Privacy Concerns in the World of Autonomous Vehicles

STS Research Paper
Presented to the Faculty of the
School of Engineering and Applied Science
University of Virginia

By

Matthew Orlowsky

4/12/2020

On my honor as a University student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments.

Signed: Matthew Orlowsky

Approved: _____ Date April 17, 2020
Rider Foley, Department of Engineering and Society

Introduction

One of the most exciting innovations today is the promise of autonomous vehicles. Autonomous vehicles (AVs) are expected to occupy 25% of the global market by 2040 (Taeihagh, 2019). The automotive industries working with this future technology propose that the autonomous vehicles will achieve the remarkable by transporting passengers to their destinations with minimal human intervention. AVs have the potential for an abundance of benefits including improved safety, shorter commutes, less stress, fewer environmental impacts, and freedom for immobile persons. It is anticipated that the passenger will enjoy access to Wi-Fi, encounter less risk of accidents, and benefit from the ability to avoid traffic and road construction delays. However, all of these benefits can only be realized if the algorithms that navigate the autonomous vehicles are fed vast amounts of data via GPS, voice-recognition, cameras, and other sensors. AV technology will require extensive communication and data transfer between the car, the infrastructure, and surrounding environment. The AV will need constant interaction with location tracking satellites and will require sensors such as radar and LIDAR (Light Detection and Ranging) to create and maintain internal maps, and calculate and derive the routes.

One thing that must be considered is what will happen with all of this information. The data includes not only locations and preferred routes but also shopping habits, on-line activity, and even voice recordings. What if hackers were to get a hold of this data? Will robbers know when you are away from home? What possible things will the security agency be able to do with the collected data? Although autonomous vehicles promise many benefits, what are the potential privacy concerns that exist with this technology and can actions be taken now to alleviate these risks? The results of the research can be helpful for identifying areas for needed action and also

to raise awareness to the importance of cybersecurity. Threats and dangers communicated globally may motivate others to become involved in the development of design solutions and mitigations. Investigation of current vulnerabilities associated with autonomous vehicle privacy, security, and personal autonomy will aid in the consideration of potential mitigations for user's privacy protection. The hope is that our privacy will be protected and that the car windows will be the only way that the outside world will be able to "see" us.

Autonomous Vehicle Connections with Society

There are wide-spread human and social dimensions of autonomous vehicle technology. This technology will have an enormous positive impact for many people including the handicapped, the young, and the elderly, who in the past, were unable to operate a car and are now promised greater freedom to travel. "Introducing the technology will potentially bring changes across the entire sphere of mobility, impacting many levels of society. At the same time, triggering a fundamental transformation in the way we get around" (Fraedrich, 2016, p. 622). Although the possibilities are exciting, there is growing public concern regarding data privacy. In light of the July 2019 Capital One data breach where data from 106 million people was compromised, state and municipal legislature are considering new regulations regarding what data will be collected and how it will be used (Eliot, 2019).

The sonar, radar, cameras, and wireless communication necessary for a safe and efficient ride come with numerous security dangers. User activity data has the potential to be used for manipulation and profiled advertising, and surveillance data could be used for legal and illegal tracking. The vehicle itself would be a repository of personal information that would be at risk

for hacking, burglary, and misuse (Glancy, 2012). Imagine a scenario of a future of autonomous vehicles where the unsuspecting rider is manipulated by sponsored retailers and unseen persuasions due to data mining of personal travel history, emails, and shopping preferences. Some vehicle data remains anonymous, however, when information can be identified to an individual, then it becomes personal information and must be protected. The moral obligations, the legal rights, and political considerations should ensure protection for individuals' civil liberties and freedoms (Glancy, 2012). Due to the social and human impacts, there must be emphasis on accountability of the roles, objectives, and design approaches of the developers including corporations, government, and researchers (Blyth et al., 2015).

Consumer habits are another area that can suffer undesirable affects. Will the vehicles utilize advertisements and marketing to pressure us to alter our shopping and dining behaviors? How will this advertising be regulated? Although it is true that a better understanding of the customer can offer a better user experience, this "understanding" comes in the form of vast amounts of data which could become very valuable to marketers. The anticipated efficiency of the vehicle will be compromised if marketers are able to identify travel paths with higher potential economic value (LaFrance, 2016).

Additionally, autonomous vehicle implementation requires that cities must be prepared for the changing role of the human inside and outside the vehicle, and the resulting impacts on sociotechnical structures and practices (Blyth et al., 2015). Technology has grown rapidly and the regulatory environment has not grown proportionally (Collingwood, 2017). One complexity is that neither privacy issues nor autonomous vehicles are clearly defined matters, so it is difficult to specify the interactions between them (Glancy, 2012). Uncertain topics include the types of information being collected, the reason for collecting information, accessibility, and

permissions during storage. Customer consent is not a solution if the customers do not fully understand what they are agreeing to. Anonymizing the data will also not work if deanonymizing algorithms can re-identify an individual. Additionally, in the case of surveillance, if the AVs are not personally owned, i.e. when used in transportation services, does it mean that they are a public space and subject to allowable surveillance (Taeihagh, 2019)?

The advances of autonomous vehicles can be understood with Thomas Hughes' framework of technological momentum. The development and expansion of the automobile is a conservative invention (Hughes, 1987). Automobile transportation in the United States has been sustained for nearly a century. Such systems attain technological momentum when they "have a mass of technical and organizational components; they possess direction, or goals; and they display a rate of growth suggesting velocity" (Hughes, 1987, p.76). The components of the system of autonomous vehicles include all stakeholders with interests in the technology such as automobile manufacturers, technology firms, communications providers, federal and state regulatory groups, National Highway Traffic Safety Administration (NHTSA), state departments of transportation (DOTs), state departments of motor vehicles (DMVs), advertising agencies, and public commuters (Anderson et al., 2016). The evolution of large systems progress through invention, development, innovation, transfer, and growth, competition, and consolidation. These phases are not always sequential and can overlap. Generally, inventor-entrepreneurs solve critical problems during the invention and development phases. During the development phase, the invention or idea, transforms into a complete system capable of operating in an environment with outside factors and forces (Hughes, 1987). Engineers determine solutions in the growth and momentum phase where there is dynamic expansion of the system. As systems grow, reverse salients develop. "Reverse salient are components in the system that have fallen behind or are out

of phase with the others" (Hughes, 1987, p.73). The technological advancements of autonomous vehicles necessitate data communication, transmission, and storage not previously required in past automobiles. This new functionality creates unanticipated obstacles due to the potential exposure of this data. The need for protection will require the alteration of other automobile components. This must be addressed by the system builders, engineers, car manufactures, and legislators. These problem solvers must construct centralized solutions for the privacy issues and enable coherence for future advancements (Hughes, 1987). Issues that are identified up front during the design phase can be solutioned more successfully and with less effort while the technology is still in development.

