Revolutionizing Our Roadways
Data Privacy Considerations for Automated and Connected Vehicles
CONTENTS

Executive Summary ................................................................................................................ 1
Chapter One: Introduction ....................................................................................................... 3
Chapter Two: Data Privacy Protection Is an Important Concern for Americans ................. 6
  Federal Privacy Protection .................................................................................................. 6
  Texas PII Laws and Regulations ......................................................................................... 9
Chapter Three: Best Practices in Data Privacy Protection ...................................................... 13
Chapter Four: New Technologies May Challenge Existing Data Privacy Protections .......... 14
  The Debate about Data Anonymization .......................................................................... 14
  The Auto Industry Lacks Consensus Concerning How to Protect Drivers’ Data Privacy ... 15
  Many Types of Data Are Gathered and Transmitted by Passenger Vehicles .................... 16
  Consumers Are Willing to Trade Their Personal Data to Obtain a Benefit ....................... 17
  Eighty-Six Percent of Adults Surveyed by AAA Think There Should Be Laws and Policies to
  Protect Their Vehicle Data ............................................................................................... 18
Chapter Five: Case Study of Transportation-Related PII in Texas ........................................ 19
  How Are Vehicular Data Being Collected in Texas? ........................................................ 19
  How Are Transportation Data Being Used? ....................................................................... 21
  Who Owns the Data? .......................................................................................................... 22
  What Is the Value of Vehicular Data? ................................................................................ 23
Chapter Six: State and Federal Approaches to Automated and Connected Vehicle
  Data Privacy ......................................................................................................................... 24
  Federal Efforts Concerning Automated and Connected Vehicle Data Privacy ................. 24
  State Actions to Address Automated and Connected Vehicle Data Privacy ....................... 26
Chapter Seven: Policy Considerations for Texas ................................................................. 27
  Is There a Need for New Policies to Manage Vehicular and Transportation Data Collected
  and Used by the State as Connected and Automated Vehicles Are Deployed? ................. 27
  Should There Be a State or Federal Approach to Managing Vehicular Data Issues from
  Connected and Automated Vehicles? ............................................................................... 28
Chapter Eight: Findings .......................................................................................................... 29
  There Is Not a Consistent Statutory Definition of “Personal Identifying Information”;
  Section 521.002 of the Texas Business and Commerce Code Appears to Be the Most
  Encompassing Definition of PII for Texas ........................................................................ 29
  The Adequacy of Data Anonymization as the Basis for Privacy Protection of PII Gathered by
  State Agencies, for Transportation or Other Uses, May Need to Be Reconsidered .......... 30
  Stakeholders Are Uncertain about Who Owns the Data from Vehicles ............................. 30
  There Is a Rush by Stakeholders to Monetize Vehicular Data from Automated and
  Connected Vehicles ........................................................................................................ 30
  State Agencies Have Made Few Preparations to Deal with the Extremely Large Volume
  of Vehicular Data That Experts Predict Will Be Produced by Automated and Connected
  Vehicle Technologies ...................................................................................................... 31
References ............................................................................................................................. 32
Executive Summary

The 2015 Texas legislative session highlighted Texas policy makers’ concerns about protecting citizens’ personally identifiable information (PII). Texas legislators introduced seven measures that addressed the protection of PII, including protection of the data privacy of certain populations (e.g., juveniles) and protection of citizens’ data privacy in specific circumstances (e.g., the use of location data in search warrants). These disparate measures reflect the current body of law and regulation concerning PII and data privacy—both in Texas and nationally—which is segmented into legal protections governing specific types of data.

The Texas Business and Commerce Code defines PII as data that “alone or in conjunction with [emphasis added] other information identifies an individual” (Section 521.002). This definition is important for statutory protections concerning PII, as well as requirements about how organizations handle certain data types. Some laws, for example, require federal and state agencies to take special care when handling someone’s name, address, financial data, and medical data. As part of the research reported here, the research team learned that there are many different definitions of PII, both in Texas and nationally. At the federal level, the White House proposed new legislation in January 2015 to create uniform state laws concerning PII protection.

The protection of PII is of interest to both state and federal policy makers, especially in the context of emerging technologies that may capture a great deal of PII. One example is the capture and use of data about an individual’s location by a smartphone or by automated or connected vehicle technology.

The U.S. Supreme Court recently handed down an important decision about law enforcement use of a global positioning system (GPS) tracking device to gather data about an individual’s location (United States v. Jones) and the need for a warrant to search an individual’s smartphone (Riley v. California). The focus of this report is on data privacy in the context of transportation, not data privacy generally. Researchers provide the federal and Texas statutory/legal framework for data privacy protection, and use automated and connected vehicle technology as a case study to explore how a new technology may impact the protection of PII and the vehicular data privacy of Texas citizens.

Researchers found that privacy protections for PII are segmented or siloed in the United States at federal and state levels. Researchers also found that there are different definitions of "personally
identifiable information” in Texas as well as nationally and internationally; this discrepancy can lead to confusion about which data are covered by laws and regulations. The research revealed several areas concerning PII and vehicular data privacy that Texas might address to prevent problems arising from new technologies in the near future:

- There is not a consistent statutory definition of “personal identifying information”; Section 521.002 of the Texas Business and Commerce Code appears to be the most encompassing definition of PII for Texas.

- There is a rush by stakeholders to monetize vehicular data from automated and connected vehicle technologies.

