DIGITAL ACCESS TO SCHOLARSHIP AT HARVARD
DASH.HARVARD.EDU

HARVARD LIBRARY

# Mind, Machine, and Society: Legal and Ethical Implications of Self-Driving Cars

## Link
http://nrs.harvard.edu/urn-3:HUL.InstRepos:38811556

## Terms of use
This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material (LAA), as set forth at
https://harvardwiki.atlassian.net/wiki/external/NGY5NDE4ZjgzNTc5NDQzMGIzZWZhMGFlOWI2M2EwYTg

## Accessibility
https://accessibility.huit.harvard.edu/digital-accessibility-policy

## Share Your Story
The Harvard community has made this article openly available.
Please share how this access benefits you. Submit a story

# Contents

# Introduction

The self-driving car embodies, in many ways, what people imagine when they consider what society will look like in the future. The automation of integral aspects of our daily lives through technology has rapidly increased since the advent of the computer in the mid-1900's, but unlike the original preoccupation with applying the computer as a tool for people to augment their informational activities – communication, computation, data storage, etc. – computers are now substituting for humans in increasingly complex functions, by emulating the human mind within operating mechanical systems (Brynjolfsson 2014). The question of how or whether the machine actually makes autonomous decisions is still largely contested: if the machine is programmed by humans, can it ever really make decisions on its own; can machine learning ever break out of human biases when trained on human-generated data? It is evident nonetheless that machines are replacing increasing numbers of functions that have historically required human decision-making in real time, as is notably the case in the area of driving.

The development of artificial intelligence in such an integrated and high-stakes application as the autonomous vehicle creates a pressing need to inspect the risks and consequences of this technology in society, and that is what this thesis sets out to do. This thesis also comes at an important moment in the self-driving car's development process, just two weeks after what is considered the first pedestrian death involving an autonomous vehicle in the United States (Garcia 2018). In that crash, a self-driving Uber car fatally hit a pedestrian crossing the road in Tempe, Arizona. The car was fully equipped with technology that was designed to stop in time to avoid hitting the pedestrian, and there was a driver at the wheel to take control in case of emergency. Camera footage shows that the driver was not paying attention to the road at the time of the crash (Roberts 2018). While government agencies including NHTSA have opened investigations to identify the cause of the crash, the accident pushed to the center of public attention the larger question of how regulators and litigators ought to effectively allocate responsibility and incentives for the development of safe self-driving cars. For example, the accident called attention to previously unreported ties between the Arizona governor and Uber that led to a testing program with relatively little expert oversight (Harris 2018).

In this thesis, I use an analysis informed by both engineering and law to address the systematic risks and consequences that may arise from putting the human mind into the machine to make life-critical decisions in driving. I will turn to legal precedent and societal dynamics to identify who is and ought to be responsible for accidents involving self-driving cars.

Liability for car accidents today can be examined in relation to three elements in the driving ecosystem: the car, the driver, and the physical context. Unpredictable contingencies within each of these elements, as well as those that arise in the interface between elements, will be treated as focal points of potential regulation and as sources of legal liability for accidents that do occur.

There has been considerable precedent in products liability law for how to categorize different types of failures and who to hold responsible. However, these precedents may have to be reconsidered in relation to the considerable software component of the self-driving car. Defaults may arise at three levels of complexity in the car: the mechatronic, component-specific level; the autonomous mechatronic system which acts as the software control system for the physical components; and the cross-linked mechatronic system which connects a given car to other cars and to a smart grid. Failure at any of these three levels can warrant different types and levels of responsibility.

The driver's responsibility also changes in the self-driving car. In the traditional car, a driver was historically subject to well-defined rules specifying safe conduct on the road. Shifting the weight of control and decision-making to the machine presents new challenges for deciding who ought to be responsible. Distracted driving offers an example. Distracted driving has served as a label for dangerous behaviors such as texting while driving or driving under the influence of alcohol, and regulators outlawed these behaviors accordingly. However, self-driving cars provide new incentives for distracted driving which may be more challenging to regulate. When most driving functions are automated, the need to pay attention to the road is vastly diminished, indeed most of the time made unnecessary, and the nominal driver may be a poor safety net to step in and override the car's functions in those critical moments when the car's response functions prove inadequate. Moreover, even a small incentive for distraction can encourage far riskier behaviors, as was the case of a fatal semi-autonomous car accident on May 7, 2016.

Joshua Brown had been driving his Tesla Model S car on autopilot at the speed limit when a lorry cut in front of it. The bright sunlight reflecting off the white trailer did not register on the Tesla's computer vision technology fast enough for it to brake and avoid a crash. Although it is unclear what Joshua Brown had been doing in the moments immediately leading up to the accident, he had previously posted several YouTube videos of himself driving hands-free on the highways though he received numerous automated warnings not to do so. Also, the automated functions of the car could have been overridden, but Brown did not react to the trailer or attempt to brake before the crash (Kurtz 2016). This is just one example of the broader behavioral phenomenon that people pay less attention, not more, to tasks that involve automated processes. The case calls for a nuanced analysis of how to regulate for the risks arising from altered driver behavior in self-driving cars.

Additional risks arise from the interaction of the self-driving car with the drivers of other cars. Human drivers take into account a myriad of factors not explicitly addressed by law and not accounted for by current technology, including eye contact at intersections, social rules of courtesy, and perceived behavior of the other driver based on the age, brand and even color of a car. When a driver is sitting at the wheel of self-driving vehicles, however, it may not be apparent to other drivers on the road whether the car is under the control of a human or a machine, and whether or not expected social norms of driving apply. This suggests that the strict application of current driving rules and right of way principles may not be adequate for regulating a driving ecosystem with self-driving cars.

The current legal system governing driving has been built and modified on the basis of human behavior and values, and not a mechanical system of rules and laws that are executed exactly as enacted. A good example is speed limits. Human drivers and even law enforcement officers do not assume that the posted speed limit is the actual limit that will be enforced; in fact, driving below the speed limit may in many cases be more dangerous than driving a small number of miles per hour above. On the other hand, if it is raining heavily or the roads are icy, the safe (and enforceable) speed limit may come down below the posted limit. Self-driving cars will have to account for the variation in driving speeds based on such external factors, as well as account for sudden acceleration or deceleration from nearby cars that is common during driving. When determining the allocation of legal responsibility for accidents that arise from social contexts, current law will not only have to be reexamined and potentially modified, but regulations will have to be tailored to the behavioral impacts on the driver and the car.

The third element of the self-driving car, the physical context, is perhaps the most challenging one to analyze, because of the wide variety of contingencies it presents. Anomalies on the road that require rapid response are frequent and inexhaustible, including breaches of law by other car drivers and pedestrians – jaywalking, speeding, changing lanes without a signal – and unexpected phenomena such as animals crossing, potholes, weather conditions, damaged roads, or flying debris. Self-driving cars will have to react to all of these safely, and responsibility for safely navigating these unpredictable scenarios will need to be allocated in the imperfect stages of the development process, or even after full development and commercial deployment. In the case of the Tesla Model S accident, the bright sunlight reflecting off of the white trailer was a contingency not accounted for by the technology, and yet responsibility was not placed on the manufacturers of the sensors. Tesla had, after all, warned the driver several times about the need to pay attention to the road, and the white trailer's driver had shown reckless driving by cutting so close. However, the question remains: which responses are programmed in for the car to react to, and which are left out? And what degree of situational variety should the people who are making these decisions be held responsible for?

In addition, the broader body of driving rules and its enforcement system are also products of many competing interests. Responsibility for the car's development can be analyzed at three levels. First, the designers of the car have the first pass at building high-quality automobile parts and sensors, coding reliable and fast-responding control systems, and thinking of what scenarios the car should be responding to, and how. Designers have a responsibility to be transparent about where things might go wrong with the parts they are designing, and they are also responsible for making ethical decisions about what to program into the car and what to leave out, which may have life or death consequences (Donde 2017).

Responsibility also lies with funders of self-driving car development. These include car manufacturing companies looking to turn a profit, academic institutions that want to be at the forefront of cutting-edge developments in this arena (Mcity 2018), and government projects, grants, and competitions that motivate the development of energy-efficient vehicles (DARPA 2018). One common argument for autonomous vehicles made by the media and car manufacturers is the potential for self-driving cars to decrease the number of car accident-related deaths every year. However, the numbers of deaths in themselves do not motivate reform. During the Vietnam War, there were 58,220 recorded U.S. military deaths, and 940,677 deaths from car accidents were recorded during the same period. The far greater public protest and mobilization around the Vietnam War deaths were rooted more in the nature of and context for the deaths than in the sheer numbers of lives lost. Likewise, the reason for why an issue becomes a national concern is significant in itself, and where money is invested has a large say in what the spotlighted issue will be.

The third contextual level consists of the political, cultural, and economic systems in which the technology is developed. I will be treating the autonomous vehicle as an engineered product in the context of a society that is also engineered by human behavior, biases, and value systems (Gusfield 1996).

I begin in Chapter 2 by breaking down the different layers of the self-driving car from a technological perspective. This will allow us to see where the areas of unpredictability and liability lie in engineering design.

In Chapter 3, I examine legal precedents concerning product liability affecting cars, and how responsibility for accidents is allocated. The chapter will explore liabilities taken on by pedestrians, car drivers, and manufacturers, and how regulation and litigation developed with respect to each of these responsible parties.

Finally in Chapter 4, I propose areas in which autonomous vehicles offer increased safety compared to current driving conditions, as well as new areas of liability that the car introduces. Many of the legal precedents in place today will be readily applicable to the new areas of liability, but there are additional areas that require new regulation, including distracted driving, data security, and ethical standards for decision-making and vehicle design by manufacturers.

Considering the context, risks, and liabilities entailed is critical in the development of any revolutionary technology. When disasters such as the Fukushima nuclear meltdown or thousands of smog deaths in New Delhi are called "unintended consequences," responsibility is shunned by all parties involved, and discussion of how such consequences might be prevented becomes almost futile. As leading STS scholar, Sheila Jasanoff, has explained, "[T]he language of unintended consequences… implies that it is neither possible nor needful to think ahead about the kinds of things that eventually go wrong" (Jasanoff 2016, 23). The label "unintended" seems to indicate that the distinction between intended and unintended can only be determined retrospectively. In other words, the concept is biased towards reassuring people that *intended* technological change is generally positive. The term "unintended" also tends to fix the point of determining intentions at a specific moment in time, even though technology continues to reiterate and develop through many phases, feedbacks, and decisions. In the spirit of challenging the concept of "unintended consequences", I now move to analyze the specific technologies and systems involved in the development of the self-driving car, and to consider how preventive regulatory measures can be designed and liability and responsibility can be allocated when accidents occur.

# 2 A Technological Overview of the Self-Driving Car

In order to explore the legal implications of switching from traditional to self-driving cars, I begin with a technological overview that can help reveal the practical capabilities, limitations, and novel risks of autonomous vehicles. Recognizing that there is rapid ongoing development in these technologies, this chapter lays out the systems that are currently being explored and the architecture of how they fit together.

The self-driving car is composed of three systems: the mechatronic, autonomous mechatronic, and cross-linked mechatronic (Liu-Henke 2017). These can be thought of as layers of increasing computational abstraction, with the mechatronic system focusing on the mechanical, physical components of the car, the autonomous mechatronic system incorporating its software processing and control functions, and the cross-linked mechatronic system involving communication among multiple vehicles, possibly using a smart-car grid or other centralized information system.

The boundaries between these systems are far from clear in practice. For example, the distinction between mechatronic and autonomous mechatronic systems is not simply that of physical components and data systems. The information to be processed at the autonomous mechatronic level must be collected at the mechatronic level, and the variance and noise in data collected at the mechatronic level need to be interpreted at the autonomous mechatronic level. Thus, mechatronic sensors and components already process data in choosing how to represent the outside world informationally, and autonomous mechatronic software programs can help further manipulate the data into useful forms before running action-planning algorithms.

Likewise, each of the three levels informs the design of the others, and thus are only loosely divided into distinct areas in which researchers have been focusing their development efforts.