Implications of Autonomous Vehicle Technology

Autonomous vehicle technology cannot exist without vast communication, data transmission and storage between the automobile and its surroundings (Figure 1). Sensors such as radar, cameras, LIDAR, thermal imaging devices, and even sonar will be utilized to analyze and compute the routes for the vehicles (LaFrance, 2016). The different designs of self-driving systems all create and maintain internal maps of their surrounding areas using these sensors and radar. Advanced software will be used to plot efficient travel paths. Artificial intelligence will be used to integrate internal data within the vehicle with external environmental data to analyze and determine how the automobile behaves. In addition, AVs may use algorithms utilizing past user statistics gathered from EDRs (event data recorders) and artificial intelligence to provide the optimal user experience. Some of the information will be used to create highly specialized mappings and some companies will use the metrics for research to better understand everyday

driving scenarios and investigate improvements using simulation labs and test tracks (Gardner, 2019). All this data is necessary for successful operation of the autonomous vehicle.

Security in automated vehicles is more important than traditional automobiles because in the case of a cyber-attack, the passenger may not be able to recover the automobile. Techniques such as increased data redundancy are necessary because the redundancy will allow identification of conflicting data and allow proceeding to the recovery decision making process (Petit and Shladover, 2015). The very nature of the diverse radio communication required by the technology will require secure data collection and protection by means of data authentication, integrity, access control, encryption, and sanitization techniques (Mahmood et al., 2019).

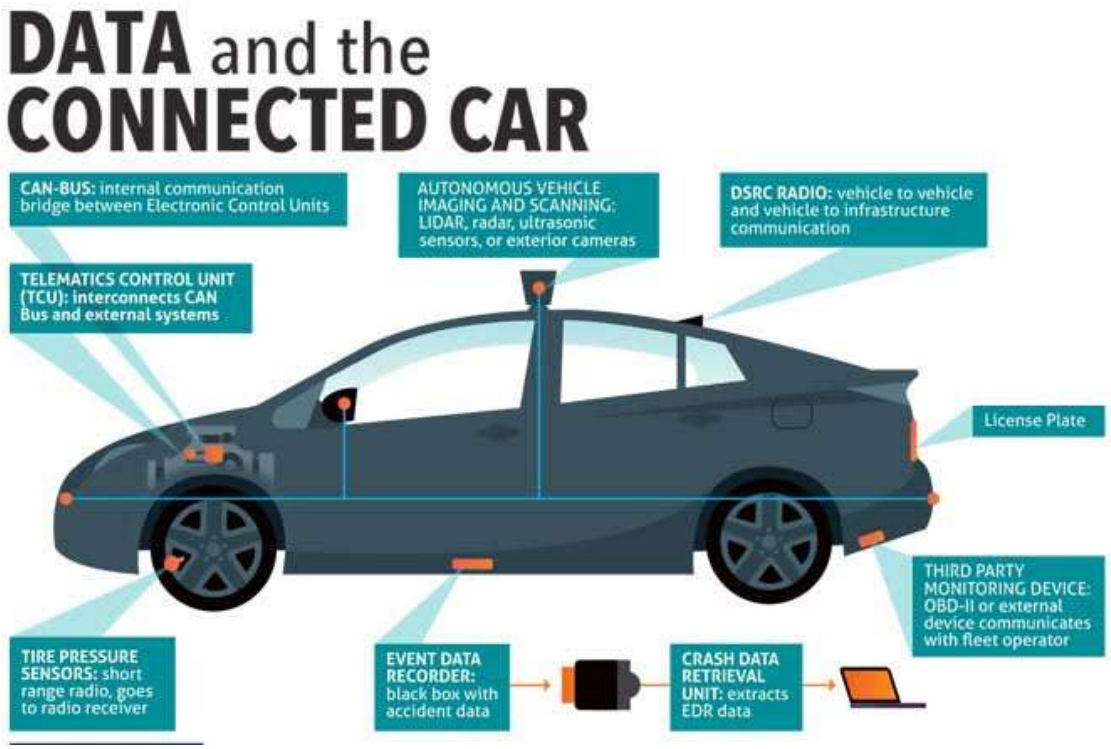

Figure 1. Data and the Connected Car (Image source: Future of Privacy Forum, 2017).

Although privacy protections can be built into the architecture, to date, US state regulations have failed to address the extensive problems associated with collection, use, storage and dissemination of data generated by autonomous vehicles. Currently only seventeen states have passed laws relating to the data retrieval from event data recorders (EDRs) (BCLP, 2018). Government avoidance in putting strict regulations in place is most likely due to the desire to promote the emerging technology (Taeihagh, 2019). Nonetheless, legislators need to act swiftly and decisively as these issues may impact the degree to which the technology is adopted, causing delays in implementation. Without clear privacy protections in place, autonomous vehicles could encounter public resistance from users who perceive them as a threat (Collingwood, 2017).

Complete control over this rapidly emerging technology will be difficult. Automotive computers contain over 100 million lines of code. As companies race to beat competitors to the market, rigorous process testing might be rushed, resulting in system failures and security vulnerabilities. Additionally, new features might integrate components from multiple sources which brings a greater challenge when completing integration testing (Nash et al, 2017). Car manufactures and other companies endorsing AV technology should employ "privacy by design" in order to ensure that the methods and regulations are defined up front during development (Glancy, 2012). This will establish consistency across the industry and eliminate the need for future modifications that may be difficult or impossible to put in place after the technology has advanced too far.

Privacy Definitions

Privacy is defined by Merriam-Webster (n.d.) as “freedom from unauthorized intrusion”. The Oxford Learner’s dictionary (n.d.) defines privacy as “the state of being alone and not watched or interrupted by other people”. When considering issues of privacy in autonomous vehicle technology, it is useful to further define privacy by elaborating on its properties. According to Pfitzmann and Hansen (2010), the six properties of privacy are anonymity, unlinkability, unobservability, undetectability, pseudonymity, and identity management. Anonymity necessitates that a person is not able to be identified within a group or set of individuals. Unlinkability is when that two actions or individuals cannot be related together. Unobservability requires that the subject is imperceptible or unnoticeable. Undetectability means that an attacker cannot determine there is a valuable piece of information available. Pseudonymity is when pseudonyms are used as identifiers instead of real names of the individual and identity management is the handling of partial identities, such as pseudonyms (Pfitzmann & Hansen, 2010).