- State agencies have made few preparations to deal with the extremely large volume of vehicular data that experts predict will be produced by automated and connected vehicle technologies.

- The adequacy of data anonymization as the basis for privacy protection of PII gathered by state agencies for transportation or other uses may need to be reconsidered.

- Stakeholders are uncertain about who owns the data from vehicles.
Chapter One: Introduction

This report is designed to provide policy makers with information regarding the current and developing challenges regarding the protection of personally identifiable information and the available protection techniques. Specifically, the report uses new technologies regarding automated and connected vehicle data as a case study to examine these challenges and techniques.

The Texas Business and Commerce Code defines PII as data that “alone or in conjunction with [emphasis added] other information identifies an individual” (Section 521.002). How PII is defined is important for the scope of statutory PII protections, as well as requirements about how organizations handle certain data types. Some laws, for example, require federal and state agencies to take special care when handling someone’s name, address, financial data, and medical data. As part of this research, the research team learned that there are many different definitions of PII, both in Texas and nationally.

It is difficult to identify any generally accepted definitions of the types of data that may be gathered, produced, or transmitted by automated and connected vehicles; some of these data may contain PII. An auto care industry association task force developed a set of definitions in a white paper to distinguish between and among different types of telematics data that are gathered, produced, or transmitted by automobiles currently. The following definitions illustrate the many different types of data that are currently gathered, produced, or transmitted by passenger vehicles:

- **Driver personal data**—includes passwords, account passwords, phone numbers, logins, geolocation data, personal history, biometrics, driver behavior, etc.
- **In-vehicle infotainment**—includes songs, movies, games, maps, applications, and other third-party copyrighted material.
- **Forensic information**—data used by car companies, legal entities, and insurance companies to determine driving parameters following a crash. It can also include any information gleaned from a vehicle following any crime where the vehicle can reveal evidence, even the driver’s personal data.
- **Inspection data**—includes emissions and safety-related data and codes used for official periodic inspections.
- **Diagnostic data**—codes and parameter identifications used to diagnose vehicle faults. This also includes prognosis information including oil quality monitors and other data used to predict or communicate service scheduling.
- **Vehicle manufacturer proprietary information**—includes onboard software, some security-related information, and calibration information. Calibration files contain vehicle configuration data, such as upper and lower operating limits, code-setting thresholds, and other data unique to a specific year, make, model, and engine configuration.
- **Intelligent transportation system (ITS) information**—includes GPS, radar, sonar, yaw, and accelerometer information used for the vehicle (I).
This report uses the term “vehicular data” to include the data outlined by the white paper’s definitions of “driver personal information,” “in-vehicle infotainment,” and “intelligent transportation system information.” All of these categories of data can contain information that can identify an individual and so have data privacy implications. Location data, which identify where an individual driver is at a particular time, or the origin and destination of a trip linked to an individual driver, can be part of any of these three categories. For example, smartphones, in-vehicle maps, GPS, and entertainment systems may use an individual’s location data to provide services.

Automated and connected vehicles will use a combination of onboard sensors and communications technology (e.g., dedicated short-range communications) to identify the location of the vehicle itself and other vehicles on the road, which is part of the safety function of these new technologies. This type of location data may or may not be traceable to an individual driver. If location data can be used to identify an individual “alone or in conjunction with other data,” as defined in the Texas Business and Commerce Code, then location data may be encompassed by the Texas Business and Commerce Code’s definition of PII. Researchers refer to “transportation data” in this report to describe aggregated and anonymized travel and traffic information that is used by state agencies for traffic management purposes, which does not include PII.

Through an extensive review of the literature on data privacy and a series of interviews with subject matter experts, the research team learned that many Americans highly value the privacy of their vehicular data, but currently available and emerging vehicle technologies may make it difficult to secure this information. For example, a 2015 survey conducted by U.S. Senator Markey’s office determined that “nearly 100 percent of cars on the market include wireless technologies that could pose vulnerabilities to hacking or privacy intrusions” (2). A 2013 survey of 2000 adults by the Auto Alliance found that privacy is an issue for consumers, with 75 percent of survey respondents indicating that they were very/somewhat concerned that companies would collect data from the software operating self-driving cars (3). A 2013 telephone survey conducted by the American Automobile Association (AAA) found that 86 percent of the 1,007 U.S. adults surveyed thought there should be laws and policies to protect their vehicle data (4).

The 2015 Texas legislative session highlighted Texas policy makers’ concerns about protecting citizens’ PII. Texas legislators introduced seven measures that would address the protection of PII, including protection of the data privacy of certain populations (e.g., juveniles) or protection of citizens’ data privacy in specific circumstances (e.g., the use of location data in search warrants). These proposed measures reflect the current body of law and regulation concerning PII and data privacy—both
individual’s location by a smartphone or by automated or connected vehicle technology. The U.S. Supreme Court recently handed down decisions about law enforcement use of a GPS tracking device to gather data about an individual’s location (United States v. Jones), and the need for a warrant to search an individual’s smartphone because of the vast amount of personal data contained in a smartphone (Riley v. California).

The proposed Personal Data Notification and Protection Act would expand upon current state (including Texas) definitions of “personally identifiable information” or “sensitive personally identifiable information.”