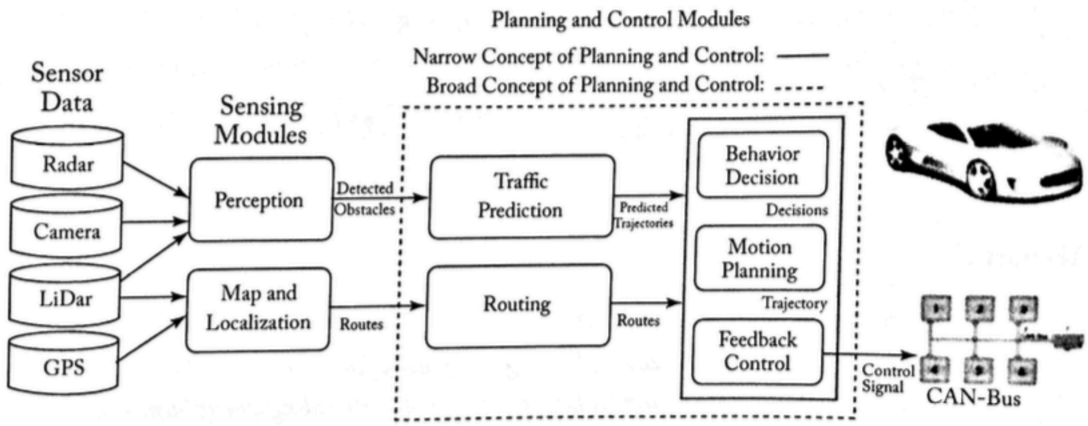


Figure 1: Modules in the mechatronic and autonomous mechatronic systems (Liu 2018, 84)

## 2.1 The Mechatronic System

The traditional car's mechanical subsystems will mostly remain relevant for the autonomous vehicle, but automating the driver's functions will require many additional components. Each of the original subsystems will need to measure information through added sensors, relay the information through various channels, and actuate the mechanical components according to determined courses of action for the car. Current sensors exist in the car for monitoring tire pressure, engine temperature, rear images, etc., but measurements are only used to inform the driver. In the self-driving car, many more sensors will need to be added to replace human perception, and data inputs will be processed by the autonomous mechatronic system.

Functions that the human driver normally performs on the road, such as estimating road conditions, traffic, and other environmental factors, as well as assessing of the status of the car, including gas consumption, strange noises, deflated tires, and the car's response to braking or other instructions, must all be measured precisely and accurately by sensor technology. We can analyze how sensors are integrated into the self-driving car by focusing on their categorical functions: knowing the location of where the car is and wants to go, accurately observing and estimating the car's surroundings, and getting the car to drive according to the autonomous system's commands (Schweber 2018).

***Knowing where you are going: Tracking location***

In order for the car to navigate successfully to its destination, it needs to keep updating its current location and route and check both against its destination, recalibrating for unexpected road blocks and diversions. The Global Positioning System (GPS) is a commercially available solution already integrated into many cars. It reads signals from four low-orbiting satellites closest to the car and triangulates its location relative to the satellites. GPS is readily available as a sophisticated system on a chip (SoC) integrated circuit (IC), or a multi-chip chipset that works with the autonomous mechatronic system to triangulate the location using only power and an antenna.

The GPS system has an error margin of about one meter and a slow update rate of about 10Hz, or once every six seconds (Boucher 2010). This makes it suitable for location tracking but not precise enough to guide driving in real-time, which would require navigating within one meter of obstacles and a reaction time of a fraction of a second. Furthermore, the GPS signal is often out of range during significant stretches of a typical ride, including when the car goes into tunnels, when radio signals interfere, and when the car drives in poorly connected areas such as mountains. In a scenario where passengers of self-driving cars are not expected to be knowledgeable about the cars' routes, it would be important to have a reliable navigation system at all points along the route. Cross-communication with the car's real-time localization technologies covered in the next section may help remedy these locational blind spots. Cars may also require an inertial measurement unit (IMU), a system of gyroscopes and accelerometers that measures the relative motions of the vehicle from its last GPS-determined location and can provide real time updates at around 200 Hz (Hoflinger 2013).

Additional barriers to GPS usage include the problem of how an autonomous vehicle that starts its journey in a location not reachable by GPS signal can navigate, as well as the danger of a third-party intentionally manipulating or interfering with GPS signals (see Chapter 4). IMUs that can track over long distances are currently available in space crafts as a spinning-wheel gyro and gimbaled platform, but they are too large, costly, and power intensive for an autonomous vehicle. Thus, manufacturers are designing Micro-Electro-Mechanical System (MEMS)-based gyros and accelerometers that would be compact enough to be practical for use in the self-driving car (Schweber 2017). Given that cars today already use simple versions of MEMs accelerometers that are serve as triggers for airbag deployment, expanding functionality to location-tracking is easily imagined.

### *Seeing where you are going: Estimating the car's surroundings*

In addition to knowing its geographic location, the car will need to visualize its immediate surroundings to navigate safely. There are three localization systems currently undergoing development for the self-driving car: camera, radar, and LiDAR technologies. The most advanced autonomous vehicles today use all three.

A self-driving car would use input signals from eight or more cameras placed around the exterior of the car to stitch together and make a 3D model of its surroundings. One systems architecture that does this has the cameras capture, compress, and send the images to the Camera

Control Unit (ECU). The ECU then decodes, filters, and geometrically corrects the images, before sending them to the autonomous mechatronic system to stitch the images together and identify objects. An alternative systems architecture has the cameras process the images themselves before sending it to the autonomous mechatronic system, eliminating the ECU and allowing a completely decentralized process before computer vision algorithms are carried out by the autonomous mechatronic system (Lins 2016).

One challenge in obtaining an adequate camera system is the need for an extremely high dynamic range (a minimum of 130 dB) to obtain clear pictures even with sunlight glare. This is close to the most expensive and cutting-edge camera currently available (145 dB) costing a couple thousand dollars per camera. With eight or more cameras installed in the car, this visual system alone would cost $16,000, which is not economically feasible given the $30,000 price point for mass-producible cars that consumers are paying for today (May 2017). In addition, processing live video streams from eight or more 1080p cameras at 60 frames/second would generate about 1.8 GB of raw data (Liu 2017, 3). This is a massive amount of data that needs to be compressed and processed in real time, requiring several additional FPGAs, which are integrated circuits programmable after manufacturing, with current technology. That goal is also not yet possible for reasons of cost, space, and power for a commercially available autonomous vehicle (Rudolph 2017).

The second localization system uses Radio Detection and Ranging (RADAR), placing a combination of short-range radar (SRR) and long-range radar (LRR) sensors in various formations around the car to detect and locate objects surrounding the car using radio waves. For ideal collision-preventing operation, the SRR sensors would use a 79-GHz frequency band with a 4-GHz bandwidth, but this is outside regulated global frequency specifications, which only allows a 1-GHz bandwidth at a 77-GHz frequency band. The resulting compromise on the relative resolution yields much higher margins of error than necessary for safe navigation.

Baseband processing is currently carried out on a radar monolithic microwave integrated circuit (MMIC), but the development of a raw data radar sensor would enable far more efficient and precise processing. The latter option would provide the unfiltered raw data to the radar process controller through a MIPI CSI-2 interface, which would allow data transmission without any loss or compression. If the raw data radar and baseband functions are integrated together on a monolithic SoC with further development, medium and long-range object detection would be highly effective in the autonomous vehicle (Rudolph 2017).

The third and most accurate system uses Light Detection And Ranging (LiDAR) to transmit a laser beam and calculate the distance of an object based on the time it takes for the beam to return to a highly sensitive receiver. Using a rotating laser and MEMS in an infrared LiDAR system, where a pad of Single-Photon Avalanche Diodes (SPAD) receives the emitted beams, can produce a 360 degree, 3D rendering of the surrounding environment. Another version undergoing development is a solid-state LiDAR which utilizes over 100 laser diodes with nanosecond pulse widths and a correspondingly large array of receivers. A LiDAR unit usually rotates at around 10 Hz and takes 1.3 million readings per second, allowing it to create HD maps, localize surrounding objects, and detect obstacles (Liu 2017, 3).

All of the most advanced autonomous vehicles beta tested so far, most notably the Google Car, use a LiDAR system to detect their surroundings. The largest barriers for LiDAR being used in mass deployment, however, include accuracy under changing external conditions, such as the capability to reliably detect objects up to 300 yards away in extreme weather conditions including rain, snow, and temperature variation (Rudolph 2017), as well as the current cost of several tens of thousands of dollars per system. To address these issues, lasers today have temperature control built into their packaging, and systems use multiple lasers at different wavelengths to address issues of absorption through moisture in the air (Pfennigbauer 2011).

A combination of the three localization systems provides a costly but highly effective solution for visualizing a car's surroundings in real-time. The main barrier here is the price point of the technology, as exemplified by LiDAR costing $75,000 per car at the end of 2017 (Mitchell 2017). This is far higher than the average price consumers are currently paying for a new car, which is around $30,000 (Mays 2017). Given that LiDAR is only one of the myriad technologies that need to be added to the traditional car to make it self-driving, cheaper solutions must be found before mass production is possible.

***Getting where you are going: Augmenting traditional car subsystems***

The final part of the mechatronic system covers the components needed to operate the car to drive where it needs to go. For each of the traditional car subsystems, additional sensing and actuating components must be added to stand in for the roles currently played by the human driver (Liu-Henke 2017).

In the tire system, the autonomous vehicle needs to control for the contact between the tire and the road surface. This not only requires sensors for tire pressure, wearing of the tire treads, and other maintenance factors, but also an approximation of the friction between the tires and the road surface. The human driver intuitively accounts for the response of the car on icy, bumpy, or damaged road conditions, and the autonomous car will need to measure all these and respond accordingly.

One of the most challenging aspects of automating the driving and braking system, or more specifically motor and torque control, is accurately measuring the vehicle's velocity relative to its surroundings and other moving vehicles. Traditional cars most commonly use odometry, or a rotary encoder that uses an LED and a detector in the wheel to measure rotations and position changes, but it is subject to errors ranging from 5-25% due to wheel slippages, surface roughness, and other misalignments (Boucher 2010). Thus, even though the speedometer is used as a benchmark while driving, the driver has to sense the relative motion of the vehicle in different ways when braking and merging in traffic, avoiding obstacles, and other critical real-time actions. Likewise, the GPS has errors in the order of meters compared to the vehicle's true location, and IMUs generate cumulative errors as velocity errors are a function of time (Hoflinger 2013). The most cutting-edge solution to this problem is to use localization systems such as those described above to estimate the vehicle speed from processed images, leaving the autonomous mechatronic system to compute optimal motor and torque control from this information.

The suspension system must balance between safety and comfort. Specifically, it must gauge the ability to accelerate, brake, and corner while optimizing the comfort of the human rider on a bumpy road (Harris 2005). Human drivers usually maneuver a car in various road conditions judging the safety of the particular maneuver and the comfort to themselves. I return later in the chapter to the implications of leaving the balance of these factors and the computation of optimal control to the autonomous mechatronic system's algorithms.

Energy management, including gas consumption and battery power, is another system that is normally left to the driver's discretion. To be self-driving, the car will not only need to accurately model the nonlinear battery life and other energy components, but will need to plan for when and where these components must get replaced or recharged.

There are many types of steering even for the traditional car, including rack-and-pinion steering, and recirculating-ball steering. Steering design for the autonomous vehicle will need to include an axle for controlling the steering angles, and one for stabilization (Nice 2001). The car needs sensors to measure the relative movements in these axles. If actuators are used to move the car once an action has been decided upon, another challenge lies in determining when to allow human overriding and when to allow the actuators to intervene in human driving to avoid dangerous situations. The design will need to be modified according to the role that the rider is given in the car, a role affected by some of the legal and regulatory choices discussed later in this thesis.

All in all, advanced sensor technologies that can make most if not all necessary measurements required for safe vehicle monitoring and navigation already exist, but there are still barriers that need to be overcome before having a successfully self-driving car. The cost of localization systems and programs that integrate communication between sensors needs to come down drastically in order for mass-market production to be feasible. Also, as exemplified by the recent Tesla Model S and Uber self-driving accidents, manufacturers still need to design a reliable way to process the large amounts of sensory data and account for extreme but not infrequent sensory inputs such as sun glare.

## 2.2 The Autonomous Mechatronic System

The autonomous mechatronic system, comprising the data algorithms of the autonomous vehicle, takes in the sensory information provided at the mechatronic level, and computes an output action that it wants the mechatronic system to actuate. This system can be considered in three parts: algorithms that perceive objects from the localization mechatronic systems and predict the movement of those objects; algorithms that decide on an action and control for the car; and the managing client system that runs an operating system to execute interactions between the car's software and hardware components.

***Perception and Prediction in Autonomous Driving***

Perception of exogenous elements, or the identification of other vehicles, pedestrians, traffic signals, and other objects in 3D, is absolutely critical for the autonomous vehicle to drive safely. Yet, perception remains one of the most complex and challenging features undergoing development.

Computer vision algorithms and object classifiers have been used in localization systems to identify objects. These are enabled by training on large datasets. The fundamental problems solved by these algorithms include image classification, semantic segmentation, optical flow, stereo, and tracking. Image classification is normally done by preprocessing the images that extract distinctive features to reduce process time, computing histogram of orientation (HOG) features over a sliding detection window, and using a linear support vector machine (SVM) to classify objects (Dalal 2005). The 3D spatial information about these objects is computed using algorithms such as stereo, which matches similar points and edges between multiple images based on a cost weight function that calculates the difference of intensities between pixels (Hirschmuller 2008); optical flow, which pairs the motion of intensities between multiple images taken at slightly different times; and scene flow, which uses super-pixels to build a 3D parametric plane for moving objects, exploiting the assumption that there is piecewise rigidity in the images used (Menze 2015).

Computer vision algorithms in the self-driving car also utilize a deep neural network to deal with large datasets for significantly more accurate ground-truth labeling, and modeling visual perception with end-to-end training as opposed to extracting manually designed features plus subsequent classification or regression steps (Li 2014, 69). One of these deep neural networks is the Convolutional Neural Network (CNN), where each neuron in a layer has a localized receptive field, and exploits translational invariance in visual data to share the local connection weights spatially across the layer (Krizhevsky 2012).