Research Questions and Methods

Violations of an individual’s privacy are serious and potentially devastating events that require important consideration. What are the current vulnerabilities associated with autonomous vehicle privacy, security, and personal autonomy and what potential mitigations can be put in place to protect the privacy of users? Can we identify current U.S. state and federal legislature, as well as auto industry actions in the area of autonomous vehicle privacy to interpret trends of involvement and regulatory action?

The methods used for analysis included technical exploration and prior legislative policy research. The communication, data transmission, and storage technology of the autonomous vehicle was explored to understand the components associated with privacy issues in three areas: personal information, personal autonomy, and surveillance. Vulnerabilities and the potential solutions were identified. Content analysis was used to evaluate each identified potential mitigation and each was categorized according to the area of applicability, complexity, and extent to which the problem is solved. This analysis will indicate technology and methods that can potentially ensure privacy within the AV technology and can be used as guidelines or reference for future deliberations on privacy regulations of autonomous vehicles in hopes to minimize the risk of further privacy vulnerabilities. Current and past legislative actions, hearings and testimony were examined in order to elucidate discourse from federal, state, and auto industry participants on AV security and privacy topics. The status and trends of government and auto industry regulatory action were summarized in accordance with technological momentum. As the autonomous vehicles evolve within the automobile industry, these organizations and people committed by various interests, will contribute to the growth and durability of the system.

Results

By examining autonomous vehicle function and technology, the current vulnerabilities associated with autonomous vehicle privacy, security, and personal autonomy can be identified and potential mitigations to protect the privacy of users can be proposed. By determining the current state and federal legislature, as well as auto industry actions in the area of autonomous vehicle privacy, the extent of involvement and regulatory action can be shown. The results detail the findings of the AV vulnerabilities along with discussion on strategies for mitigation in terms

of data privacy, personal autonomy, and surveillance. The evidence shows that although privacy risks exist across all categories, possible resolutions are available and span all categories as well. There are several common mitigation techniques that solution multiple areas of privacy concern and these include data minimization, access controls, anonymity, legal regulations, and confidentiality agreements. Although the many possible mitigations indicate that none of the vulnerabilities are insurmountable, there are also associated difficulties which must be considered. Within the area of personal data, there are numerous opportunities for data security however the administration can be difficult because specifics can vary depending on the activity being administered. Confidentiality agreements and user consent can be valuable tools unless users are unfamiliar with the technical details or have trouble understanding what is being agreed to. Properties such as the open aspect of AV data collection will inevitably require regulation and auto industry actions to set standards for the industry. Legislative involvement will be needed to solution many of the data security and privacy vulnerabilities, however, this legislation has been slow to develop. The findings for current federal and state legislature, and auto industry actions in the areas of AV privacy and cybersecurity, have unfortunately, to date, shown very minimal engagement.

Personal Information Vulnerabilities

Personal information (or personal data) is any data that can be identified to an individual. When considering the very nature of the autonomous vehicle technology, it is not surprising that extensive amounts of data will be collected, stored, and transmitted. This data may include information about the passengers themselves, their destinations, frequent activities, the timing of where and when the person travels, as well as how they travel. Issues of importance when

considering personal information include what data is being collected, what the data is being used for, how it will be kept, and who will have access to it. Many different user groups will have interest in the AV information including vehicle developers, marketers, advertisers, transportation researchers, law enforcement, and insurance companies (Glancy, 2012). The travel patterns of users will be the most valuable personal information collected by the autonomous vehicle (Glancy, 2012). Glancy notes that this information can be used to annoy a user through targeted marketing and advertising both within the vehicle, and even after departing the AV through means such as social media or email. The personal data collected can also be used to steal a user's identity. Additionally, stalkers could use this information to frighten, threaten or even harm people. Government agencies, including law enforcement and intelligence agencies, could use collected information to find suspects for investigation or prosecution. The recordings of past locations of the AV user could be used to predict where the person will likely be located in the future and if this personal data is associated with other information then it can possibly be used to predict a person's actions. For instance, if the vehicle is regularly parked in a high-income neighborhood, then the prediction may be that the user activities include high end retail shopping or expensive dining. In such a case, the user can at risk to be manipulated by targeted advertising (Glancy, 2012).

There are several basic privacy enhancing technologies (PETs) that are effective for building in protections to both guard sensitive data as well as tighten security. One resolution for the handling of personal information is to utilize anonymous information (Eckhoff & Wagner, 2018). The goal is to modify the information so that it cannot be identified to an individual and in this way protect the individual by providing anonymity. K-anonymity is a prevalent method used to protect privacy in public releases of statistical databases (Eckhoff & Wagner, 2018). K-

anonymity is used when the database in question contains identifying information (i.e. name) as well as sensitive information. First the identifying information is removed and then database rows are grouped into equivalence classes with at least k rows that are indistinguishable with respect to their quasi-identifiers (are identifiers that do not identify users by themselves, but can do so when correlated with other data) (Eckhoff & Wagner, 2018). "Each equivalence class contains all rows that have the same values for each quasi-identifier, for example all individuals with the same ZIP Code, date of birth, and gender" (Eckhoff & Wagner, 2018, p. 9). One problem with anonymizing data is has the potential to be deanonymized using such techniques as data mining and relational database procedures (Glancy, 2012). K-anonymity can allow re-identification of individuals along with their sensitive data when the data is correlated with other data and due to this, other variations of k-anonymity have been proposed to greater ensure sensitive values are well-represented (Eckhoff & Wagner, 2018). If the anonymous aggregated data can be connected back to identify an individual or user groups, it becomes more advantageous to summarize data so that individual records do not exist instead of just removing identifiers (Glancy, 2012). The publication of this aggregated data can lessen privacy concerns. Data can be grouped by time periods, individuals, or geographic areas. "Aggregation is most effective if the raw data is hidden even from the service provider, which can be achieved by using cryptographic protocol" (Eckhoff & Wagner, 2018, p. 14), to secure a confidential communication channel.

