subtly. A paradigm shift is needed to bring ethical considerations within the everyday activities of design practice.

REFERENCES