Several object classification and localization algorithms are currently being developed utilizing CNNs, such as Faster R-CNN (Ren 2015), which uses a Region Proposal Network (RPN) to propose possible regions for the object and then judging the existence of the object by projecting the image onto a fixed-size feature map by Region of Interest (ROI) pooling layer and converting it into a feature vector. Other algorithms that skip the proposal step include SSD (Liu 2016), YOLO (Redmon and Divvala 2016), and YOLO9000 (Redmon and Farhadi 2016), which are being developed to reach real-time object detection, optimizing the tradeoff between runtime and accuracy. For the stereo method mentioned before, CNN can be applied using Siamese architecture to reach good 3D depth estimates, exemplified by the Content-CNN algorithm (Luo 2016). CNN can also be applied to optical flow, using an encoder-decoder architecture called FlowNet (Fischer 2015), or a coarse-to-fine approach called SpyNet (Ranjan 2017).

In addition to object perception, semantic segmentation, or scene parsing for areas such as the road and sidewalk, is critical. A Fully Convolutional Network (FCN) serves this purpose. It is formed by removing the softmax layer of a CNN and substituting a 1x1 convolution layer for the CNN's last fully connected layer (Long 2015). One of the most successful segmentation algorithms today is the pyramid scene parsing network (PSPNet), which won the 2016 ImageNet scene parsing challenge, a competition to accurately identify objects from images, with a score of 57.21% on the ADE20K dataset and 82.6% on the VOC2012 dataset (Zhao 2016).

While this is a large step forward relative to previous segmentation capabilities, it is still far from what would enable safe autonomous driving, where every error is a potential accident. PSPNet's accuracy scores can be compared to the current accident rate due to recognition errors, which account for 39% of vehicle accidents. As calculated by car insurance companies, the average person will be involved in 3-4 accidents in a lifetime (Aho 2016). Although it is hard to make a direct estimation of the number of car accidents that would correlate to a given accuracy rate for image classification, it is clear that the algorithms will need to improve accuracy significantly before they outperform humans.

Once the objects are identified, algorithms need to estimate the location, speed, and acceleration of the objects over time, otherwise known as tracking. Some of the important challenges include the tracking of rapid direction shifts especially in pedestrians and cyclists (tragically illustrated by the fatal Uber accident), partially or fully occluded objects, and dynamic changes to the appearance of the object due to position and lighting. Tracking has traditionally been accomplished by recursively solving a Bayesian filtering problem, alternating between the prediction step where the current object state is predicted from the previous state based on a temporal evolution model, and the correction step where the posterior probability distribution of the object's state is calculated from an observation model relating observation to the object state (Giebel 2004). The algorithm has found various applications including robotics, stock prediction, ballistics, and other problems that require tracking. Due to the fact that the particle filters used for this Bayesian approach make recovery from temporary detection failure difficult, another process under development is the tracking-by-detection method, where object detection runs across consecutive frames are linked together. A Markovian Decision Process is used to reduce the error of missed detection and false position from the object detectors (Xiang 2015).

Once the functionality of classifying and tracking objects is resolved, the next computational step for the autonomous vehicle is to predict the flow of traffic and routing decisions in the near-term. Behavioral predictions of drivers and pedestrians are derived not only from physical rules and attributes, but from other factors including historical patterns of behavior, scenario-based features, and map features. These behaviors can also be categorized by formalizing them into a classification problem, and running machine learning methods on them to solve the underlying regression problem. Commonplace examples of routing algorithms that can be adapted for the autonomous vehicle include Dijkstra (Lu 2012), which uses a simple shortest-path graph theory algorithm on a point-based model of vehicle routes, and a popular artificial intelligence algorithm A*, which is heavily reliant on a heuristic in order to search the highest-potential node first. Altogether, the routes predicted by the algorithms in this section serve as inputs in deciding what action the autonomous vehicle should take.

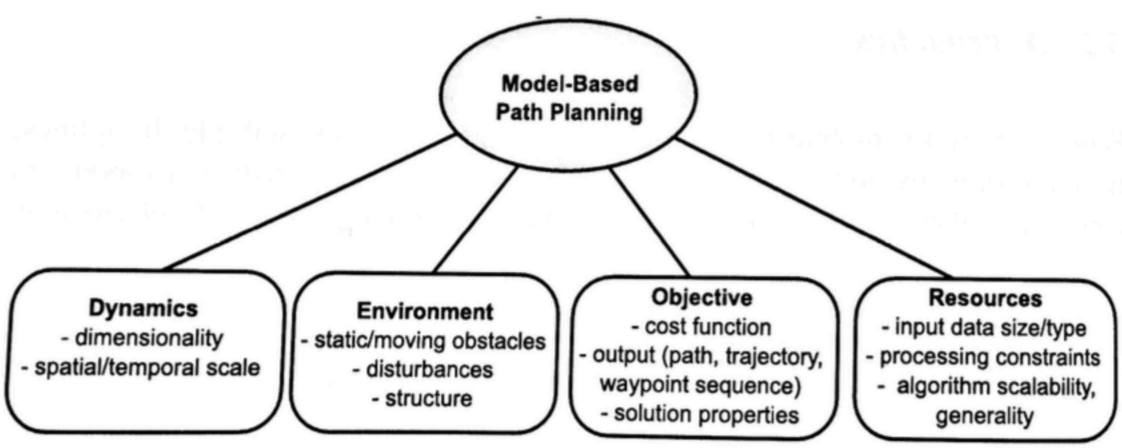


Figure 2: Aspects of model-based path planning approaches (Fossen 2007, 192)

### ***Decision, Planning, and Control in Autonomous Driving***

The behavior decision module processes a majority of the sensor and perception data, including traffic lanes, vehicle attributes and dynamics, past behavioral decisions, obstacles around the vehicle, map objects, and driving laws, to make an efficient and safe driving decision. Because of the highly diverse and dynamic nature of the data to be processed, instead of using a uniform mathematical model, advanced software algorithms use Bayesian models, Markov Decision Processes (MDP),and Partially Observable Markov Decision Processes (POMDP). These are all used to allow some probabilistic modeling of the data. These processes model the environment based on a state space, conditional probability transition model between states, a behavioral decision output space, and a reward function for the possible actions. Furthermore, using Dynamic Programming and incorporating Bellman's value iteration algorithm, allows for a computation of the optimal policy action without explicitly computing each of the probability distributions at each stage (Brechtel 2011).

The reward distribution of weights between the different modes of action will be of greatest legal and regulatory interest. Designing the optimal mode of action will require balancing the objectives of reaching the destination, safe and legal driving, and riding comfort, in a way that is appropriate and optimal in each situation. Who gets to decide this prioritization and how much responsibility the decision-maker has for different types of scenarios when the algorithm fails will fundamentally influence the design of the final mass-produced algorithm. Although the large number of variations in scenarios has prevented a direct-solution, deterministic algorithm thus far, a categorical rule-based system is still being researched, as the simplicity and reliability of a deterministic system would be ideal for mass production.

Once the behavioral decision is made, the motion planning module generates a vehicle trajectory for physical control execution with the street modeled as a 2D plane, finding the optimal spatial-temporal trajectory within spatial-temporal constraints. A cost function is used to ensure obedience to the behavioral decision outputted by the previous module, vehicle safety constraints around objects and traffic laws, and also constraints on the curvature of the trajectory such that the downstream feedback control module can comfortably execute physical maneuvering of the vehicle. Many algorithms are being developed to converge on the actual optimal policy based on these assigned costs. The most successful one to date decomposes the problem to first plan a trajectory shape on the 2D plane and then solves the speed planning based on that trajectory (Mcnaughton 2011). Clearly, there are drawbacks to optimizing these two factors separately, so alternatives such as an SL-coordination system using longitudinal and latitudinal optimization are also being researched.

Finally, the optimal trajectory is passed onto the feedback control module, which constantly compares the actual position of the autonomous vehicle with the proposed trajectory to correct for differences. Feedback control relies on a bicycle model preventing lateral movement and rotating by steering the front wheels in a 2D plane under physical constraints (Paden 2016), and running a PID control algorithm on the model to reduce error between the vehicle's pose and the proposed trajectory. The PID controller would track vehicle position as a cartesian location of the center of the vehicle along with its rotation angle from the x axis, and minimize a weighted function of the error, integral, and differential components to maneuver the car along the trajectory.

To robustly account for all of the unrestricted real-world road cases, and not just a limited set of scenarios, learning based approaches that use the big data amassed from human drivers can enhance the decision, planning, and control modules. Specifically, reinforcement learning is used by the autonomous vehicle to take in feedback sensory data after it takes an action to get rewards from certain results and iteratively improves its underlying model assumptions based on the feedback. One of the most popular reinforcement learning algorithms is Q-Learning, which calculates the optimal expected return for different actions, and then chooses the action with the largest expected return. The Q function is then solved in a backward propagated fashion by Dynamic Programming. However, due to the large number of possible scenarios and actions, computation of the optimal expected return is impractical for real-time industry use, and so deep neural networks are introduced for computation efficiency. Example combinative approaches include Deep-Q-Learning (DQN), which is similar to the neural networks described previously in computer vision, except without pooling layers (Mnih 2013), and the Advantage Actor-Critic (A3C) algorithm (Mnih 2016).

In such a learning based approach, the biggest challenge and responsibility lies in defining the scope of the episodic learning tasks or valid sequence of states, designing the reward function generated from the environment, and designing a value function that weights the desirability of different states. Thinking about how to define these aspects for each of the modules becomes then the most important aspect of the problem. With respect to training the systems after design, the most effective way to train reinforcement learning algorithms is to use the massive amounts of human driving data being generated, but a drawback of this approach is that it is highly sensitive to the quality of the data, and would easily learn biases and errors that are inherent in human behavior as well. Thus, striking a balance between efficient data processing, which uses simplified models and subsets of data, and robust algorithm training using more complex parameters and large data sets, is an important decision that may need to be regulated.

### *Client Systems for Autonomous Driving*

The third major piece of the autonomous mechatronic system consists of the client systems that allow the modules to function together: the operating system which coordinates the interactions between different components and the computing platform which is in effect the "brain" of the car.

The operating system coordinates the communications among all the components in the sensing, perception, computation, and execution processes, and allocates limited CPU resources to whatever functions are most critical. Most autonomous driving operations today use a variation of the Robot Operating System (ROS) to provide these capabilities (Quigley 2009). The ROS treats the different modules which each perform a task as nodes, with transmission of information between the nodes treated as messages. The Master node stores all the names of the nodes, connections, and services, with services being the nodes that accept requests from the computing platform and return the results of a request. The highest risk this system poses is the likely probability of a node crashing; most dangerous is the possibility of the Master node crashing. Thus, decentralization and control systems that use combinations of different nodes and multiple master nodes are also undergoing development, such as the ZooKeeper mechanism (Hunt 2010). Furthermore, this operating system is the part of the autonomous vehicle that is most vulnerable to hijacking. If any one of the nodes is compromised, the system can easily misallocate or run out of computing resources, leading the vehicle to malfunction. Robust cyber security measures and possible design modifications must be taken into consideration at this stage.

The main considerations affecting the computing platform are cost, time efficiency, and power consumption. Progress is being made on each of these fronts in such diverse and dynamic ways that it is difficult to discuss common processes specific to the autonomous vehicle. Research has been done to ballpark the computational power needed compared to current capabilities, as well as the cost of meeting such requirements today. One example of a first generation computing platform consists of two compute boxes with Intel Xeon E5 processors and four to eight Nvidia K80 GPU accelerators connected to a PCI-E bus. Altogether, the system delivers 64.5 TOPS/s at 3,000 W. If one of the boxes fails and both boxes run at their peak in the worst case scenario, they would require over 5,000 W of power consumption, generating an unacceptable amount of heat. Each box as of the end of 2017 costs $20,000-$30,000, making the solution as a whole not economically viable for mass commercial deployment.

To get a sense of the computing hardware solutions under development, I briefly list some of the available products. These include GPU-based solutions with decentralized memory and DNN acceleration, DSP-based solutions optimized for computer vision acceleration, FPGA-based solutions optimized for sensor fusion, and ASIC-based solutions that allow heterogeneous accelerators that are programmable with their own set of algorithms. In terms of the computing units most suitable for different workloads, DSP was shown to be the most efficient computing unit over CPU and GPU for convolutional and feature extraction workloads, taking 5 ms per convolution at 7.5 mJ and 4 ms per feature extraction convolution at only 6 mJ (Liu 2016, 164). Further research on the limits of computation for each of these hardware elements and their suitability for self-driving functions such as localization, planning, and control will be key in the design of the autonomous mechatronic system.