Data minimization is another valuable alternative. The goal of data minimization is to minimize the amount of data by only collecting the least amount needed for the successful operation of the car. A difficulty with data minimization is that the AVs advanced technology's sensors most likely collect more data than is required for the task being completed (for instance a

traffic camera records not only vehicles but pedestrian's movements as well). This unrelated data is referred to as collateral data (Eckhoff & Wagner, 2018). The system should therefore be designed to limit recorded data to specifically what is needed for job or task. Data minimization can also be done through separation of data whereby distinct data about the vehicle remains separate from the data linked to an individual and only the vehicle data is collected and stored (Glancy, 2012). One potential problem with data minimization is that it could limit the innovative uses of the data. Large datasets offer opportunities for societal advancements when used by analysts or data scientists (Bowdish, 2015).

Requiring user consent is another opportunity to protect personal data. This could be achieved by requiring the user to be aware of the information that will be collected and be given an option whether to agree or not. One challenge with this may be that users may be unfamiliar with the technology and may find it difficult to understand the terminology and what they are agreeing to (Glancy, 2012).

If any of the personal information is transmitted during the functioning of the AV, it will be at risk from unauthorized access. Encryption, data security, access authentication, and confidentiality agreements would have to be employed to assure that the data is protected from other users on the network (Eckhoff & Wagner, 2018). Encryption protects the confidentiality of messages or other types of data. There are several types of encryption methods. The traditional two-party shared encryption/decryption allows encryption with the public key and corresponding decryption with the private key (Eckhoff & Wagner, 2018). Identity-based encryption is a method where the public key is an arbitrary string, such as a user's email address which allows encryption for a recipient without a public/private key pair (Eckhoff & Wagner, 2018). Attribute-based encryption utilizes that the private keys and the ciphertexts depend on user attributes

(Eckhoff & Wagner, 2018). A user's set of attributes must match the ciphertext's set of attributes in order to de-encrypt (Eckhoff & Wagner, 2018). One concern with encryption is that highly encrypted data may cause delays in the data transmission as more time is required at both ends of the communication to implement the password protection (Meyers, 2018).

Differential privacy is an approach to database privacy aimed at providing unobservability. This is done by adding a small amount of random noise to the results of the database queries so that the results of the query will be the same regardless if it contains an individual's record or not (Eckhoff & Wagner, 2018).

In addition to loss of personal data, access to personal information through legal process is easier when the information is retained by someone other than the "data subject" (Glancy, 2012, p. 1204). "Constitutional protections do not apply to law enforcement and national security officials when they seek access to personal information, not from the person, but from others who have the personal information" (Glancy, 2012, p. 1204). If the autonomous vehicle is deemed public space, then the information collected may not be considered private information.

One further risk to personal information is the problem with pre-owned or scrapped automobiles. The CPU (central processing unit) of these AVs will potentially contain a plethora of personal information. For example, "a data recorder from a rideshare vehicle may well contain a list of the previous owners linked smart devices, with addresses and ID numbers, along with a full history of everywhere the donor vehicle went in the year before the accident that wrecked it, as well as hundreds of account numbers and logs that can be used to link passengers to phone numbers, addresses, and payment histories" (Nash et al, 2017, p. 1). Automakers and mobile providers should use data wiping techniques such as factory resets or other data removal when vehicles transfer ownership or are in accidents or at end of life (Nash et al, 2017). In addition, the

auto industry should consider frequent routine data destruction. If personal information is required to "perform a particular function (such as toll payment), that personal information should be automatically destroyed when that transitory purpose (paying the toll) has been accomplished" (Glancy, 2012, p. 1238).

Personal Autonomy Vulnerabilities

Autonomy is "the act or power of making one's own choices or decisions" according to the Merriam-Webster thesaurus (n.d.). Autonomy is "concerned with individual control and self-determination - people's abilities to make independent choices about themselves" (Glancy, 2012, p.1188). This includes a person's ability to control where they move, when, and how they get there. Autonomy could be affected in 4 ways within the world of AV technology: control, choice, intrusion protection, and anonymity (Glancy, 2012).

The control aspect of the user's personal autonomy could be at risk if the AV is vulnerable to external control. This might be the case if the AV is programmed to function in a specific way when given external information (i.e. when it is automatically re-routed due to an upcoming road blockage or traffic accident). Another risk is if the autonomous vehicle has the ability to be controlled by a command or instruction from the network as this could also put the user at risk for loss of autonomy as the vehicle would be manipulated remotely (Glancy, 2012). System security and access control is important because the protection of privacy is dependent on the security of the systems and subsystems. In AVs particularly, the LIDAR and other sensors have been seen to be compromised, affecting the driving decisions of the vehicles (Eckhoff & Wagner, 2018). Access control is also a necessary protection especially with systems that have

an Internet connection as these can be compromised and taken over by remote-control (Eckhoff & Wagner, 2018).

The customer's freedom of choice will be affected if the autonomous vehicle can be used for location-based targeted advertisements. The user could become a "confined" audience, forced to the influences of the highest paid marketer (Glancy, 2012). Therefore, AVs should be designed to prevent this type of undesirable intrusion. The focus of legal protections against interferences with AV privacy emphasizes 3 key objectives: "(1) protecting user decision-making and control over whether and how an autonomous vehicle is used, (2) requiring respect for a user's choice and consent with regard to both vehicle operation and information autonomous vehicle travel, and (3) preventing intrusions including unwanted sensory inputs, such as advertising thrust on an individual using an autonomous vehicle" (Glancy, 2012, p. 1194).

Some autonomy privacy concerns can be addressed by obtaining users' agreement and consent to how the autonomous vehicle will operate, however the consent must be fully understood by the individual before this can be an effective control. Unfortunately, AV technology is quite complex and a "major challenge for autonomous vehicle developers will be to make sophisticated technical information about the consequences of using these vehicles understandable by potential users" (Glancy, 2012, p.1195).

Anonymity is another possible solution to avoid intrusions on personal autonomy. Many people will want to be free to travel without others knowing where or when they are travelling. Anonymity techniques can be used to satisfy this but only if there are not security pressures to be able to trace illegal activity and unlawful network action as this could limit the use of anonymization (Glancy, 2012).

Surveillance

In considering surveillance, the possible use of autonomous vehicle technology as a means of tracking people's travels is frightening to the public. If a person believes he is being spied on, especially if it is by the government, then they will become distrustful and paranoid within their own country. "Indeed, surveillance using autonomous vehicles could threaten the political and social well-being of our society" (Glancy, 2012, p. 2012).

"Targeted surveillance keeps track of a particular identified human person, who would otherwise expect to be let alone, and certainly not to be followed" (Glancy, 2012, p. 1209). Within the vehicle, customers could potentially be victims of surveillance if the automobile is equipped with devices to sense smoking, drinking, or other types of actions. The past history of locations, times, and dates may be collected, so the information could potentially be used by law enforcement or other private and public agencies to conduct remote surveillance (Glancy, 2012).