The remainder of this report is organized in the following way:

- Chapter Two examines the state and federal legal and regulatory framework for data privacy protection and the lack of uniformity in the definition of PII.
- Chapter Three discusses various data privacy protection frameworks, including international and national best practices to protect PII, which have several consistent requirements.
- Chapter Four explains how new technologies may challenge existing data privacy protections.
- Chapter Five examines a case study regarding transportation-related PII in Texas.
- Chapter Six examines state and federal efforts to protect automated and connected vehicle data privacy.
- Chapter Seven identifies policy considerations regarding transportation-related PII in Texas.
- Chapter Eight discusses the findings of this research.

In Texas and nationally—which is segmented into legal protections governing specific types of data.¹

At the federal level, the White House proposed new legislation in January 2015 to create uniform state laws concerning PII protection, including:

- The Personal Data Notification and Protection Act.
- The Student Digital Privacy Act.
- A Voluntary Code of Conduct for Smart Grid Customer Data Privacy.
- A Consumer Privacy Bill of Rights.

The proposed Personal Data Notification and Protection Act would expand upon current state (including Texas) definitions of “personally identifiable information” or “sensitive personally identifiable information.” The protection of PII arises as an issue for both state and federal policy makers, especially in the context of emerging technologies that may capture a great deal of PII. One example is the capture and use of data about an indi-

¹For example, Texas legislators introduced measures relating to data privacy in the following areas: security breach notification and prosecution of the offense of breach of computer security; removal of PII from car crash information maintained by the state; removal of PII from state juvenile court records; removal of PII from announcements about lottery winners; handling of biometric information by state agencies; handling of medical data with PII by state agencies; and a requirement for search warrants to obtain location information.
Chapter Two: Data Privacy Protection Is an Important Concern for Americans

Over the last 40 years, state and federal law has developed into an inconsistent patchwork of privacy protections for specific data types or individuals. These protections’ efficacy will likely be tested as connected devices continue to proliferate.

The Internet of Things (IoT), according to the Federal Trade Commission’s (FTC’s) report, refers to “the ability of everyday objects to connect to the Internet and to send and receive data” (5). The FTC report about the recent IoT workshops notes that “six years ago, for the first time, the number of ‘things’ connected to the Internet surpassed the number of people” (5). The FTC report predicts that this technology trend is just beginning and cites expert estimates that in 2014 “there will be 25 billion connected devices, and by 2020, 50 billion” (5). This chapter describes the segmented approach to data privacy protection taken by the U.S. Congress, the federal courts, and the Texas Legislature.

Federal Privacy Protection
FTC is the chief federal agency concerned with privacy policy and enforcement (6), and its mission is to use “law enforcement, policy initiatives, and consumer and business education to protect consumers’ personal information and ensure that they have the confidence to take advantage of the many benefits of the ever-changing marketplace” (5).

The following cases and statutes illustrate the current segmented approach to data privacy that the U.S. Congress has taken. In 1978, in response to the Supreme Court’s rejection of a limited right to privacy in financial documents in United States v. Miller, Congress enacted the Right to Financial Privacy Act of 1978 (RFPA). RFPA protects consumers by preventing banks and financial institutions from disclosing an individual’s personal financial information without a judicially granted subpoena or a search warrant. RFPA bolstered the consumer protections in the Bank Secrecy Act of 1970 and the Fair Credit Reporting Act of 1970. The Gramm-Leach-Bliley Act limits the information financial institutions may share with third parties without the prior consent of customers.

Congress has taken a narrowly focused approach to defining rights to privacy in other areas of PII. In addition to the privacy of financial documents, federal statutes protect health information (Health Insurance Portability and Accountability Act of 1996 [HIPPA]), children’s use of the Internet (Children’s Online Privacy Protection Act of 1998 [COPPA]), video rental and purchase information (Video Privacy Protection Act of 1988), and PII held by cable companies (Cable Communications Policy Act of 1984 [1984 Cable Act]). The Family Educational Rights and Privacy Act of 1974, as amended, outlines privacy protections for parents and students, including rights to confidentiality, restrictions on disclosure of PII, and the right to inspect records. The Controlling the Assault of Non-solicited Pornography and Marketing Act of 2003 establish-
es national standards for the sending of commercial email and requires FTC to enforce its provisions. The Telephone Consumer Protection Act of 1991 limits the use of automatic dialers for placing unsolicited calls to consumers, among other privacy intrusions by telephone. In addition, 47 states, the District of Columbia, Puerto Rico, and the Virgin Islands have all enacted legislation requiring customer notification when PII has been released, as well as remedial action if such information has been disclosed or lost pursuant to a data breach (7). Many (but not all) of these laws impose administrative, civil, or criminal penalties for non-compliance with their requirements.

Each of these statutes reflects a policy choice to enhance the protection of certain types of data over the interests of other stakeholders.

Statutes that restrict the use of PII include protections on the collection of data (COPPA), the maintenance and transmission of data (HIPPA), and the use of data (the Video Privacy Protection Act). However, some statutes cover all three aspects of PII protection, such as the 1984 Cable Act, which restricts the collection, maintenance, and dissemination of cable subscriber information. Data protection rules in the European Union create restrictions based on functions such as data controller, data subject, and data processor; current U.S. statutes have not incorporated a similar functional framework for protecting PII (8).

These disparate federal laws illustrate how Congress has taken the approach of addressing discrete areas in which the collection and use of PII present a specific potential harm to individuals, such as the handling of a patient’s medical information, the release of a customer’s financial information, and children’s use of the Internet. Each of these statutes reflects a policy choice to enhance the protection of certain types of data over the interests of other stakeholders. For example, commercial enterprises may seek access to an individual’s health-related information to better target pharmaceutical advertising or to children’s email addresses to target them for new games and products. Law enforcement organizations may seek access to an individual’s financial information for investigative purposes. However, the U.S. Congress has made a policy choice to place constraints on access to those specific types of personal information. Although the statutes described above provide constraints concerning certain types of data collection, they do not clearly address the PII that may be part of new technologies, such as automated and connected vehicles.