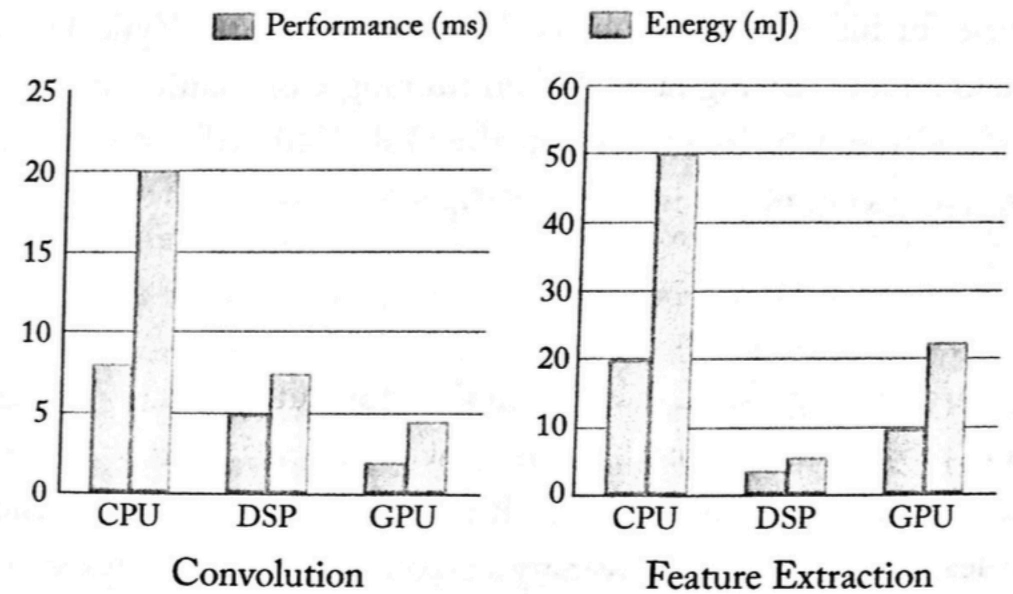


Figure 3: CPU, DSP, and GPU comparisons in performance (left bar) and energy (right bar) (Liu 2018, 164).

## 2.3 The Cross-Linked Mechatronic System

An aggregation of driving data across multiple cars opens up the potential to learn driving behaviors from many different scenarios, create an open-source visualization of road conditions, and communicate between vehicles. Autonomous driving clouds would support the functions of the autonomous vehicle by providing capabilities such as distributed simulation tests for new algorithm deployment, offline deep learning model training, and HD map generation. The infrastructure of such a driving cloud includes distributed data storage and computing capabilities that allow for flexible allocation of resources across applications, as well as a heterogeneous computing layer that prioritizes and accelerates different kernels on GPUs or FPGAs depending on what optimizes computational and energy efficiency.

The most widely adapted solution for in-memory storage has been the Hadoop Distributed File System (HDFS), which provides persistent storage, but memory-centric distributed storage systems, such as Alluxio, which utilize cluster frameworks, are being developed to manage multiple storage layers at the same time (Li 2014). For distributed computing, the Hadoop MapReduce engine is the most commonly used framework. It reads input data from disk, maps a function from the data, reduces the results, and then stores them on disk (White 2012). However, Spark, a more recently developed distributed computing solution, uses a resilient distributed dataset (RDD) which keeps track of a working set of distributed programs, overcoming MapReduce's forced linear dataflow structure (Zaharia 2010). OpenCL for heterogeneous computing acceleration (Stone 2010). The last infrastructural problem of constructing a heterogeneous computing layer serves to increase power and computational time efficiency by allocating resources to be processed on CPU or GPU, depending on whether the process requires low power consumption (CPU), or multiple convolutions and parallel processing (GPU). Systems utilizing applications such as OpenCL, YARN, and LXC are being developed to provide this service across the distributed computing programs.

On this cloud infrastructure, artificial intelligence systems can be used to augment driving performance, including but not limited to running offline simulations of improved autonomous vehicle capabilities, deep learning which allows the vehicle to upload data it has generated while driving and update its models based on processed results, communication with surrounding connected vehicles, as well as the building of a centralized 3D high-definition map.

## 2.4 Accident Reduction Enabled by Self-Driving Technology

The technology and algorithms of the self-driving car present significant economic and social opportunities if the initial costs of development can be overcome. Here, I focus on their potential to reduce vehicle accident rates in the United States, because it is most relevant to our discussion of liability. According to the National Safety Council, there were 40,200 traffic deaths in 2016, up 6% from the year before (Boudette 2017). This makes car accidents the leading cause for non-disease-related deaths in the United States (Nichols 2017). Each of the reasons for vehicle accidents can be examined in turn to determine the benefits that self-driving technology has to offer towards preventing crashes. In 2015, the National Highway Traffic Safety Administration (NHTSA) released a report that collected statistics on the causes of lightweight vehicle crashes over two and a half years. Recording the immediate reasons for a pre-crash event, or "critical reason" for each accident, NHTSA found that 94% of the critical reasons were driver-related (Table 1).

**Table 1. Driver-, Vehicle-, and Environment-Related Critical Reasons**

| Critical Reason Attributed to | Estimated | |
|---|---|---|
| | Number | Percentage* ± 95% conf. limits |
| Drivers | 2,046,000 | 94% ±2.2% |
| Vehicles | 44,000 | 2% ±0.7% |
| Environment | 52,000 | 2% ±1.3% |
| Unknown Critical Reasons | 47,000 | 2% ±1.4% |
| Total | 2,189,000 | 100% |

*Percentages are based on unrounded estimated frequencies
(Data Source: NMVCCS 2005–2007)

These driver-related reasons will all become vehicle-related in the self-driving car. Legal provisions would then need to address the technical performance of the vehicle as built into the mechatronic, autonomous mechatronic, or cross-linked mechatronic systems. Using these three technical levels as lenses of analysis, I will now identify forces that may influence the number of accidents to increase or decrease for each critical reason.

For environment-related reasons such as slick roads, glare, view obstructions, and road design, the autonomous vehicle can be expected to perform more poorly than human drivers in the early years of development. To respond safely to these extreme conditions, we would need sensors and algorithms to process data points outside of typical driving patterns. As mentioned in the mechatronic section, economically viable sensor technologies do not yet collect information at the resolution or accuracy that human drivers are able to process. For example, the Tesla Model S accident described in Chapter 1 occurred when a sensor did not register glare off a white trailer, which a human driver would immediately have seen. However, as sensors continue to improve and become cheaper, they may trend towards the point where they surpass human sensing in driving and decrease environment-related accidents. Sensors are already widely used to augment human sensing and provide information beyond the limits of raw human perception, such as those in microscope cameras or gas detectors. A solution that deploys such sensors effectively in the autonomous vehicle can certainly be imagined.

At the autonomous mechatronic level, the programs that engineers write to manipulate the car safely in everyday conditions may not account for all of the extreme environmental situations that cause accidents. Human ability to extrapolate from past experiences and apply intuition to new challenging situations allows drivers to navigate icy road conditions safely, avoid blind spots due to sun glares, and navigate under other strenuous conditions. Here, reduction in environment-related accident rates will come not only from faster and more reliable processing at the autonomous mechatronic level, but also from the cross-linked mechatronic level. Artificial intelligence and machine learning algorithms including supervised and unsupervised deep learning would enable cars to learn from experience much as humans do. Moreover, aggregate data from all cars on the road and the ability of any car to gain the experience of scenarios encountered by every other autonomous vehicle will in principle allow the car to surpass human levels of learning. Environment-related accidents can accordingly be expected to increase at first, but eventually to drop below current levels with advancing computer algorithms.

Vehicle-related factors that cause accidents currently include tire, brake, steering, suspension, transmission, and engine problems. Because failures in these systems are closely tied to natural wear-and-tear and to production defects, they would be expected to continue in the autonomous vehicle as well, regardless of the level of automation. In the traditional car, mechatronic-level problems consist of mechanical components that respond suboptimally to control commands by the driver. A greater number of sophisticated sensors and parts in the autonomous vehicle will undoubtedly increase the complexity of these problems. Moreover, the autonomous vehicle's mechatronic system will also include actuation systems and mechanical control inputs. This requires reliable coordination and execution of a myriad mechanical processes normally carried out by the driver, from signaling to turn left to braking at the right times and speeds. These added functions will introduce additional possibilities of error leading to accidents.

The autonomous vehicle also has advantages over the traditional car in the management of mechatronic errors. For one, the sensors in the autonomous vehicle allow for far more precise and regular monitoring of components in the car that may be subject to wear and tear. Being able to replace components before they become a liability may help to prevent some fraction of vehicle-related accidents. Relying on algorithms to actuate driving functions might also allow for elimination of human errors such as forgetting to turn on the signal before changing lanes, a slip of the hand on the steering wheel while reaching for a drink, etc.

The last category of critical reasons for vehicle accidents is also the largest: driver-related accidents. The act of removing the driver completely from the process of driving means that the design and capabilities of the autonomous vehicle will directly define improvements and failures in this category, and thus have a direct impact over a vast majority of scenarios that that currently result in vehicle crashes. NHTSA provides the following breakdown of types of driver error (Table 2).

**Table 2. Driver-Related Critical Reasons**

| Critical Reason | Estimated (Based on 94% of the NMVCCS crashes) | |
|---|---|---|
| | Number | Percentage* ± 95% conf. limits |
| Recognition Error | 845,000 | 41% ±2.2% |
| Decision Error | 684,000 | 33% ±3.7% |
| Performance Error | 210,000 | 11% ±2.7% |
| Non-Performance Error (sleep, etc.) | 145,000 | 7% ±1.0% |
| Other | 162,000 | 8% ±1.9% |
| Total | 2,046,000 | 100% |

*Percentages are based on unrounded estimated frequencies
(Data Source: NMVCCS 2005–2007)

Here we see that recognition error, which includes "driver's inattention, internal and external distractions, and inadequate surveillance," accounts for a full 41% of driver-related reasons for accidents. Decision error, which includes "driving too fast for conditions, too fast for the curve, false assumption of others' actions, illegal maneuver and misjudgment of gap or others' speed," accounts for another 33%. In about 11% of cases, performance error such as "overcompensation, poor directional control, etc." was the primary reason for accidents, and in about 7%, non-performance errors were critical, with the most important cause being sleeping at the wheel. For each of these errors, the autonomous vehicle would at the very least offer regularization and increased reliability. Automation would eliminate significant variation-related errors that come from inconsistent human behavior and attention, as well as from conscious risk-taking to break the law. After all, the car does not daydream, fall asleep, or make illegal U-turns. Instead, “driver-related” error in autonomous vehicles would come from software processes that do not respond appropriately to unaccounted for or incorrectly programmed road situations, or to hardware flaws from limited processing power or deficiencies in sensor capabilities.

## Conclusion

It is evident that components and algorithms in each of the three layers of computational complexity of the self-driving car are still undergoing intensive development. Each system shows promising signs of technological improvement in precision and efficiency, while reducing cost and power consumption, and presents tangible methods of reducing vehicle accident rates.

However, each also incorporates its own risks and limitations that carry legal implications. In addition to the technical barriers mentioned above, there is a more general need for a normative discussion on how a desirable reward function ought to be determined for reinforcement algorithms, prioritization of systems processes, and weighing of the costs of different driving behaviors. At present, optimization algorithms are primarily concerned with economic and efficiency metrics and are only loosely tied to the human behavioral consequences and liabilities that would be brought about by each of the car’s decisions and actions. Each modularized technological problem discussed in this chapter should therefore be recognized as carrying highly consequential ethical choices in its design. The legal implications and proposals in the following chapters would not only prepare and compensate for shortcomings in the autonomous vehicle, but would also assign legal responsibility for conscious choices made in the autonomous vehicle’s design.

# 3 Apporttioning Responsibility

Traditionally, liability laws for the automobile influenced vehicle safety design by placing a financial burden on manufacturers, and by damaging, in some cases even altogether destroying, the car's brand image. From the 1950's to the present day, there has been a trend toward liability liberalization, where plaintiffs can claim that a design defect made an accident worse even if it did not cause the accident (Graham 1991). In addition, there has been an increasing number of arguments for strict liability, such as considering the manufacturer's post-market safety measures as evidence that the product had a defect in its original design (Graham 1991). In light of these trends, questions of legal responsibility seem to influence vehicle safety design.

Yet, economists such as John Graham and Murray Mackay have questioned whether liability law has such a controlling influence on vehicle design. From a series of interviews with industry and government experts, Mackay drew the conclusion that "strict liability… has held back new designs, consumed resources that might otherwise have been directed at design improvement, and added on costs to the consumer" (Mackay 1991, 29). He argued that federal regulation of automotive design, internationalization of design and production, and automobile industry consolidation had much greater impact on the pace and direction of safety-related innovation since the 1960's. However, all of these conclusions were based on his perception that American safety design lagged behind European standards. He did not consider a variety of factors other than liability laws that may have suppressed safety improvements in the United States, such as consumer aversion to safety features, or availability of safety equipment.

In a slightly more quantitative analysis, John Graham correlated the death rate from 1950 to 1990 with variables such as alcohol consumption, adjusted measures of the volume of liability lawsuits, and regulation levels, and he concluded that there was insignificant correlation between liability awards and fatality rates (Graham 1991, 67). His conclusions did not account for the fact that fatality rates are not synonymous to vehicle safety design, however, and also the fact that fatality rates are caused by all vehicles on the road, including outdated and unsafe designs, and not limited to vehicles equipped with the newest safety innovations. Neither Graham's nor Mackay's negative conclusions about the impact of liability on car safety design are wholly convincing, but they do indicate that there are a myriad socioeconomic, political and cultural factors that may influence technological safety. Therefore, this chapter analyzes the development of regulations and liability law addressing each party responsible for vehicle liability.