Firm controls over access to the network will be vital and personal information will need to be encrypted and made anonymous to avoid targeted surveillance. Otherwise, "law enforcement, national security, and other types of public and private agencies can conduct remote surveillance of the vehicle's user" (Glancy, 2012, p. 1210). This issue becomes even more important when destinations are of a private nature, for instance abortion clinics, churches, political gatherings etc. "Targeted surveillance compromises an important aspect of individual autonomy—the ability to resist being categorized, manipulated psychologically, intimidated, or mechanistically predicted by society or the government" (Glancy, 2012, p. 2012).

Mass surveillance is another cause of concern in autonomous vehicles. Mass surveillance "involves indiscriminate and comprehensive collection of personal information from everyone

within an area or sector. This type of large-scale surveillance of a population can also function as an instrument of control over the behavior of every individual within that population" (Glancy, 2012, pp. 1211-1212). This mass data could be used by marketers or advertisers to collect and identify patterns of activities or user types to be manipulated by "behavior advertising" (Glancy, 2012, p.1214). All of these ways of surveillance end up impacting the consumer's autonomy because they affect the actions of the individual and limit the ability of the individual to travel where and when they want. The overall summary of autonomous vehicle risks and possible mitigations are presented below (Table 1).

Privacy Category	Sub-Category	Risk	Mitigation	Potential Difficulties With Mitigation
Personal information	**Data**	• Unauthorized access to data • Targeted marketing and advertising • Identity theft • Stalking • Law enforcement locating • Available data to civil litigants and private investigators, in such cases as divorce actions and vehicle investigators	• Data minimization • Encryption • Frequent data destruction • Legislation and regulation for strong network protection • Only gather vehicle data and no user information • Anonymize • Summarize and remove individual data records • Require consent from users • Confidentiality agreements • Access controls • Disclose in user's manual that EDR exist • Access authentication • Differential privacy methods • Factory reset at end of life or new ownership	• Data minimization my limit innovative uses of data • High encryption may cause transmission delay • Legislature is slow/reluctant • Hard to separate individual vs. vehicle data • Deanonymize can identify users • Anonymity could hinder ability to trace a person doing wrong or illegal things • May be difficult for users to understand technical details in order to give consent or sign confidentiality agreements
Personal Autonomy	**Control**	• External control of the vehicle	• Access controls • Legal regulations	• Legislature is slow/reluctant
	Choice	• Location based targeting advertisement	• Prevent intrusions from unwanted sensory input	• Agreed upon regulations take time
	Intrusion Protection	• Law enforcement and intelligence agencies can use vehicle data to find and prosecute suspicious people • Use past location data to predict future location • Physical and psychological intrusion by censors or snoopers	• Obtain consent • Legal regulations	• May be difficult for users to understand technical details
	Anonymity	• Personal information can be correlated with other information to predict user's actions • Profiles can be used to manipulate user's choices as to where to travel and where to eat	• Anonymity • Restrict collection of data at state and federal levels	• Security pressure to be able to trace illegal actions and unlawful network activity • Data mining and relational database techniques can re-identify a person • Anonymity could hinder ability to trace a person doing wrong or illegal things
Surveillance	**Targeted Surveillance**	• Threaten social and political well-being of country if citizens feel they are being watched • Can monitor by sensing and recording activities like drinking or smoking • Secret collecting of personal information • Government and private sector investigators can subpoena or get administrative orders to access data • Harmful to psychologically health	• Data Encryption • Anonymity • Access controls • Legal regulations	• Data mining and relational database techniques can re-identify a person • Anonymity could hinder ability to trace a person doing wrong or illegal things • Legislature is slow/reluctant
	Mass Surveillance	• Mass data can be used by marketers and advertisers to predict and manipulate consumer behavior • Data can be used by news media, insurance companies, political campaigns and product manufacturers • Government control of mass surveillance can result in too much power	• Anonymity • Data Security • Legal regulations	• Data mining and relational database techniques can re-identify a person • Anonymity could hinder ability to trace a person doing wrong or illegal things • Legislature is slow/reluctant

Table 1. Autonomous Vehicle Privacy Concerns and Potential Mitigations.

Principles of Privacy by Design

Designers and developers with upfront knowledge of AV privacy vulnerabilities will be able to incorporate "privacy by design" to minimize potential risks. According to Cavoukian (2013), privacy by design encompasses seven principles: (1) proactive privacy protection instead of remedial action after privacy violations have happened, (2) privacy as the default setting, (3) privacy embedded into the design, (4) full functionality with full privacy protection, (5) privacy protection through the entire lifecycle of the data, (6) visibility and transparency, (7) respect for user privacy. Applying this to AV privacy, this would involve analyzing functional requirements, contemplating threats and attacks, and defining required security implementations. Four strategies that deal with the data itself would be to "minimize, hide, separate, and aggregate" (Eckhoff & Wagner, 2018, p. 7). Four strategies that would apply to process would be to "inform, control, enforce, and demonstrate" (Eckhoff & Wagner, 2018, p. 7). These strategies defined by Danezis et al. (2014) include (1) minimizing data collected by using select-as-you-collect, and anonymization and pseudonymization design patterns; (2) hiding data by using encryption (when in transit or when at rest), traffic hiding techniques, etc. (3) separating personal data as much as possible by means of distributed approaches (4) aggregating data to process it at the highest level of aggregation and with the least possible detail in which it is still useful by using the k-anonymity family of techniques or differential privacy (5) informing in a transparent way the subjects of the system by having adequate interfaces and detecting potential privacy breaches (6) providing control to users over data by using techniques such as user-centric identity management, end-to-end encryption, etc. (7) enforcing privacy policies by appropriate access control mechanisms (8) demonstrating the compliance with privacy policies by activities such as logging and auditing.