Currently, there is no federal statute that comprehensively protects individual privacy. Instead, there is a patchwork of different federal statutes about the protection of specific types of personal information. The U.S. Supreme Court established many of the limitations on government intrusion into an individual’s privacy through its decisions concerning searches and seizures by the government constituting a violation of the Fourth Amendment. The Supreme Court recently decided two cases that are relevant to the protection of individual data privacy.

In the most recent decision, Riley v. California, the court decided unanimously that police officers must obtain a warrant to search an individual’s cell phone. The court stated in the syllabus of the case that, contrary to the physical intrusion of inspecting an arrestee’s pockets, “more substantial privacy interests are at stake when digital data is involved” (9). Connected and automated vehicles contain a great deal of digital data, much of it PII, according to the experts the research team interviewed. For this reason, new technologies such as connected and automated vehicles may present the same privacy issues that the court identified in Riley (9):

This is an illustrative, rather than a comprehensive, examination of U.S. statutes that constrain the collection and use of data. Other statutes that constrain the collection and use of data include the Privacy Protection Act of 1980, which restricts the search or seizure of work product materials in the possession of third parties by government officers, and the Computer Matching and Privacy Protection Act of 1988, which regulates the federal government’s practice of comparing individual information stored across different agency databases. See, for example, the Children’s Online Privacy Protection Act of 1998, Health Insurance Portability and Accountability Act of 1996, and Gramm-Leach-Bliley Act.

In this chapter, connected vehicle is used to refer to the broader concept of a car having a large amount of information being exchanged with different entities for infotainment purposes, not the federal government’s connected vehicle program. This program is discussed in greater detail in Chapter Six.
First, a cell phone collects in one place many distinct types of information that reveal much more in combination than any isolated record. Second, the phone’s capacity allows even just one type of information to convey far more than previously possible. Third, data on the phone can date back for years. In addition, an element of pervasiveness characterizes cell phones but not physical records. A decade ago officers might have occasionally stumbled across a highly personal item such as a diary, but today many of the more than 90% of American adults who own cell phones keep on their person a digital record of nearly every aspect of their lives.

Given that some industry experts have said that connected vehicles of the future are likely to resemble a “smartphone on wheels,” this decision could have important implications for government use of connected vehicle data. A recent white paper concerning connected vehicles suggests that “the car is no longer just a way to get from point A to point B. In fact, it’s not just a car: it’s the control centre for your mobile life” (10).

Also in Riley, the Supreme Court highlighted location data as a particular privacy concern, stating (9):

(Data on a cell phone can also reveal where a person has been. Historic location information is a standard feature on many smart phones and can reconstruct someone’s specific movements down to the minute, not only around town but also within a particular building. See United States v. Jones, 565 U. S. ___, ___ (2012) (SOTOMAYOR, J., concurring) (slip op., at 3) (“GPS monitoring generates a precise, comprehensive record of a person’s public movements that reflects a wealth of detail about her familial, political, professional, religious, and sexual associations.”).

The Supreme Court concluded in Riley (9):

Modern cell phones are not just another technological convenience. With all they contain and all they may reveal, they hold for many Americans “the privacies of life,” Boyd, supra, at 630. The fact that technology now allows an individual to carry such information in his hand does not make the information any less worthy of the protection for which the Founders fought. Our answer to the question of what police must do before searching a cell phone seized incident to an arrest is accordingly simple—get a warrant.

United States v. Jones, the recent Supreme Court case cited in the court’s Riley decision, involved the installation of a warrantless GPS tracking device on the vehicle of a suspected drug trafficker by the police. The court unanimously ruled that the installation of the tracking device was a violation of the defendant’s Fourth Amendment rights, although the court split 5–4 on the reasons why it was a violation. The majority reasoned that a physical
GPS installation on the vehicle constituted a “trespass” against the defendant, while the minority cited Justice Harlan’s concurrence in Katz v. United States, arguing that the long-term GPS surveillance was a violation of the defendant’s reasonable expectation of privacy. A GPS device provides location data concerning the vehicle as part of the way it functions. This case is relevant to the data produced by connected and automated vehicle technology, which will provide a vehicle’s location information to third parties, such as other automobiles and state-owned infrastructure. Justice Sotomayor’s concurring opinion in United States v. Jones indicated that the Supreme Court may soon find it “necessary to reconsider the premise that an individual has no reasonable expectation of privacy in information voluntarily disclosed to third parties” (17). The decision in Jones raises the question of whether state-owned facilities may gather—and track—an individual’s vehicle location without his or her consent, and without a warrant to do so.

The court unanimously ruled that the installation of the tracking device was a violation of the defendant’s Fourth Amendment rights, although the court split 5–4 on the reasons why it was a violation.