## 3.1 Traffic laws require pedestrians to assume greater responsibility for accidents

Vehicle safety begins with the people involved firsthand in a car crash, pedestrians (including bike riders) and car drivers. Allocation of responsibility between these two parties was the most contentious issue when the car was first introduced to the road. The car's legitimate place on the road first needed to be established. When the first commercially successful automobiles were introduced by Ford Motor Company in 1908, in the form of the Model T, the road was still regarded as infrastructure reserved for walking, horse-drawn carriages, and children playing in the streets. The car was seen as an interloper that had to conform to the road rules established over centuries of human civilization, and any accidents that happened between car and pedestrian were automatically attributed to the driver's fault. A social reconstruction of transportation was needed to establish vehicles as legitimate occupiers of road systems, along with subsequent traffic rules that shifted some of the burden of safety to pedestrians. This repurposing of the roads only evolved over the course of several decades through heated debate and interest group power struggles.

In 1906, automobile lawyer Xenophon P. Huddy published a now well-known book on the nature and status of the automobile on the roads, *The Law of Automobiles* (Huddy 1906). In it, he discussed how to determine what happened at the scene of an accident, and the transfer of responsibilities and liabilities in different car rental and sale situations. He also considered issues of manufacturer and regulatory responsibility, although at the time “no reported case in reference to the manufacturer’s part played in automobiling has as yet appeared” (Huddy 1906, 403). There was still an ongoing debate over “whether or not the United States government should, to any extent, control the operation of automobiles and seek to take the matter out of the hands of the states” (Huddy 1906, 457). No firm rules on matters of vehicle safety were in place at the time, but Huddy’s reflections show a growing awareness of distinctions among different parties involved in ensuring automotive safety. As more automobiles took to the roads, alliances were shifted or created, each advocating its own perspectives on the road and discourses to legitimize its campaigns. The alliances were dynamic, but recognizable groups included pedestrians, parents, educators, business associations, traffic engineers, legislators, public transportation operators, motorists, and of course vehicle manufacturers.

According to, *Fighting Traffic: The Dawn of the Motor Age in the American City*, a book by Peter Norton, an STS professor at the University of Virginia, each of these groups advocated for its own perspective on the street, using unique rhetorical strategies to promote its cause (Norton 2008). Three major groups were identified.  First, the social group, including parents of children and pedestrians, used morally charged arguments for justice, depicting the introduction of the automobile as an invasion. The second group, the “traffic controllers,” consisted of traffic engineers and business associations who saw unregulated roads as an inefficiency that ought to be optimized. The third, and ultimately the most influential, group in the reconception of the road, was motorists, the main beneficiaries of automobile sales. This group appealed to American traditions of economic and political freedom to advocate for its view that roads ought to be purposed primarily for vehicles. Many records show that automobiles gained ground through the efforts of a powerful elite that lobbied for favorable legislation, despite a fervently adverse public, indicating that the proliferation of automobiles was not driven by public demand. Motordom's interests were clearly established as dominant by 1930, and the streets were widely accepted as a legitimate place for cars. While the term "jaywalking" would have been highly contested by pedestrians in the 1920s, by 1930, traffic laws that give the automobile the right of way were widely accepted (Norton 2008, 3-4).

Traffic rule enforcement was strengthened and enabled by a shift in legal treatment of responsibility for automobile accidents. According to a 1936 law review article by the lawyer and future president, Richard Nixon, initially "the rules of law applicable to automobile cases, were no different from those which had been developed in the days of the horse and buggy," and counsel at the time insisted that "the absolute liability of the keeper of vicious dogs or evilly-disposed mules should be imposed upon the drivers of 'devil wagons'" (Nixon 1936, 476). This evolved into a case-by-case evaluation, decided by juries, as to whether the driver or plaintiff had exercised the "required degree of care" (Nixon 1936, 477).

Jury verdicts almost always led to decisions in favor of the plaintiff regardless of the circumstances around the accident, so the courts shifted instead to trials by judges based on applicable standards of care. These standards were established whenever judges saw patterns of negligence in groups of similar cases. One example of such a standard was the Stop, Look, and

Listen Rule which gave pedestrians the right to cross as long as they first stopped, looked, and listened (Nixon 1936, 479). Even with these standards, the assignment of responsibility gradually shifted to place more of the burden on the pedestrian, though plenty of contradictory rulings remained (Nixon 1936). Judicial decisions thus reflected both the applicable traffic rules and the societal view of correct pedestrian and automobile uses of the road, and rulings were influenced by changing cultural views and external forces such as those mentioned in *Fighting Traffic*. What is consistent, however, is that notions of safety and the social reconstruction of the road around the automobile were driven by various advocacy groups lobbying for their opinions, and not by a predetermined set of legal principles.

The incentive to build roads during World War II to strengthen the nation's defense infrastructure had a large impact on evolving ideas of automobile safety. Official federal authority for liability regulations was only established in 1970, with the formation of the National Highway Traffic Safety Association. In the meantime, socioeconomic and political factors driven by agendas not directly related to motordom or pedestrian interests were critical in establishing the legitimacy of the vehicle on the road.

## 3.2 Driver responsibility and the social construction of laws

Traffic laws that regulate the safe flow of cars on a road, such as right of way, designated lanes, and speed limits, reflect a societal consensus on expected behaviors on the road and allocation of fault. It is no secret that the application of these rules in real life is far from a simple and literal application of a text to actual situations. One example with a high degree of consensus is that drivers and even law enforcement officers do not assume that the posted speed limit is the actual speed limit that will be enforced; in fact, driving below the speed limit will in many cases, especially on crowded highways, be more dangerous than driving 5-10 miles per hour above. On the other hand, if it is raining heavily or the roads are icy, safe driving behavior would require speeds below the posted limit. Thus, the focus of this section is not on how laws might be codified in precise or comprehensive ways, but rather to identify factors that contribute to acceptable driver behaviors and responsibility on the road. This treatment not only allows for a more nuanced consideration of driving regulations and expectations, but also offers flexibility in response to changes that may come from the adoption of the autonomous vehicle.

Of particular interest are the limits of driver control of the vehicle. Arguments over how much of the consequences of a car accident or precarious driving situation can be attributed to the driver's fault are often really arguments over how much of the underlying actions ought to be seen as lying within the driver's control. A salient example of how this boundary is constructed can be seen in the sociologist Joseph Gusfield's famous analysis of how driving under the influence, or drinking and driving, came to be established as unacceptable behavior.

In his book *Contested Meanings: The Construction of Alcohol Problems*, Gusfield explores how a new social reality was constructed out of a diverse body of competing viewpoints, and how perceptions of fault can change in different contexts. As Gusfield notes, even among groups that successfully pushed forward the adoption of public policies against drinking-driving, the way that "traffic professionals, alcohol experts, and citizen anti-drinking-driving movements conceive the problem of auto crashes and alcohol impairment differs considerably both in their commitment to specific countermeasures and in the emotional content of their support" (Gusfield 1996, 13). Today's condemnation of drinking-driving as an unavoidable moral judgment can be seen as just one viewpoint that won out through a series of historical developments.

It seems logical that an assessment of the dangers of a driver impaired by alcohol would lead to legislative measures that seek to prevent this behavior, but this judgement was far from straightforward. Acknowledging the myriad factors that go into vehicle safety, it is significant that society deemed driving under the influence of alcohol as a reasonable and worthwhile target to regulate. More concretely, a study from the 1950's showed that death and injury in car crashes came more significantly not from the impact of the crash, but from the "second collision" with dangerously designed equipment, objects, and ejection of the passenger due to poorly designed door locks (Gusfield 1996). To eliminate the most precarious aspects of a car crash, then, more concern might have been directed toward developing vehicles that keep the passengers safe after a car crash and safer road objects than to the sobriety of the driver. The very construction of vehicles and roads would then have been designed with the assumption that drivers would sometimes be impaired – whether from talking on cellphones, speaking with passengers in the car, sleep deprivation, or drinking – and fault would not be assigned solely for driving while impaired. Furthermore, a significant difference was created between the legitimacy of alcoholism as a reason for accidents, and other health issues such as driving while mentally or physically impaired. This discrepancy indicates that a positive correlation between a behavior and car accidents is not enough by itself to explain the extent to which a driver is expected to exercise control over the car. The allocation of responsibility depends on background social judgments of what is blameworthy conduct.

The specific political, religious, and socioeconomic factors that went into establishing drinking-driving as a prohibited practice are not relevant to our purposes. But we can step back from the example of drinking-driving and view regulation of driving behaviors as having a broader symbolic influence that not only allocates resources but serves to define "the public norms of morality and to designate which acts violate them" (Gusfield 1996, 171). While not all traffic laws have a moral purpose, such as for example bans on driving in the wrong lane, the boundaries of a driver's control and responsibility over the movement of the vehicle can be seen as a product of consensus – not on a fundamental analysis of how things are, but on a socially constructed understanding of how things ought to be.

## 3.3 Automobile manufacturer responsibilities for safe design

Discussions about vehicle safety first involved those most visibly involved in accidents, pedestrians and drivers, but questions of liability soon arose regarding safe design of the machine involved, or the automobile itself. While many professionals are involved in the design and deployment of a vehicle, from engineers to corporate heads, litigation usually seeks to hold corporations responsible as single, undivided entities. Because the law's main purposes are to compensate victims and deter future accidents, it made sense to pin liability on the entities with the deepest pockets, as opposed to isolating individuals who may have had a direct hand in bringing about the harm in question. Thus, this section focuses on decisions made by the leaders of corporations, whether in deciding on a design option or in producing a given design and allowing it to be used by the public. I use the term manufacturer in this chapter to refer to manufacturing companies as a whole.

The following discussion is organized around four factors that motivate vehicle manufacturers to pay attention to the safety characteristics of their design, or the underlying forces that shape the system of automobile product liability today. According to *Product Liability and Motor Vehicle Safety*, written by economist John D. Graham, the four core dimensions are: (1) self-imposed standards of ethics and professional responsibility, (2) consumer demand for safety, (3) regulations that compel safety design, and (4) laws that impose responsibility (Graham 1991, 3).

The first line of defense is the manufacturer's self-imposed standard of ethics and responsibility, which serves as a guideline even during the initial phase of vehicle design and which existed even before public awareness or governmental regulation. When Ford Motor Company first became a publicly traded company in 1956, it offered an optional safety package with its new Continental model. After the company invested millions of dollars in seat belts, safety door latches, sun visors, and other safety features, Ford realized that "safety doesn't sell." From 1955 to 1956, Ford's competitor Chevrolet was able to widen its sales lead over Ford from 70,000 to 20,000 cars by marketing powerful V-8 engines and jazzy wheels (Eastman 1984). Safe vehicle design may be an ethical standard that some manufacturers entertain, but often it competes with profitability, whether through direct investment in parts or investment in research and development of innovative designs. Since consumer demand is directly correlated to profitability, manufacturers' decisions are in effect hand-tied to public opinion.

Public opinion also played a large role in the establishment of government agencies to regulate car design and the development of product liability laws surrounding the automobile. When cases first shifted focus from pedestrian-driver conflicts to driver-manufacturer ones, litigation was not very friendly to car drivers. The first automobile product liability decisions in the United States established a standard of strict scrutiny, in line with the approach in most other western countries. In 1965, *Suvada v. White Motor Co.* established that plaintiffs must prove that their injury or damages resulted from a condition of the product, that the condition was unreasonably dangerous, and that the condition existed at the time the product left the manufacturer's control (Suvada v. White Motor Co 1965). This clearly placed more of the burden of proof on the plaintiff than the manufacturer. However, the public response served as a critical factor in shifting the weight of responsibility.

Rising demand for manufacturer accountability was best expressed in Ralph Nader's famous activist book, *Unsafe at Any Speed: The designed-in dangers of the American automobile* (Nader 1965). The book called for regulations on automobile manufacturers who Nader alleged had gained power by lobbying Congress and successfully resisting attempts to reform automobile safety in the 1950's and 1960's: "Accumulated power of decades of effort by the automobile industry to strengthen its control over car design is reflected today in the difficulty of even beginning to bring it to justice... when a car is involved in an accident, the entire investigatory, enforcement, and claims apparatus that makes up the post-accident response looks almost invariably to driver failure as the cause" (Nader 1965, 42). Nader cited several concrete examples of how manufacturers failed to incorporate basic safety measures and referenced several industry experts to add to his credibility. The pioneering nature of his accusations and substantial evidence he presented led the book to become a national nonfiction best-seller in 1966.