Current federal and state legislative actions, and auto industry participation

The collection and transmission of AV user information has been recognized as a serious problem for personal information privacy and has led to the desire for restrictions and regulations. This is especially important in the case where personal data is retained indefinitely without the individual's knowledge (Glancy, 2012). Although these issues have resulted in an increased interest in congress, lawmaking that would encourage the development and testing of autonomous vehicles has faced controversy. In 2017, during the 115th Congress, numerous committee hearings were held within both the House of Representatives and the Senate on autonomous vehicles technology and "possible federal issues that could result from their deployment" (Canis, 2020, p. 14). In the House of Representatives, the House Committee on Energy presented H.R. 3388, the SELF DRIVE Act. This bill, which passed in the House on September 6, 2017, stated that no AV could be sold domestically without an automaker's cybersecurity plan in place defining mitigation of "cyberattacks, unauthorized intrusions, and malicious vehicle control commands" (Canis, 2020, p. 17), a defined cybersecurity point of contact, plans for limiting of system access, and plans for employee training and policy maintenance. The House passed bill would have required manufacturers to specify the data collection and storage generated by the vehicles and the method of conveying that information to the vehicle owners and occupants (Canis, 2020). The manufacturer could, however, exclude from the privacy policy, methods that encrypt or make anonymous the sources of data (Canis, 2020, p. 18). Unfortunately, the Senate never took up the SELF DRIVE Act and therefore it stalled in congress.

The Senate Committee on Commerce, Science, and Transportation created a separate bill, S. 1885, the AV START Act, which emphasized prioritizing safety, promoting innovations,

reinforcing separate but complimentary federal and state regulatory rules, strengthening cybersecurity, and educating the public (Canis, 2020). The Senate bill would have mandated "written cybersecurity plans to be issued, including a process for identifying and protecting vehicle control systems, detection, and response to cybersecurity incidents, and methods for exchanging cybersecurity information" (Canis, 2020, p. 17). Additionally, a cybersecurity point of contact would be required at the manufacturer or vehicle developer (Canis, 2020). This bill would not have explicitly required privacy plans by developers, but it required National Highway Traffic Safety Administration (NHTSA) to establish a motor vehicle privacy database that would include the types of information collected during operation of the vehicle along with details on about how personally identifiable information (PII) would be "collected, retained, and destroyed when no longer relevant" (Canis, 2020, p. 18). Unlike the House passed bill, S. 1885 would require the Department of Transportation (DOT) to create incentives so that "vehicle developers would share information about vulnerabilities, and would have specified that all federal research on cybersecurity risks should be coordinated with DOT" (Canis, 2020, p. 17). Unfortunately, the bill did not make it through the 115th congress and efforts to revive it in the 116th congress failed due to objections from a group of Senate Democrats that said it did not do enough to address consumer safety and cybersecurity issues (Miller, 2019).

Reasons for failure of these bills include disagreements on the amount that Congress should modify the traditional division of vehicle regulation, whether federal standards should require technology to report and prevent hacking of critical vehicle software, and how much information should be available to car buyers (Canis, 2020, p. 17). Additionally, there were differences on "the extent to which vehicle owners, operators, manufacturers, insurers, and other parties have access to data that is generated by autonomous vehicles, and the rights of various

parties to sell vehicle-related data to others" (Canis, 2020, p. 1). The U.S. Department of Transportation and NHTSA have issued three reports since 2016 that speak to federal autonomous vehicle policies, suggesting "best practices that states should consider in driver regulation; a set of voluntary, publicly available self-assessments by automakers showing how they are building safety into their vehicles; and a proposal to modify the current system of granting exemptions from federal safety standards" (Canis, 2020, p. 1).

States have also not kept pace with the autonomous vehicle technology. "The National Governors Association (NGA) has noted that state governments have a role with respect to vehicle and pedestrian safety, privacy, cybersecurity, and linkage with advanced communications networks" (Canis, 2020, p. 19). Between 2013 and October 2019, at least 41 states and the District of Columbia have considered legislation related to autonomous vehicles. During that time, 29 states and the District of Columbia enacted legislation, governors in 11 states issued executive orders, and 5 states issued both an executive order and enacted legislation (Canis, 2020). Given all these actions, however, not one state has enacted laws addressing cybersecurity and only one single state has enacted legislature relating to privacy (Figure 2).

To date, only 2 states, Massachusetts and Pennsylvania, have drafted bills regarding future AV cybersecurity concerns (Essex & DuBois, 2020). Massachusetts has 4 bills and Pennsylvania has 1 bill, all of which are pending as of February 2020. Massachusetts bills MAS 1945 (S1945, 2017) and MAH 1829 (H1829, 2017) aim to set requirements for data collection of AVs for privacy protection, as well as to set security and accuracy standards for data collection systems and data sharing. The objectives of Massachusetts bills MAS 179 (S179, 2017) and MAS 2056 (S2056, 2019) are to ensure the security and confidentially of customer information, and protect against threats and unauthorized access. The Pennsylvania bill, PA S 427 (S427,

2017) states details about obtaining and maintaining a permit to deal with information collected by an autonomous vehicle that includes personal information from a vehicle tester or rider.

Types of Autonomous Vehicle Laws Enacted by the States

2017-2019

Topic	Number of States That Enacted Legislation
Commercial	22
Cybersecurity	0
Definitions	21
Infrastructure and connected vehicles	5
Insurance and liability	8
Licensing and registration	3
Operation on public roads	14
Operator requirements	11
Privacy of collected vehicle data	1
Request for study	6
Vehicle inspection requirements	0
Vehicle testing	10
Other	8

Figure 2. State Enacted Law for Autonomous Vehicles (Canis, 2020).

While little has been done in state government to address cybersecurity threats of future AVs, 10 states have drafted bills addressing privacy concerns in AV technology: Arizona, California, Georgia, Hawaii, Massachusetts, Minnesota, New Jersey, North Dakota, Oregon, and Texas (Essex & DuBois, 2020). Only one US state, Georgia, to date has successfully enacted one of these bills into legislation (Essex & DuBois, 2020). Bill GA S 219 was enacted as Act No. 214 on May 8th, 2017 (S219, 2017). This bill deals with the matters of AV testing, privacy of collected vehicle data, and insurance and liability concerns (Essex & DuBois, 2020). Five of 10 US states who have drafted bills regarding AV privacy concerns have failed to enact their bills in legislative committees (Arizona, California, North Dakota, Oregon, Texas) (Essex & DuBois, 2020). The main topics discussed in many of these rejected bills were how personal data should

be managed and maintained within AV internal systems. Four of ten US states who have drafted bills regarding AV privacy concerns have a current pending status on their legislation (Hawaii, Massachusetts, Minnesota, and New Jersey) (Essex & DuBois, 2020). Many of the bills still pending in state government agencies address similar topics to the rejected bills drafted by North Dakota and other states but Hawaii's bill, HIS 620 also expands its scope to include any automated systems such as drones and unmanned aircraft (S620, 2019).