Texas PII Laws and Regulations
Texas has many statutes and regulations that refer to PII. Specifically, there are at least 42 Texas statutory provisions that reference PII and at least 25 regulatory provisions that refer to PII in the Texas Administrative Code. Many of the provisions in the Texas Administrative Code concern early childhood intervention services, privacy considerations related to health benefit plans, and law enforcement fusion centers, as well as a statute prohibiting the unauthorized collection or culling of PII.\(^5\) Section 521.002 of the Texas Business and Commerce Code defines personal identifying information as information that:

...alone or in conjunction with other information identifies an individual, including an individual’s:

a. name, social security number, date of birth, or government-issued identification number;

b. mother’s maiden name;

c. unique biometric data, including the individual’s fingerprint, voice print, and retina or iris image;

d. unique electronic identification number, address, or routing code; and

e. telecommunication access device as defined by Section 32.51, Penal Code.

The same section of the Texas Business and Commerce Code (Section 521.002) states that sensitive personal information means, subject to Subsection (b):

(A) an individual’s first name or first initial and last name in combination with any one or more of the following items, if the name and the items are not encrypted:

(i) social security number;

(ii) driver’s license number or government-issued identification number; or

(iii) account number or credit or debit card number in combination with any required security code, access code, or password that would permit access to an individual’s financial account; or

(B) information that identifies an individual and relates to:

(i) the physical or mental health or condition of the individual;

(ii) the provision of health care to the individual; or

(iii) payment for the provision of health care to the individual.

Neither of these definitions specifically describes an individual’s location data available from GPS and other telematics devices in an automobile. Location information is important personal information, for as Justice Sotomayor explained in United States v. Jones, “GPS monitoring generates a precise, comprehensive record of a person’s

\(^5\)See, for example, Texas Administration Code Title 40, § 108.235 (Procedural Safeguards and Due Process Procedures, Division for Early Childhood Intervention Services); Texas Administration Code Title 28, § 21.3806 (Texas Department of Insurance, Eligibility Statements, Privacy Issues); Vernon’s Texas Statutes and Codes Annotated, Government Code, § 421.084, “Fusion Centers Operating in This State: Rules and Monitoring”; Vernon’s Texas Statutes and Code Annotated, Business and Commerce Code § 324.051, “Unauthorized Collection or Culling of Personally Identifiable Information.”
public movements that reflects a wealth of detail about her familial, political, professional, religious, and sexual associations” (11). In order to protect location information, the Texas Business and Commerce Code’s definition of personal identifying information encompasses information that “alone or in conjunction with [emphasis added] other information identifies an individual,” which is the broader of the two definitions. The Texas Penal Code criminalizes the fraudulent use or possession of PII, which is referred to in Section 32.51 of Title 7 as “identifying information.” Specifically, the Texas Penal Code states:

(1) “Identifying information” means information that alone or in conjunction with other information identifies a person, including a person’s:
   (A) name and social security number, date of birth, or government-issued identification number;
   (B) unique biometric data, including the person’s fingerprint, voice print, or retina or iris image;
   (C) unique electronic identification number, address, routing code, or financial institution account number; and
   (D) telecommunication identifying information or access device.

(2) “Telecommunication access device” means a card, plate, code, account number, personal identification number, electronic serial number, mobile identification number, or other telecommunications service, equipment, or instrument identifier or means of account access that alone or in conjunction with another telecommunication access device may be used to:
   (A) obtain money, goods, services, or other thing of value; or
   (B) initiate a transfer of funds other than a transfer originated solely by paper instrument.

This definition may include location information, but the distinction between what PII may be the subject of criminal and civil penalties, in the event of a major fraud or data breach, is unclear.

Perhaps the most important aspect of PII protection from the perspective of legislators is how state agencies handle this information. The Texas Government Code, Title 5, Chapter 559: State Government Privacy Policies, describes the rights of individuals and imposes requirements on state entities regarding data privacy. These protections include the following:

- **The right to be informed about information collected**—“An individual is entitled to be informed about information that a state governmental body collects about the individual.”

- **The right to notice about certain information laws and practices**—A “state governmental body that collects information about an individual” on paper or through an online form must prominently inform individuals that “the individual is entitled on request to be informed about the information that the state governmental body collects about the individual,” “the individual is entitled to receive and review the information,” and “the individual is entitled to have the state governmental body correct information about the individual that is incorrect.”

- **The right to correction of incorrect information**—“Each state governmental body shall establish a reasonable procedure under which an individual is entitled to have the state governmental body correct information about the individual that is possessed by the state governmental body and that is incorrect. The procedure may not unduly burden an individual using the procedure.”

- **State agencies may not charge a fee** to individuals requesting to have their information corrected.

Perhaps the most important aspect of PII protection from the perspective of legislators is how state agencies handle this information.
this state.” Texas Government Code 2054.1125(2) defines sensitive personal information in the context of a data breach for which notification is required as having “the meaning assigned by Section 521.002, Business and Commercial Code.” As defined above, sensitive personal information means, subject to Subsection (b), an individual’s first name or first initial and last name in combination with any one or more of the following items, if the name and the items are not encrypted:

(A) social security number;
(B) driver’s license number or government-issued identification number; or
(C) account number or credit or debit card number in combination with any required security code, access code, or password that would permit access to an individual’s financial account.

(3) “Victim” means a person whose identifying information is used by an unauthorized person.

(b) For purposes of this chapter, the term “sensitive personal information” does not include publicly available information that is lawfully made available to the public from the federal government or a state or local government.5

---

A data breach concerning individual motorists’ location data—which are highly valued by marketers and other stakeholders—may not be covered by the definition of sensitive personal information. This is a plausible scenario since experts interviewed indicated that there is likely to be an extremely large volume of data resulting from automated and connected vehicle technology in the near future. They indicated that state agencies are not equipped to deal with such an overwhelming influx of vehicular and transportation data, and they predicted that third parties will be used to store and analyze state vehicular and transportation data. Therefore, care must be taken in devising state government agreements with third parties to manage vehicular and transportation data; for example, it will be important to specify a definition of personally identifying information in contracts and agreements that is as broad as possible in order to encompass all of the types of PII that could be subject to breach or other harm.