With Ralph Nader as its leader, the motoring public pushed for new governmental agencies. The Department of Transportation was established in 1967 and the National Highway Transportation Safety Administration in 1970. These agencies would be primarily responsible for safety regulations in the following years, as well for critical inquiries into notorious cases. One famous example is NHTSA's investigation into the fuel tanks of the Pintos and GM Vegas in 1974. The inquiries were spurred by concerns raised by the public as well as the Center for Auto Safety, founded by Ralph Nader and car safety expert Byron Bloch. With the creation of federal regulatory authorities to set minimum vehicle safety standards, Congress did not allow manufacturers to rely on compliance with existing standards as a general defense in liability lawsuits. Moreover, manufacturers were held responsible for complying with the common law standards of the fifty states, to account for interstate travel.

The trend towards holding manufacturers accountable for safe vehicle design can also be seen in judicial interpretations of the common law in liability cases. The first case to establish the manufacturer's duty to provide "crashworthy design" was *Larsen v. General Motors* in 1968. "Noting that 'automobiles are not made for the purpose of colliding with each other," the court recognized that collisions are a "frequent and inevitable contingency of normal automobile use," and that the manufacturer has a duty "to use reasonable care in the design of its vehicle to avoid subjecting the user to an unreasonable risk of injury in the event of a collision" (Larsen 1968). More generally, there was a trend in product liability law towards allowing plaintiffs to claim that a design defect aggravated their injuries even if it didn't directly cause the accident. This shift coincided with growth in litigation frequency, jury award amounts, and frequency of awards. Furthermore, there was a shift from contributory negligence, where rewards could be denied if the plaintiff's actions contributed to the accident in any way, to comparative negligence, where the plaintiff still receives a portion of the damage reward, even if the plaintiff's negligence contributed partially to causing the injury.

A famous early case of automobile manufacturer liability was *Grimshaw v. Ford Motor Co.* in 1981 (Grimshaw 1981). Here, a 1972 Ford Pinto was hit from behind on the highway at a speed in excess of 50 miles per hour. Due to Ford's decision to place the fuel tank behind the rear axle, the car burst into flames, killing the driver and permanently disfiguring the passenger. Ford was already aware that the Pinto's design decreased the crush space available upon impact, but had placed the fuel tank in the rear in order to add a hatchback and station wagon in the future. In addition, some company engineers had suggested that placing the fuel tank behind the rear axle would increase safety by preventing gasoline vapors from entering the car.

Immediately after the initial ruling, Ford appealed on the basis that the approaching car had crashed at a speed far above the 20 miles-per-hour crash test requirements set by federal law at the time. NHTSA had been proposing standards of fuel system integrity since 1969, when it required cars to pass crash tests at 20 mph. Ford's cars had met these standards, but the $3.5 million penalty that was eventually upheld on appeal was not because of explicit regulatory non-compliance. NHTSA conducted an investigation that compared the crash performance of various brands of comparable cars, and the Ford Pinto was shown to burst into flames much more frequently than the other cars.

In 1981, the California appellate court upheld the initial ruling, stating that "Ford's conduct constituted 'conscious disregard' of the probability of injury to members of the consuming public" (Grimshaw 1981). The court did not acknowledge Ford's regulatory compliance as a sufficient safety measure, and instead focused on the fact that if Ford had made the design choice to place the fuel tank over the rear axle, it would not have caught fire even in a crash of 50 mph. This landmark decision is a prime example of how litigation contributes to design safety expectations that are placed on manufacturers, even after preventive regulations have been made and met. In effect, manufacturers are held to a higher level of responsibility for safety design than can be spelled out by regulations, and the court system helps to tighten those standards.

The complex nature of how manufacturers make design decisions that trade off safety against cost can be seen in Ford's cost benefit analysis, revealed in the court's documents (Figure 4).

**Benefits**

*Savings: 180* burn deaths, *180* serious burn injuries, 2100 burned vehicles
*Unit Cost: $200,000* per death, $67,000 per injury, $700 per vehicle
*Total Benefit: 180 x ($200,000) + 180 x* ($67,000) + 2100 *x* ($700) = **$49.5 Million**

**Costs**

*Sales:* 11 million cars, 1.5 million light trucks
*Unit Cost: $5.08* per car, *$5.08* per truck
*Total Cost: 11,000,000* x ($3.96) + 1,500,000 *x* ($3.96) = **$49.5 Million**
Therefore, if the cost to replace the fuel tank was $3.96 per vehicle, the costs and benefits would equal each other out (all other things remaining the same).

**Figure 4. The Break Even Point of the Cost/Benefit Analysis (Grimshaw 1981)**

The most controversial item in the break-even calculations was the valuation of human life at $200,000 per death. Using these approximations, Ford did not find it worthwhile to alter fuel tank placement. Ford had followed the precedent of risk/utility standards for manufacturer's liability established by Judge Learned Hand in *United States v. Carroll Towing* when making this decision in 1947, and NHTSA was the organization that provided the automobile industry with the $200,725 per life estimate used in Ford's 1972 risk/benefit analysis (Carroll Towing 1947). Nevertheless, the court ruled that the formula provided by Judged Learned Hand was for specific accidents, while Ford's analysis involved use of the product as a whole. Ford had made a legal decision in not making design changes that would have made the Pinto safer, but this did not make the decision ethical in the jury's perspective (Leggett 1999). As ethically questionable as it may be to put monetary value on human death or injury, manufacturers have to make those choices every day, and it is impractical to pursue safety features while altogether disregarding the price for doing so.

While I do not attempt to provide a formula for how manufacturers can reliably make design decisions to ensure that they meet their responsibilities, it is clear from this discussion that no one factor – economic analysis, public opinion, regulations, or even ethical standards – can on its own lead to safe vehicle design. If autonomous vehicles are mass produced, society will need to formalize the standards by which manufacturers decide a far larger number of design decisions with safety tradeoffs. These decisions will include safety tradeoffs normally left to drivers, from when to change lanes to when to brake abruptly. Regulations and litigation will continue to play an important role in defining those standards.

## Conclusion

In examining the structure for automobile liability that is currently in place, it is evident that there are many parties who are responsible for directly making safety related decisions, as well as those who hold accountable those who are responsible. The parties identified here included pedestrians, automobile drivers, and manufacturers, made accountable through regulations, litigation, and public opinion. For each of these parties, responsibility for safe vehicle design and control is shaped by an emerging social consensus. Safety expectations are socially constructed by an understanding of how things ought to be, and not by a fundamental analysis of how things are.

That said, many real-time decisions that ensure or endanger a person's safety will need to be formalized and systematized in the technological design of the autonomous vehicle. This will lead to major shifts of responsibility, incorporating many critical decisions regarding design of the car, real-time decisions normally left up to the driver, and systematized information processing and storage decisions at a higher level. The challenge will then be to adapt our current liability system to prevent dangerous vehicle design decisions and to adequately compensate victims when something goes wrong.

# 4 Preparing for Autonomous Vehicles

The autonomous vehicle has made great strides toward furthering safe and convenient driving. However, integrating human sensing and perception into the mechanics of the car will also introduce new risks that require new regulation. In this chapter, I will analyze each of the parties potentially responsible for automobile accidents described in the previous chapter, focusing on how their responsibilities will change with the transition from the traditional car to the self-driving car. The parties to be analyzed are pedestrians, drivers, and manufacturers.

## 4.1 Pedestrian Responses to the Self-Driving Car

Pedestrians will continue to share the roads with vehicles, even with a transition to self-driving cars. The main challenges will come from the fact that the rules that pedestrians follow in practice are socially determined and not followed exactly as written into traffic rules. One example of this is jaywalking in New York City. While it is technically illegal in New York, jaywalking is an almost instinctive practice by the city's residents, and law enforcement officers turn a blind eye a vast majority of the time (Martin 2014).

The autonomous vehicle will have difficulty safely responding to these deviations of human behavior from what is formally required by law. In the case of the fatal Uber self-driving car and pedestrian accident discussed in Chapter 1, the technology was designed and fully expected to stop the car in time to avoid colliding with the pedestrian in a situation similar to the one at the time of the crash (Roberts 2018). While the details of the crash are still being investigated, the accident raises questions as to whether the self-driving car ought to be responsible for responding to abrupt pedestrian street-crossings, or whether pedestrians ought to change their behavior to accommodate the self-driving car. The Wall Street Journal reports that the fatal pedestrian crash in Arizona has spurred studies on whether people will change their patterns of walking out in front of a vehicle if they know that an oncoming vehicle is likely to be a self-driving car (Roberts 2018). Studies that monitor such shifts in human behavior will be critical for regulations addressing the self-driving car, and the car will need to be designed taking these behaviors into account.

According to the Governors Highway Safety Association, the number of pedestrian fatalities rose 27% from 2007 to 2016, much faster than the growth of general traffic fatalities during the same time period. Russ Martin, the director of government relations for the organization, attributes this increase to a general rise in the number of people walking and driving, as well as increased cellphone use while both walking and driving (Roberts 2018). Behaviors such as the trend towards ubiquitous cellphone use are not likely to change in the face of the self-driving car. Regulators will need to be proactive in instituting design tests that measure the safety of pedestrian crossings, similar to the fuel-tank crash safety tests described in Chapter 3. Furthermore, the self-driving car will need to utilize deep-learning algorithms to be trained on human behavior patterns, instead of hard-coded algorithms that strictly follow written law.

## 4.2 Driver Responsibilities in the Transition to the Self-Driving Car

While the end goal of the self-driving car may be to eliminate the role of the driver altogether, there are many steps in the transition that will still involve the human driver. The increasing levels of automation are likely to add new driver-related risks, such as an incentive for distracted driving.

Many types of influences currently contribute to distracted driving, including drinking, texting, and dozing off, among others. As more functions of the vehicle become automated, including lane changes and parking, drivers will tend to pay less attention to those activities, possibly becoming overly reliant on these functions, or neglectful of potential shortcomings in the technology. In light of this trend, manufacturers will want the law to hold drivers accountable for paying attention even when using these augmenting functions. However, as shown by both the Tesla Model S and the self-driving Uber accidents in Chapter 1, the human driver is a poor back up safety measure for when autonomous systems fail.

Car manufacturers have long been held to regulatory standards and crash tests, but as mentioned in the previous chapter, simply meeting these standards have not exempted them from liability. Likewise, having the driver agree to a waiver has not exempted manufacturers from responsibility for safe design. Even in cases where a design defect is not the direct cause of an accident, if the defect further endangers a person and increases damages during an accident, the manufacturer is still liable. Thus, precedents for holding manufacturers accountable without needing to articulate every possible scenario are already in place.

Car manufacturers should also be held responsible for clearly communicating the limitations of their technologies, as well as building in safety functions that account for increased distracted driving. To respond properly, drivers will have to be aware of critical limitations in processing power in the central control programs. From a technological standpoint, default safety systems such as automated braking and anti-collision measures should be counteracted with programs that allow the driver to override defaults. The tradeoffs between human and mechanical solutions will have to be assessed for every available automated feature when making design choices, keeping in mind human tendencies to over-rely on automated driving functions. Post-accident litigation will play a particularly critical role in allocating responsibility between driver and manufacturer design, because of the complex nature of determining whether attentiveness and response could be reasonably expected from the driver when a car does not perform as it ideally should.

## 4.3 Manufacturer Responsibility to Design a Safe Self-Driving Car

Established product liability laws will be directly applicable to many situations involving component and systems failures, but there may also be new sources of liability arising from the mechatronic, autonomous mechatronic, and cross-linked mechatronic levels.

***Mechatronic system***

For the mechatronic system, the existing product liability laws and trends discussed in Chapter 3 will have direct applicability. Manufacturers will continue to bear the responsibility for reliable performance of sensor technologies, actuating motors, and other mechanical components. Rigorous tests analogous to fuel tank crash tests will be useful for ensuring the quality and placement of these new components. Litigation will play a key role in expanding the scope of responsibility beyond that which is explicitly mentioned by law, as in the Ford Pinto case (Grimshaw 1981).

***Autonomous mechatronic system***

The autonomous mechatronic system presents more novel risks, by virtue of incorporating human driver decisions into the design of the vehicle. The shift in how the law ought to treat manufacturers is far from a simple transfer of driver responsibilities to car manufacturers, however. Previously, courts have relied on retrospective analysis of case-by-case decisions that people involved in the accident made, as well as the condition of the car both when it left the manufacturer's facilities and at the time of the accident. Now, those decisions will be algorithmically determined based on programs written into the car before the accident. This means that cases will need to take account of manufacturers' foresight, instead of the reasonableness of any real-time decision.

Two sources of possible breakdown arise from the design of autonomous driving algorithms: design failure errors and design limitation errors. Traditionally, vehicle-related design decisions have focused on increasing the reliability and durability of the car's components. In the software programs of the self-driving car, accurate and reliable real-time data management in the car's systems will be necessary. Failure to process the data as programmed is a well-understood problem, analogous to component failures.

On the other hand, design limitation errors present new challenges. These range from the challenges of achieving optimal driving performance to making ethical driving choices. Especially in the early stages of development of the autonomous vehicle, design choices regarding which processing power capacities are suitable for different driving scenarios will be critical. At each step along the process of developing a fully autonomous vehicle, car manufacturers will offer additional automated functions, such as parking, switching lanes, and cruise control. In each case, the human driver's expectations for what the car can be relied on to do on its own will be key to safely navigating the car. Regulators will thus play an important role in defining the boundaries of what each automated function can safely complete in order for manufacturers to offer it as a package in their cars and market it as an automated function.