The automotive industry representatives have passed their own self-regulatory guidelines to address the data privacy issues of autonomous vehicles. In 2014 the Alliance of Automobile Manufacturers and the Association of Global Automakers passed a set of seven Privacy Principles (BCLP, 2018). "Participating automobile manufacturers commit to comply with the seven Privacy Principles, which govern the collection, use, and disclosure of driver behavior information retrieved from self-driving vehicles" (BCLP, 2018, p. 1) (Figure 3).

Alliance of Automobile Manufacturers and the Association of Global Automakers 7 Privacy Principles	
1. Transparency	**Members should provide owners and registered users with ready access to clear, meaningful notices about the member's collection, use, and sharing of covered information.**
2. Choice	**Members should offer owners and registered users with certain choices regarding the collection, use, and sharing of covered information.**
3. Respect for Context	**Members should use and share covered information in ways that are consistent with the context in which the covered information was collected, taking account of the likely impact on owners and registered users.**
4. Data Minimization	**Members should collect covered information only as needed for legitimate business purposes and retaining covered information no longer than they determine necessary.**
5. Data Security	**Members should implement reasonable measures to protect covered information against loss and unauthorized access or use.**
6. Integrity and Access	**Members should implement reasonable measures to maintain the accuracy of covered information and give owners and registered users reasonable means to review and correct personal subscription information.**
7. Accountability	**Members should take reasonable steps to ensure that they and other entities that receive covered information adhere to these Privacy Principles.**

Figure 3. Alliance of Automobile Manufacturers and the Association of Global Automakers Seven Privacy Principles (BCLP, 2018)

Another action that was taken in 2016 to address potential privacy concerns was that the motor vehicle manufacturers established the Automotive Information Sharing and Analysis

Center (Auto-ISAC), which contains a set of cybersecurity principles which the Department of Transportation designates as a central clearinghouse for manufacturers to share reports of cybersecurity incidents (Canis, 2020).

Discussion

The emerging technology of autonomous vehicles currently has identifiable privacy vulnerabilities that pose risks to consumers in the area of personal information, autonomy, and surveillance. There are many opportunities for these risks to be resolved using privacy by design methods to incorporate mitigation techniques while the technology is still in the development phase. Additional support can be obtained with legislature and auto industry regulations and by implementing common standards to unify the advancement of the technology across the industry. This research expands to issues outside of the field of autonomous vehicle and is applicable to the broad technology of automation in general. Similar potential privacy issues will be present in advancing technologies such as aerial vehicles, unmanned aircraft systems, and drones, in addition to the communication links and components used for their operation.

Although this research has shed light on many potential answers to privacy risks, there are several limitations to the work that has been done. One limitation is that the scope of the legislative and policy research covered only the United States and did not consider other countries. By not addressing foreign policy as part of the research, there may have been valuable strategies or mitigation tactics that were missed. The research would have been more complete by including other countries' approach to autonomous vehicle technology as well as a comparison of this to the U.S. response. A second limitation is that analogies to other emerging technologies were not explored. Because there are existing technologies, either related or

unrelated to autonomous vehicles, that share similar requirements for data collection, transmission and storage, it would be useful to examine these technologies in order to understand their methodology for data security implementation. This investigation would help identify actual known risks along with aid in the discovery of mitigation successes that could be applied to autonomous vehicle technology. Another caveat with this research project is that, because the technology is new, ongoing development will most likely bring variations on the technical details which may differ from prior investigation. At this point in time, the total scope of the technology is not yet known and future decisions and final technical implementations may result in both new needs for privacy as well as new opportunities for mitigations.

If I were to continue this study, I would expand the scope of investigation to solicit input from the general population using surveys or interviews to identify the extend of public acceptance and trust in the technology as well as to identify critical topics. Understanding these concerns would shed light on preferred focus areas. Additionally, I would talk first hand to advocacy groups, leaders pushing for state and federal legislature, as well as other stakeholders. Lastly, I would enlarge the area of research to explore privacy outside of the vehicle including infrastructure privacy issues and implications. There are other factors that will influence the successful solutioning of AV privacy concerns such as the degree of city preparedness for autonomous vehicles and whether infrastructure changes have progressed at a pace necessary to be ready as the technology emerges.

This research project has been influential and will benefit me in my future career in systems engineering. One thing that has been reinforced through the completion of this project is the importance of thorough research and an unbiased attitude toward the subject of investigation. Although I had some knowledge of autonomous vehicles at the beginning of this project, and

understood the need for privacy considerations, I did not fully comprehend the extent of these implications until after completing the research and obtaining a more comprehensive understanding of the technology. Another thing that became clear to me through this project was the importance of the wide-reaching effects and consequences of evolving technologies and the understanding that it is not appropriate to focus solely on the efficiencies and benefits without also evaluating social and ethical implications. Like many others, I too, was initially drawn in by the beneficial promises of the technology, and without this project, would have not been aware of the many linked risks and concerns that must also be addressed. This experience will aid in my future engineering work for evaluation and research of proposals as it emphasizes the need for comprehensive and responsible decision making and planning. Understanding the total impacts, both positive and negative, will be the best way to design and implement truly successful deliverables.

Conclusion

There are many available technologies for autonomous vehicle developers to utilize to address privacy issues. By communicating and incorporating these techniques, the designers will be able to build trust and reassurance among autonomous vehicle users as well as the general public. The timing is right to focus on these issues as the technology is still in progress and flexibility exists with the ongoing design. Both the government and the auto industry must define and implement regulations and consistent standards to safeguard privacy and cybersecurity concerns. Others can expand on this study by examining additional vulnerabilities that can have potential impact on the autonomous vehicle technology especially in the area of safety and liability concerns. Another recommendation would be to emphasize more collaboration with

multiple groups such as AV technology developers, the auto industry, advocacy groups, the Department of Transportation, and state and federal government. The promise of the autonomous vehicle technology is thrilling and we want to be able to fully utilize it, while at the same time, knowing confidently that our privacy and personal information are guarded and secure.

Bibliography

Anderson, J. M., Kalra, N., Stanley, K. D., Sorensen, P., & Oluwatola, O. A. (2016). *Autonomous vehicle technology: a guide for policymakers*. Santa Monica, CA: Rand Corporation.

BCLP. (2018). Autonomous Vehicles – Data Privacy Issues. Retrieved October 9, 2019, from https://www.bclplaw.com/en-US/thought-leadership/autonomous-vehicles-data-privacy-issues.html.

Blyth, P., Mladenovic, M., Nardi, B., Su, N., Ekbia, H. (2015). Driving the self-driving vehicle: Expanding the technological design horizon. *2015 IEEE International Symposium on Technology and Society (ISTAS)*, 1-6.doi: 10.1109/ISTAS.2015.7439419.