---

5 The Texas Government Code contains provisions concerning open government in Title 5, Subtitle A, Chapter 552, concerning public information. Section 552.130 of Chapter 552 provides an exception that protects the confidentiality of certain motor vehicle records, such as a driver’s license, although it does not specifically provide an exemption for personally identifiable information.
The Texas Government Code, Title 10, Subtitle B, Chapter 2054, Information and Resources, Section 2054.133 requires state agencies to develop and keep updated an information security plan to protect the agency’s information. Specifically, Texas requires state agencies to:

1. Consider any previously developed vulnerability reports.
2. Incorporate any existing network security services.
3. Identify and define the responsibilities of agency staff who produce, access, use, or serve as custodians of the agency’s information.
4. Identify risk management and other measures taken to protect the agency’s information from unauthorized access, disclosure, modification, or destruction.
5. Include:
   a. the best practices for information security developed by the department; or
   b. a written explanation of why the best practices are not sufficient for the agency’s security; and
   c. omit from any written copies of the plan information that could expose vulnerabilities in the agency’s network or online systems.

In summary, the provisions of the following are not harmonized: Section 521.002 of the Texas Business and Commerce Code (defining personal identifying information); Section 32.51 of Title 7 of the Texas Penal Code (defining identifying information); Title 5, Chapter 559, of the Texas Government Code (describing the rights of individuals and imposing requirements on state entities regarding data privacy); Section 2054.1125(b) of the Texas Government Code (describing requirements for response to data breaches); and Texas Government Code 2054.1125(2) (defining sensitive personal information in the context of a data breach for which notification is required as having “the meaning assigned by Section 521.002, Business and Commercial Code”). The potential result could be a failure to protect important PII, such as location data, which is not specifically addressed by any of these provisions.

Similarly, it is unclear how the fraudulent use or possession of location data might be prosecuted under the Texas Penal Code. One approach to addressing these concerns might be to generally adopt the definition of PII in Section 521.002 of the Texas Business and Commerce Code, which is the broadest definition currently in use in Texas.

For example, a third-party provider of transportation data management services for the State of Texas might argue that location data were not within the scope of the PII for which it was contractually required to be responsible. For that reason, it might refuse to comply with the data breach notification provisions of Texas law. Similarly, it is unclear how the fraudulent use or possession of location data might be prosecuted under the Texas Penal Code. One approach to addressing these concerns might be to generally adopt the definition of PII in Section 521.002 of the Texas Business and Commerce Code, which is the broadest definition currently in use in Texas.

The following chapters provide a review of international and federal best practices for data privacy protection and consider how these practices are recognized in current Texas statutes and regulations.

Chapter Three: Best Practices in Data Privacy Protection

As a part of understanding the current data privacy environment, the research team investigated the best practices for ensuring data privacy. To accomplish this task, the researchers conducted a thorough review of best practices for preserving data privacy across a wide variety of sources.

The review covered best practices put forth from leading industry, federal government, and international organizations, including the:

- White House Consumer Privacy Bill of Rights.
- National Institute of Standards and Technology Guide to Protecting the Confidentiality of PII.
- Fair Information Practice Principles from:
  - The Department of Homeland Security.
- Alliance of Automobile Manufacturers.
- Organisation for Economic Co-operation and Development.

The research team learned that these groups all had best practice lists with many overlapping themes. For the purposes of this report, the researchers distilled these many lists into three overarching best practice principles, each with associated practices, which are summarized in Table 1.

While current data privacy practices may be sufficient for Texas state agencies in the current digital environment, the connected and automated vehicle data environment of the near future may require revisiting and reevaluating their adequacy. The distilled best practices included in this report could be useful as a starting point for developing uniform future policies concerning new technologies such as automated and connected vehicles, transportation data PII, and data privacy.

### Table 1: Best Practices

<table>
<thead>
<tr>
<th>Principle</th>
<th>Description</th>
</tr>
</thead>
</table>
| Transparency | Have individuals been informed about how and why their data are being used?  
               | Have individuals provided consent?  
               | Do individuals have the ability to access their personal data? |
| Security     | Have organizations minimized the data they collect by limiting collection to only what information is necessary?  
               | Have organizations taken reasonable and appropriate safeguards to protect PII from loss, unauthorized access or use, destruction, modification, or unintended or inappropriate disclosure?  
               | Are there measures in place to handle any inappropriate use, disclosure, or other breaches? |
| Accountability| Are data privacy requirements adequately enforced?  
                  | Is there a regular source of auditing for the organization to ensure compliance?  
                  | Can consumers verify and correct personal information? |
Chapter Four: New Technologies May Challenge Existing Data Privacy Protections

This chapter addresses five issues that illustrate the challenges for data privacy protection in the context of new technologies:

- The current debate about the efficacy of data anonymization to protect data privacy.
- The current lack of consensus in the automobile industry about how to protect drivers’ data privacy.
- The many types of data that are gathered and transmitted by vehicles.
- The possibility that many Americans are willing to trade their personal data in exchange for a perceived benefit.
- A study by AAA that showed that 86 percent of Americans surveyed think there should be laws and policies to protect their vehicular data (4).