For each added algorithm and data processing component, manufacturers will also need to do a cost-benefit analysis of whether it is worth the investment to add a safety-enhancing feature. The standard for this choice will have to be determined by regulators and possibly further refined through lawsuits. Can manufacturers forego the addition of a certain safety feature as long as the existing feature is less risky than in the traditional car and is thus an improvement on the current safety standard? Is there a monetary valuation of human life that can be used to determine whether the marginal increase in safety that will result from investment of corporate funds would be worth it? Is it immoral for a company to profit if that money could have otherwise gone into adding an extra feature that could have prevented an accident and saved a life? The answers to these questions will directly determine the degree of safety in the self-driving car.

Finally, manufacturers will need to formalize in their algorithms many ethical choices normally left up to the driver. Abruptly braking the car to avoid hitting a person or an animal often endangers the passengers in the car, and such a choice has been in the hands of the driver. Self-driving car manufacturers will have to formalize and implement these choices into their algorithms, sometimes making life-death decisions for more extreme scenarios. These choices call for a careful analysis of ethical dilemmas in a variety of hypothetical driving scenarios, possibly formulated like the famous Trolley Problem, in which a person needs to decide whether to allow a train to kill five people on its track or to actively switch the track and kill only one person (Goldhill 2018).

### ***Cross-Linked Mechatronic System***

Deep-learning algorithms trained on mass aggregated data can allow for self-driving cars to learn human behavior patterns. This is useful for making the self-driving car respond to socially constructed driving patterns such as speed limits instead of being restricted to follow written traffic rules. Training on real-world data raises its own legal issues, however.

As Cathy O'Neil argues in *Weapons of Math Destruction*, machine learning algorithms are very sensitive to biases in the data that they are trained on. Furthermore, typically there are not enough feedback tests to improve the algorithms and correct for those biases (O'Neil 2017). This is a realistic danger for the self-driving car as well. Given the proprietary nature of autonomous driving algorithms and the fact that they are software programs, design errors in manufacturers' algorithms will be much more difficult to scrutinize than design flaws in the mechanical components of a car. Regulators will need to take care to investigate and test these algorithms, possibly by submitting them to a team of government-appointed experts, as well as by setting up feedback mechanisms that can indicate whether the deep-learning algorithms are learning desirable driving behaviors.

Given that data collection, transfer, processing, and storage play critical roles in the autonomous vehicle, data security will also emerge as a new category of concern. Three aspects should be highlighted. First, in 2012, the Supreme Court decided in *United States v. Jones* that installing a GPS tracking device on a vehicle and tracking its location constitutes a search under the Fourth Amendment, and cannot be conducted without a warrant (US v. Jones 2012). Tracking the location of a person's car is considered a serious breach of privacy, because of the large amounts of information and detail about a person's life it can reveal. Location information alone may point to a person's religious affiliations, sexual orientation, career, personal network, and other details central to a person's identity. This information could be garnered from whether the person goes to church, visits friends' homes, drives to work, etc. The autonomous vehicle will generate ubiquitous and constant location information from its passengers, requiring especially stringent security measures for any shared economy car or cloud-based data gathering system.

Second, there has been a trend toward increasing integration of people's social media accounts and internet-based profiles with the car in order to provide more automated functions. Today's cars promote automated music selection based on the driver's preferences, bluetooth messaging to social media friends and phone contacts, and even automated ordering of the driver's go-to coffee when driving towards his or her favorite coffee chain. While the convenience and efficiency of such features can be enticing to many consumers, storage of and shared access to such rich information pose serious risks to privacy if it falls into the wrong hands. Even if the information were not to lead to identity theft, such aggregation of consumer preference data could open doors for targeted advertising campaigns, political campaigns, or psychological manipulation over which the consumer has limited control. It will be desirable for regulators to establish robust measures of anonymity, privacy, and security that need to be followed by autonomous vehicle manufacturers, service providers, and cloud-based companies.

Finally, data generated by the autonomous vehicle will not only be vulnerable to theft or misuse, but also to hijacking of the vehicle's data systems. Especially if the car is downloading information from a cloud system in real-time, hostile parties will have numerous access points for cutting off the flow of critical information, distorting the data the car receives and executes, or even taking over control of the car itself. Centralized data systems could become a prime target for terrorists who want to jeopardize the lives of masses of riders, while individual cars will also be susceptible to hackers and system failures.

## 4.4 Exogenous Forces That Impact Vehicle Safety Design

In addition to the three parties responsible for vehicle liability, there are macroeconomic forces that affect the design of self-driving cars and ought to be regulated.

The international economy is an influential force on vehicle design. Increasing internationalization of goods, including cars as well as software and data processing, has allowed for developments overseas to influence the design of cars in the United States both directly and indirectly. Trends in technological development and data security practices overseas are likely to be especially pertinent domestically. China has been steadily increasing its lead in manufacturing cheap and reliable semiconductors and technological components, while European countries have been leading development in robust software programs. Levels of competition between manufacturers from these countries that import cars into the U.S., as well as the adoption of their ideas into domestic vehicle designs, would have direct impacts on safety. Indirectly, greater internationalization of ideas and sharing across countries can help spur innovative solutions for smarter and more efficient driving decisions, and the opportunity for U.S. companies to compete across more markets could also encourage them to seek a branding edge on grounds of more reliable safety.

Another exogenous force is the role of insurance companies in mitigating risk surrounding autonomous vehicles. In the past, drivers have been held to differing levels of premiums and requirements based on characteristics such as age, driving record, and the age and value of the car. However, with the elimination of the driver and normalization of driving ability across all cars, the risk of riding in the car would become almost entirely situation and technology-dependent. Thus, it is possible that insurance companies would offer manufacturer insurance instead of vehicle and driver insurance to cover damage to the car and injured parties. Car manufacturers could in turn transfer the cost of insurance to their consumers by charging premiums for higher risk situations, such as choosing to travel in a higher-speed autonomous vehicle, traveling in harsher conditions such as during bad weather, or traveling on more difficult paths such as busy pedestrian routes or cliff-side driveways.

The role of insurance companies may even decline altogether. The benefit and profitability of insurance has come from the law of large numbers, because insurance providers can alleviate risk for individuals while still assuming less risk from the population as a whole than the total premiums they are collecting. Manufacturing companies already deal in large numbers, and may be able to self-insure for large numbers of consumers, building those costs into the price of the vehicle. If insurance companies do exist for autonomous vehicle manufacturers, it may be in the form of more general financial liability insurance to cover lawsuits and business loans. However, even if insurance companies are able to find a significant method of mitigating risk for the manufacturing companies, the desirability of that effect would be questionable. If indeed accident costs are one of the major incentives for manufacturing companies to pursue safe design choices, it may well be best to make the companies pay for damage ex post rather than be fully insured against financial harm.

## Conclusion

Integrating human decisions and processing into the machine presents novel risks for pedestrians, drivers, and manufacturers that need to be regulated. Effective regulations will take into account the fact that standards of driving and walking on the streets are a product of social consensus, and not necessarily through a strict adherence to formal law. Trends in jaywalking, distracted driving, and other human behaviors will need to be closely monitored to appropriately regulate the safety of self-driving cars. Macroeconomic forces such as international commerce trends, and the role of service industries such as insurance companies will also continue to play a role in shaping standards of vehicle safety and liability. All of these forces can be controlled through regulatory measures.

More generally, formal standards of ethics and for cost-benefit analysis need to be established for manufacturers to program into their self-driving algorithms. Regulators can enforce a battery of scenario-based tests that will formalize safe driving behaviors and design decisions. However, not all safety expectations will be completely discoverable beforehand, both because of the breadth of issues that arise from the introduction of disruptive technology and the variable, situation-dependent cases that cannot be foreseen exhaustively. These will need to be remedied by ex post facto litigation.

# 5 Conclusion

The development of self-driving cars represents the first of potentially many technologies that will significantly automate a critical decision-making function in people's everyday lives. In order to emulate the human mind in the machine, many technological solutions are being developed. These can be subdivided into the mechatronic, autonomous mechatronic, and cross-linked mechatronic systems as described in Chapter 2, and each level addresses an aspect of how the human mind works when driving. As we have seen, automated behavior can surpass human capability in various ways. Sensors at the mechatronic level have the capacity to gather information beyond the limits of the five human senses, more reliably and more accurately. Computer algorithms that process the sensory input are not prone to distraction, inexperience, or other errors in human judgment. Even patterns of human behavior and processing of social information can be enhanced using deep-learning algorithms on mass aggregated data at the cross-linked mechatronic level. The massive opportunities for economic profitability have led to established and new companies such as General Motors, Tesla, and Uber into a race to become an industry leader. Profit incentives for these companies come from the prospect of potentially huge cost reductions in trucking, taxiing, busing, and other services that currently require driver wages. At the same time, the potential for accident reduction has led regulators to allow relatively free-rein to development and beta-testing on urban streets.

The recent fatal pedestrian accident involving an Uber self-driving car discussed in Chapter 1 brought to public attention that introducing the self-driving car to the roads also comes with many new sources of harm that will require new regulation. Chapter 3 analyzed the parties that are currently responsible for vehicle liabilities: pedestrians, drivers, and car manufacturers. Each of these parties is held to standards of responsibility determined by social consensus, and not only to a strict adherence to written laws. Reasonable navigation within these socially constructed bounds has been left up to the driver, with regulators formalizing patterns of safety practices as they arose from behavioral change, as well as addressing standards for safe design as technological pitfalls came to their attention.

When putting the human mind into the machine, however, these human behaviors will need to be fed into programmed algorithms, and it will then become much more difficult for the public to monitor manufacturer's design choices. The unique risks that arise for pedestrians, drivers, and car manufacturers were addressed in Chapter 4. Many of these arose from a potential mismatch between social behaviors and the car's formal mode of operation based on formal law. For pedestrians, this includes the car's ability to safely avoid hitting jaywalkers; for drivers, there is the issue of distracted driving; and for car manufacturers, there are risks from design decisions made at each of the three technological levels in the self-driving car. Some risks, like sensor failures, are relatively well accounted for by current legislation. The areas that will be a particular focus of new regulations are the ones that incorporate driver decisions into the design of the car, namely at the autonomous mechatronic and cross-linked mechatronic levels. When it comes to programming how cars should decide ethical dilemmas, or making design decisions that have safety tradeoffs, manufacturers ideally need to have regulatory standards by which to make their decisions.

Just as the autonomous vehicle is an engineered product which incorporates many design decisions, so standards of responsibility too are a product of wide social consensus. Regulators play a critical role in shaping the standards of product design as well as responsibility, and litigation allows situations to be reviewed on a case-by-case basis, enforcing social standards and compensating victims.

There are two standards that particularly need to be established for regulating the self-driving car. First, is the systematization of driver choices so that they can be programmed into the automated system. Drivers continually make decisions that reflect their personal values – whether in speeding to prioritize time, taking the long route to enjoy the scenery, running over a squirrel to avoid a more serious crash, or risking whiplash by braking suddenly to avoid running over an animal. Regulators, and not car manufacturers, ought to have final say in what these priorities are.

Second, numerous factors must be weighed against safety by company managers, and regulators should make clear when a tradeoff ought not to be made. At what point is it acceptable to forego an added safety measure because it is too expensive? Is automating a function good as long as it outperforms human performance, or are there other benchmarks of adequacy that should be used? More broadly, what standards do manufacturers need to meet before deploying automatic features? These are just a few of the critical questions that arise from the act of automating human decisions in a realm that directly oversees physical safety, and regulators need to answer those questions.

This thesis began by analyzing the technological design of the self-driving car and made its way to the need to define standards that will hold parties responsible for new risks and new liabilities. Regulations will play an invaluable role in setting those standards, and will have the effect of working in the upstream direction, shaping social expectations of responsibility and influencing the design of the car. That said, social behaviors will continue to deviate from these standards, and it will fall to litigation to enforce and interpret the standards in the context of real-life situations.

The new regulatory and legal considerations presented by the self-driving car pave the way for future technological developments seeking to machinize a part of human consciousness and social behavior. The standards of decision-making and sources of liability identified here will only become more pertinent with the increasing adoption of artificial intelligence robots and programs. Regulators and litigators will play a key and integral role in guiding us in the right directions.

# Bibliography

## Literary Citations


Aho, Karen. "Here's How Many Car Accidents You'll Have." *Fox Business*, Fox Business, 12 Jan. 2016, www.foxbusiness.com/features/heres-how-many-car-accidents-youll-have.

Anderson, James M., Nidhi Kalra, Karlyn Stanley, Paul Sorensen, Constantine Samaras and Tobi A. Oluwatola. Autonomous Vehicle Technology: A Guide for Policymakers. Santa Monica, CA: RAND Corporation, 2016. https://www.rand.org/pubs/research_reports/RR443-2.html.

Boucher, C. and J.C. Noyer, "A hybrid particle approach for GNSS applications with partial GPS outages", *IEEE Trans. Instrum. Meas.*, vol. 59, no. 3, pp. 498-505, Mar. 2010.