Bowdish, L. (2015, April 2). The Risks of Data Minimization [Blog post]. Retrieved from https://www.uschamberfoundation.org/blog/post/risks-data-minimization/42945

Canis, B. (2020). *Issues in Autonomous Vehicle Testing and Deployment*. Congressional Research Service. Retrieved from https://crsreports.congress.gov/product/pdf/R/R45985

Cavoukian, A. (2013). *Privacy by Design - The 7 Foundational Principles*. Information and Privacy Commissioner of Ontario, Tech. Rep. Retrieved from https://iapp.org/media/pdf/resource_center/Privacy by Design - 7 Foundational Principles.pdf

Collingwood, Lisa (2017). Privacy implications and liability issues of autonomous vehicles. *Information & Communications Technology Law*, *26*(1), 32-45.

Danezis, G., Domingo-Ferrer, J., Hansen, M., Hoepman, J., Metayer, D., Tirtea, R., Schiffner, S. (2014). *Privacy and Data Protection by Design – from policy to engineering*. European Union Agency for Network and Information Security (ENISA). Retrieved from http://dx.doi.org/10.2824/38623

Eckhoff, D., & Wagner, I. (2018). Privacy in the Smart City—Applications, Technologies, Challenges, and Solutions. *IEEE Communications Surveys & Tutorials, 20*(1), 489–516. doi: 10.1109/comst.2017.2748998

Eliot, L. (2019, August 15). Cloud Breaches Like Capital One Will Strike At Self-Driving Cars.*Forbes.* Retrieved from https://www.forbes.com/sites/lanceeliot/2019/08/15/cloud-breaches-like-capital-one-will-strike-at-self-driving-cars/#20b799cf3a49.

Essex, A., & DuBois, G. (2020). Autonomous Vehicles State Bill Tracking Database. Retrieved February 28, 2020, from https://www.ncsl.org/research/transportation/autonomous-vehicles-legislative-database.aspx#database

Fraedrich, E., & Lenz, B. (2016). Societal and Individual Acceptance of Autonomous Driving. *Autonomous Driving*, 621–640. doi: 10.1007/978-3-662-48847-8_29.

Future of Privacy Forum. (2017, June 29). Retrieved from https://fpf.org/2017/06/29/ infographic-data-connected-car-version-1-0/

Gardner, G. (2019, September 17). Uber To Collect Mapping, Other Data For Possible Autonomous Service In Dallas. *Forbes.* Retrieved from https://www.forbes.com/sites/ Greggardner /2019/09/17/uber-to-collect-mapping-other-data-for-possible-autonomous-service-in-dallas/#4d2321d52aa4.

Glancy, D. (2012). Privacy In Autonomous Vehicles. *52 Santa Clara Law Review,* 1171, 1184.

H1829. House of Representatives Reg. Sess. 2017-2018. (Mass. 2017)

Hughes, T.P. (1987) The Evolution of Large Technological Systems. In W.E. Bjiker, T.P. Hughes, and T. Pinch (Eds). *The Social Construction of Technological Systems: New Directions in the Sociology and History of Technology*. MIT Press: Cambridge, MA. pp. 51-82.

LaFrance, A. (2016, June 23). The Creepy Thing About Self-Driving Cars. Retrieved from https://www.theatlantic.com/technology/archive/2016/03/self-driving-cars-and-the-looming-privacy-apocalypse/474600/.

Mahmood, A., Zen, H., & Hilles, S. M. S. (2019). Big Data and Privacy Issues for Connected Vehicles in Intelligent Transportation Systems. *Encyclopedia of Big Data Technologies*, 196–203. doi: 10.1007/978-3-319-77525-8_234

Merriam-Webster. (n.d.). Privacy. In *Merriam-Webster.com dictionary*. Retrieved February 22, 2020, from https://www.merriam-webster.com/dictionary/privacy

Merriam-Webster. (n.d.). Autonomy. In *Merriam-Webster.com thesaurus*. Retrieved February 22, 2020, from https://www.merriam-webster.com/thesaurus/autonomy

Meyers, J. (2018, December 30). Self-Driving Cars: How Automakers can Overcome Cybersecurity Issues [Blog post]. Retrieved from https://www.tripwire.com/state-of-security/featured/self-driving-cars-cybersecurity-issues/

Miller, M. (2019, September 26). Cyber rules for self-driving cars stall in Congress. *The Hill*. Retrieved from https://thehill.com/policy/transportation/463126-cyber-rules-for-self-driving-cars-stall-in-congress

Nash, L., Boehmer, G., Hillaker, A., & Wireman, M. (2017). Securing the future of Mobility. *Deloitte Insights.* Retrieved October 14, 2019, from https://www2.deloitte .com/us/en/insights/focus/future-of-mobility/cybersecurity-challenges-connected-car-security.html.

Petit, J., & Shladover, S. (2015). Potential Cyberattacks on Automated Vehicles, *IEEE Transactions on Intelligent Transportation Systems*, *16*(2), 546-556. doi: 10.1109/TITS.2014.2342271.

Pfitzmann, A., & Hansen, M. (2010). *A terminology for talking about privacy by data minimization: Anonymity, Unlinkability, Undetectability, Unobservability, Pseudonymity, and Identity Management*. Retrieved from http://dud.inf.tu-dresden.de/literatur/Anon_Terminology_v0.34.pdf

Privacy. (n.d.). In *Oxford Learner's Online Dictionary*. Retrieved from https://www.oxfordlearnersdictionaries.com/us/definition/english/privacy?q=privacy

S179. Senate Reg. Sess. 2017-2018. (Mass.2017)

S219. Senate Reg. Sess. 2017-2018. (Ga. 2017)

S427. Senate Reg. Sess. 2017-2018. (Pa. 2017)

S620. Senate Reg. Sess. 2019-2020. (Haw. 2019)

S1945. Senate Reg. Sess. 2017-2018. (Mass.2017)

S2056. Senate Reg. Sess. 2019-2020. (Mass.2019)

Taeihagh, A., & Lim, H. (2019). Governing autonomous vehicles: emerging responses for safety, liability, privacy, cybersecurity, and industry risks. *Transport Reviews*, 39(1), 103-128. doi: 10.1080/01441647.2018.1494640.

Union of Concerned Scientists. (2018). Self-Driving Cars Explained. Retrieved from https://www.ucsusa.org/clean-vehicles/how-self-driving-cars-work.