The Debate about Data Anonymization

One of the key proposals for dealing with PII from new technologies is to anonymize the data through a variety of different software processes. Researchers learned from interviews with subject matter experts that there is no consensus about whether data anonymization is actually possible. Opinions of experts ranged from “there is no such thing as data anonymization” to “anonymization would be a good way to minimize privacy concerns.” One respondent suggested that a kind of bill of rights is needed for consumer data obtained from privately owned vehicles.

For example, when researchers asked if a major automaker used data anonymization so it could research aggregated data about all models in a specific model year, an original equipment manufacturer (OEM) executive stated, “Absolutely.” He added, “It is part of the consumer agreement that we are able to do this.” When asked about the debate concerning whether data anonymization actually works, he observed, “Both camps are right. I can assure you that I can collect vehicle location data and sufficiently anonymize it so that no one can get it. I can’t guarantee that someone driving with a smartphone isn’t giving up their information and as a result of combining data, the individual’s identity could be found.”

A transportation expert observed that INRIX and other data clearinghouses seem to have collected data with a lot of PII. When asked about whether data anonymization could minimize privacy concerns, he commented, “Google has more information about you than your best friend.” He said that Google looks at every key stroke, knows users’ social network and what they are talking about, and cross-references all this information. He concluded, “There is no such thing as data anonymization.”

A senior executive at an automobile association noted, “There is only so ‘anonymous’ that you can get…eventually, the data can be reconstructed.” She suggested, “We need a sort of bill of rights for consumer data.” She said that best practices will not be helpful unless businesses and consumers follow them, so it is important to ensure that consumers know about the best practices and adhere to them. She observed that the data world is a new place to be for many companies, and it is incumbent on them to ensure they protect consumer data, treat it the way they are required to, and follow best practices to protect themselves.

The debate over whether and how to anonymize consumer data like PII has been complicated by a recent Massachusetts Institute of Technology (MIT) research study. The study found that four pieces of spatiotemporal financial metadata (data showing the time and location of a financial transaction) could identify 90 percent of individuals, despite the data set lacking any names or other traditional identifying characteristics (12). This study raises troubling questions for the efficacy of protecting PII as a regulatory framework, especially for vehicular and transportation data. Transportation data frequently include spatiotemporal data, which can be used to identify individuals with relative ease. The protection of PII as a legal/regulatory concept relies on the idea that anonymized data do not require the same high level of controls as information with PII characteristics. If location and time are, by themselves, intrinsically identifying, then many sources of transportation data may need stricter protections than previously thought. Furthermore, if the volume and spread of data continue to grow, anonymizing transportation data (especially with location information) may be impossible.

**The Auto Industry Lacks Consensus Concerning How to Protect Drivers’ Data Privacy**

A February 2015 study performed by the office of Senator Edward Markey, *Tracking and Hacking: Security and Privacy Gaps Put American Drivers at Risk* (the Markey Report), outlines how the proliferation of new technologies such as wireless communications and automated and connected vehicle technologies raise new concerns about “the ability of hackers to gain access and control to the essential functions and features” of passenger automobiles and “the opportunity for others to utilize information on drivers’ habits for commercial purposes without the drivers’ knowledge or consent” (2). Senator Markey sent letters in 2013 to the major automobile manufacturers “to learn how prevalent these technologies are, what is being done to secure them against hacking attacks, and how personal driving information is managed” (13). The Markey Report discusses the responses from 16 major automobile manufacturers (2). The report’s findings that are the most relevant to data privacy are the following (2):

- Nearly 100% of cars on the market include wireless technologies that could pose vulnerabilities to hacking or privacy intrusions.
  - Automobile manufacturers collect large amounts of data on driving history and vehicle performance.
  - A majority of automakers offer technologies that collect and wirelessly transmit driving history data to data centers, including third-party data centers, and most do not describe effective means to secure the data.
  - Manufacturers use personal vehicle data in various ways, often vaguely to “improve the

*The manufacturers are BMW, Chrysler, Ford, General Motors, Honda, Hyundai, Jaguar, Land Rover, Mazda, Mercedes-Benz, Mitsubishi, Nissan, Porsche, Subaru, Toyota, Volkswagen (with Audi), and Volvo; see the Markey Report.*
customer experience” and usually involving third parties, and retention policies—how long they store information about drivers—vary considerably among manufacturers.

- Customers are often not explicitly made aware of data collection and, when they are, they often cannot opt out without disabling valuable features, such as navigation.

The report notes that “in response to the privacy concerns raised by Senator Markey and others, the two major coalitions of automobile manufacturers recently issued a voluntary set of privacy principles by which their members have agreed to abide” (2). These privacy principles address how consumer privacy may be protected by “ensuring transparency and choice, responsible use and security of data, and accountability” (2).

The research team learned from subject matter expert interviews that in addition to Bluetooth® readers and sensors that can detect and receive automotive information from outside the vehicle, the vehicle itself and applications within it can generate and transmit data.