Boudette, Neal E. "U.S. Traffic Deaths Rise for a Second Straight Year." The New York Times, The New York Times, 15 Feb. 2017, www.nytimes.com/2017/02/15/business/highway-traffic-safety.html.

Brechtel, Sebastian, et al. "Probabilistic MDP-Behavior Planning for Cars." *2011 14th International IEEE Conference on Intelligent Transportation Systems (ITSC)*, 2011, doi:10.1109/itsc.2011.6082928.

Brynjolfsson, Erik, and Andrew McAfee. *The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies*. 1st ed., W.W. Norton & Company, 2014.

Dalal, N. and B. Triggs. "Histograms of oriented gradients for human detection". *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. CVPR 2005. Vol. 1, pp. 886–893. IEEE. DOI: 10.1109/CVPR.2005.177. June 2005.

Donde, Jay. "Self-Driving Cars Will Kill People. Who Decides Who Dies?" *Wired*, Conde Nast, 20 Sept. 2017, www.wired.com/story/self-driving-cars-will-kill-people-who-decides-who-dies/.

Eastman, J.W. *The American Automobile Industry and the Development of Automotive Safety, 1900-1966*. University Press of America, 1984.

Fischer, P., Dosovitskiy, A., Ilg, E., Häusser, et al. *Flownet: Learning Optical Flow with Convolutional Networks*. arXiv preprint arXiv:1504.06852. 2015.

Fossen, Thor I. *Sensing and Control for Autonomous Vehicles.* Springer International Publishing, 2007.

Garcia, Uriel J., and Karina Bland. "Fatal Uber Crash with Pedestrian Likely 'Unavoidable,' Police Chief Says." *USA Today*, Gannett Satellite Information Network, 21 Mar. 2018, www.usatoday.com/story/tech/2018/03/20/tempe-police-chief-fatal-uber-crash-pedestrian-likely-unavoidable/444291002/.

Giebel, J., et al. "A Bayesian Framework for Multi-Cue 3D Object Tracking." *Lecture Notes in Computer Science Computer Vision - ECCV 2004*, 2004, pp. 241–252., doi:10.1007/978-3-540-24673-2_20.

Goldhill, Olivia. "Philosophers Are Building Ethical Algorithms to Help Control Self-Driving Cars." *Quartz*, Quartz, 11 Feb. 2018, qz.com/1204395/self-driving-cars-trolley-problem-philosophers-are-building-ethical-algorithms-to-solve-the-problem/.

Graham, John. *Product Liability and Motor Vehicle Safety*. The Brookings Institution, 1991.

Gusfield, Joseph R. *Contested Meanings: the Construction of Alcohol Problems*. University of Wisconsin Press, 1996.

Harris, William. "How Car Suspensions Work." *HowStuffWorks*, HowStuffWorks, 11 May 2005, auto.howstuffworks.com/car-suspension1.htm.

Harris, Mark. "Exclusive: Arizona Governor and Uber Kept Self-driving Program Secret, Emails Reveal." The Guardian. March 28, 2018. Accessed March 30, 2018. https://www.theguardian.com/technology/2018/mar/28/uber-arizona-secret-self-driving-program-governor-doug-ducey.

Hirschmuller, H. "Stereo processing by semiglobal matching and mutual information", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(2), pp. 328–341. 2008.

Hoflinger, F., J. Müller, R. Zhang, L. M. Reindl, W. Burgard, "A wireless micro inertial measurement unit (IMU)", *IEEE Trans. Instrum. Meas.*, vol. 62, no. 9, pp. 2583-2595, Sep. 2013.

Huddy, Xenophon P. "The Law of Automobiles." 1906, Albany, NY.

Hunt, P., Konar, M., et al. "ZooKeeper: Wait-free coordination for internet-scale systems". *Proceedings of the Usenix Annual Technical Conference*, 2010.

Jasanoff, Sheila. *The Ethics of Invention: Technology and the Human Future*. W.W. Norton & Company, 2016.

Krizhevsky, A., Sutskever, I., and G.E. Hinton. "Imagenet classification with deep convolutional neural networks". *Advances in Neural Information Processing Systems,* pp. 1097−1105. 2012.

Kurtz, Rachel Abrams and Annalyn. "Joshua Brown, Who Died in Self-Driving Accident, Tested Limits of His Tesla." *The New York Times*, The New York Times, 1 July 2016, www.nytimes.com/2016/07/02/business/joshua-brown-technology-enthusiast-tested-the-limits-of-his-tesla.html.

Lee, S., and J.B. Song, "Robust mobile robot localization using optical flow sensors and encoders", *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, vol. 1, pp. 1039-1044, Apr./May 2004.

Leggett, Christopher. "THE FORD PINTO CASE: THE VALUATION OF LIFE AS IT APPLIES TO THE NEGLIGENCE-EFFICIENCY ARGUMENT." *WFU*, 1999, users.wfu.edu/palmitar/Law&Valuation/Papers/1999/Leggett-pinto.html.

Li, H., Ghodsi, A., Zaharia, M., Shenker, S., and I. Stoica, I. "Tachyon: Reliable, memory speed storage for cluster computing frameworks". *Proceedings of the ACM Symposium on Cloud Computing,* 2014, pp. 1-15. ACM. DOI: 10.1145/2670979.2670985.

Lins, Romulo Goncalves, et al. "Velocity Estimation for Autonomous Vehicles Based on Image Analysis." *IEEE Transactions on Instrumentation and Measurement*, vol. 65, no. 1, 2016, pp. 96–103., doi:10.1109/tim.2015.2477160.

Liu-Henke, Xiaobo, et al. “System Architecture of a Full Active Autonomous Electric Vehicle.” *2017 IEEE Transportation Electrification Conference and Expo (ITEC)*, 2017, doi:10.1109/itec.2017.7993246.

Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.Y., and A.C. Berg. “SSD: Single shot multibox detector”. *European Conference on Computer Vision*, October 2016, pp. 21−37. Springer International Publishing. DOI: 10.1007/978-3-319-46448-0_2.

Long, J., Shelhamer, E., et al. “Fully convolutional networks for semantic segmentation”. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* , pp. 3431−3440. DOI: 10.1109/CVPR.2015.7298965. 2015.

Lu, Jinhao, and Chi Dong. “Research of Shortest Path Algorithm Based on the Data Structure.” *2012 IEEE International Conference on Computer Science and Automation Engineering*, 2012, doi:10.1109/icsess.2012.6269416.

Luo, Wenjie, et al. “Efficient Deep Learning for Stereo Matching.” *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 5695–5703., doi:10.1109/cvpr.2016.614.

Mackay, Murray. *Liability, Safety, and Innovation in the Automotive Industry*. The Brookings Institution, 1991.

Martin, Adam. “The Savvy New Yorker's Guide to Jaywalking.” *Daily Intelligencer*, 31 Jan. 2014, nymag.com/daily/intelligencer/2014/01/savvy-new-yorkers-guide-to-jaywalking.html.

Mays, Kelsey. “Here's What the Average New Car Costs | News from Cars.com.” *Cars.com*, 11 Apr. 2017, www.cars.com/articles/heres-what-the-average-new-car-costs-1420694975814/.

*Mcity*, mcity.umich.edu/our-vision/.

Mcnaughton, M. “Parallel algorithms for real-time motion planning.” 2011, Doctoral Dissertation. Robotics Institute, Carnegie Mellon University.

Menze, M. and A. Geiger. “Object scene ow for autonomous vehicles”, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition,* pp. 3061–3070, DOI: 10.1109/CVPR.2015.7298925. 2015.

Mitchell, Russ. “Lidar Costs $75,000 per Car. If the Price Doesn't Drop to a Few Hundred Bucks, Driverless Cars Won't Go Mass Market.” *Los Angeles Times*, Los Angeles Times, 11 Dec. 2017, www.latimes.com/business/la-fi-hy-ouster-lidar-20171211-htmlstory.html.

Mnih, V., Badia, A. P., Mirza, et al. “Asynchronous methods for deep reinforcement learning”. *International Conference on Machine Learning*, 2016, pp. 1928–1937.

Mnih, V., Kavukcuoglu, K., et al. “Playing Atari with deep reinforcement learning”. 2013.

Nice, Karim. “How Car Steering Works.” *HowStuffWorks*, HowStuffWorks, 31 May 2001, auto.howstuffworks.com/steering3.htm.

Nichols, Hannah. “The Top 10 Leading Causes of Death in the United States.” *Medical News Today*, MediLexicon International, 23 Feb. 2017, www.medicalnewstoday.com/articles/282929.php.

Nixon, Richard M. “Changing Rules of Liability in Automobile Accident Litigation.” *Law and Contemporary Problems*, vol. 3, no. 4, 1936, p. 476., doi:10.2307/1189341.

Norton, Peter. *Fighting Traffic: The Dawn of the Motor Age in the American City*. Massachusetts Institute of Technology, 2008.

O'Neil, Cathy. *Weapons of Math Destruction*. Penguin Books, 2017.

Paden, B., Cap, M., Yong, S. Z., Yershow, D., and E. Frazzolo. "A survey of motion planning and control techniques for self-driving urban vehicles". *IEEE Transactions on Intelligent Vehicles* 1(1), 2016, pp. 33–55., DOI: 10.1109/TIV.2016.2578706.

Pfennigbauer, Martin, and Andreas Ullrich. "Multi-Wavelength Airborne Laser Scanning." *IMLF* , Feb. 2011, www.riegl.com/uploads/tx_pxpriegldownloads/Paper_ILMF_2011_RIEGL_Multiwavelength_ALS.pdf.

Quigley, M., Gerkey, B., et al. "ROS: An open-source robot operating system". *Proceedings of the Open-Source Software Workshop International Conference on Robotics and Automation*, 2009.

Ranjan, Anurag, and Michael J. Black. "Optical Flow Estimation Using a Spatial Pyramid Network." *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, doi:10.1109/cvpr.2017.291.

Redmon, J., Divvala, S., Girshick, R., and Farhadi, A. "You only look once: Uni ed, real-time object detection". *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition,* pp. 779−788. DOI: 10.1109/CVPR.2016.91. 2016.

Redmon, J. and A. Farhadi, A. *YOLO9000: Better, Faster, Stronger*. 2016.

Ren, S., He, K., Girshick, R., and J. Sun. "Faster r-cnn: Towards real-time object detection with region proposal networks". *Advances in Neural Information Processing Systems,* pp. 91−99. 2015.

Roberts, Adrienne. "Uber Crash Highlights Growing Safety Concern: Pedestrians." *The Wall Street Journal*, Dow Jones & Company, 23 Mar. 2018, www.wsj.com/articles/uber-crash-highlights-growing-safety-concern-pedestrians-1521810000?mg=prod/accounts-wsj.

Rudolph, Gert, and Uwe Voelske. "Three Sensor Types Drive Autonomous Vehicles." *Sensors Magazine*, 10 Nov. 2017, www.sensorsmag.com/components/three-sensor-types-drive-autonomous-vehicles.

Schweber, Bill. "It's Easier Than Ever to Use MEMS Accelerometers for the Industrial IoT." *Implementing MEMS Accelerometers*, Digit-Key Electronics, 21 June 2017, www.digikey.com/en/articles/techzone/2017/jun/its-easier-than-ever-mems-accelerometers-industrial-iot.

Schweber, Bill. "The Autonomous Car: A Diverse Array of Sensors Drives Navigation, Driving, and Performance." *Mouser Electronics - Electronic Components Distributor*, 2018, www.mouser.com/applications/autonomous-car-sensors-drive-performance/.

Singh, S. "Critical Reasons for Crashes Investigated in the National Motor Vehicle Crash Causation Survey." *Traffic Safety Facts Crash*, Feb. 2015.

Stone, J.E., Gohara, D., and G. Shi. "OpenCL: A parallel programming standard for heterogeneous computing systems." *Computing in Science & Engineering*, 2010, 12(3), pp. 66–73. DOI: 10.1109/MCSE.2010.69.

"The Grand Challenge." *Defense Advanced Research Projects Agency*, www.darpa.mil/about-us/timeline/-grand-challenge-for-autonomous-vehicles.

White, T. "Hadoop: The Definitive Guide." 2012. O'Reilly Media, Inc.

Xiang, Yu, et al. "Learning to Track: Online Multi-Object Tracking by Decision Making." *2015 IEEE International Conference on Computer Vision (ICCV)*, 2015, doi:10.1109/iccv.2015.534.

Zaharia, M., Chowdhury, M., et al. "Spark: Cluster computing with working sets." *HotCloud*, 2010, p. 95.

Zhao, H., Shi, J., Qi, X., Wang, X., and J. Jia. *Pyramid Scene Parsing Network*. arXiv preprint arXiv:1612.01105. 2016.

## Court Case Citations

Grimshaw v. Ford Motor Co., 119 Cal. App. 3d 757 (Cal. Ct. App. 1981).

Suvada v. White Motor Co. Supreme Court of Illinois. 20 May 1965.

United States v. Carroll Towing Co., 160 F.2d 482 (2d Cir. N.Y. Mar. 17, 1947).