The principles establish a framework for addressing privacy issues raised by new technologies. The privacy principles include (14):

- **Transparency:** Participating Members commit to providing Owners and Registered Users with ready access to clear, meaningful notices about the Participating Member’s collection, use, and sharing of Covered Information.
- **Choice:** Participating Members commit to offering Owners and Registered Users with certain choices regarding the collection, use, and sharing of Covered Information.
- **Respect for Context:** Participating Members commit to using and sharing 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.
- **Data Minimization, De-Identification, and Retention:** Participating Members commit to collecting Covered Information only as needed for legitimate business purposes. Participating Members commit to retaining Covered Information no longer than they determine necessary for legitimate business purposes.
- **Data Security:** Participating Members commit to implementing reasonable measures to protect Covered Information against loss and unauthorized access or use.
- **Integrity and Access:** Participating Members commit to implementing reasonable measures to maintain the accuracy of Covered Information and commit to giving Owners and Registered Users reasonable means to review and correct Personal Subscription Information.
- **Accountability:** Participating Members commit to taking reasonable steps to ensure that they and other entities that receive Covered Information adhere to the Principles.

However, as the Markey Report points out, the success of the principles in protecting consumer privacy will depend, in part, on how manufacturers interpret them. Specifically, the report notes that the ways in which transparency will be achieved are unclear (e.g., using text in the user manual) (2). The report also points out that “the provisions regarding choice for the consumer only address data sharing and do not refer to data collection in the first place,” and “the guidelines for data use, security, and accountability largely leave these matters to the discretion of the manufacturers” (2).

**Many Types of Data Are Gathered and Transmitted by Passenger Vehicles**

The research team learned from subject matter expert interviews that in addition to Bluetooth® readers and sensors that can detect and receive automotive information from outside the vehicle, the vehicle itself and applications within it can generate and transmit data. There is a lack of clarity about what data could be, or are, transmitted outside the vehicle.

There is also a lack of clarity about what external entities may be collecting data that are transmitted by the vehicle. When researchers asked an automobile asso-
ciation executive the question, “What vehicle data are capable of being transmitted,” she responded, “It isn’t always clear.” She stated that some information has to be made publicly available under federal law. Beyond that, she said, there is no law that automakers have to make vehicle information available. It is generated in a raw data form, and one has to have special equipment to collect it. The amount of information conveyed varies across the industry, and there are many different practices. There is uncertainty about both the vehicular data that are being generated and the information being transmitted. Some information can be accessed through the vehicle’s onboard diagnostic system, and other information is generated through other systems that could be transmitted. She said, “There is a huge amount of data that could be transmitted, but there is little clarification on what is actually being transmitted.” She added that there is no opt-in or opt-out choice for vehicle data at this point. She said that it is one thing to know what data are collected, but the purchase of the car is basically “opting you in.” She explained further that when a person purchases a car with a telematics system, he or she is agreeing to the terms of use for how the telematics company will use his or her data.

In terms of vehicle apps, the automobile association executive noted that the mobile app industry has grown rapidly over the last 10 years, but it is still in its infancy in the vehicle market. She said that there were “some pretty eyebrow-raising parts in [user] agreements.” She explained that it is not just the data a vehicle is generating but the data a person is generating that can be covered under the user agreement. So, if individuals are texting or calling someone on their smartphone, that could be covered under the user agreement. She added that third-party app developers are concerned about who owns the data. They may have a license to use someone’s data in many ways based on the user agreement he or she signed originally. She said that OEMs are getting data from all those apps, as well, and they are getting much more information about the driver and the users.

Consumers Are Willing to Trade Their Personal Data to Obtain a Benefit

Experts interviewed as part of this research frequently pointed to smartphone technology and the use of mobile app data for targeted retailing by third parties as illustrating a new frontier for consumer privacy issues. Specifically, the providers of navigation and telemetry systems, among others interviewed, explained how consumers agree to exchange their personal data for benefits. A telematics company executive told researchers, “People will create value by whatever means is available.” When asked if privacy concerns could interfere with the monetization of vehicular data, he said, “We are well past the privacy nervousness.” He added, “People have given up on it and can’t keep track of it all.” When asked if customers would consent to giving up their personal data if there was a sufficient benefit, he replied, “That is absolutely the case.”

Another telematics company executive said that each month, his company collects data on over a thousand diagnostic codes within the vehicle and sends them to the customer via email. He noted that app developers are beginning to build apps that can access the vehicle data and can bring the apps to market over time. These are “remote apps and in-vehicle apps so you can download them and run them in the vehicle,” he said. Like most
apps, these remote and in-vehicle apps probably will provide data such as location, age, gender, and other things back to the application program interface (API) developer. Currently, such app information is used for targeted marketing, providing customized services for users, and developing customized profiles for customers when smartphones and vehicles gather information on individuals.¹⁰

Eighty-Six Percent of Adults Surveyed by AAA Think There Should Be Laws and Policies to Protect Their Vehicle Data

AAA conducted a survey of over 1,000 U.S. adults in 2013 that contains at least two important findings for policy makers (4). When survey respondents were asked whether they agreed or disagreed with the following statement—“Consumers should always be able to decide if information generated by their car can be shared and with whom”—79 percent of the survey respondents agreed, and 66 percent strongly agreed. When survey respondents were asked if they agreed or disagreed with the statement, “There should be laws and policies to protect consumers’ rights to the information generated and captured by their vehicles,” 86 percent of the survey respondents agreed, and 69 percent strongly agreed. A 2013 survey of 2000 adults by the Auto Alliance found that privacy is an issue for consumers, with 75 percent of survey respondents indicating that they were very/somewhat concerned that companies would collect data from the software operating self-driving cars (3).

Despite the fact that many Americans routinely provide their personal data in exchange for a benefit provided by a mobile app on their smartphones, the responses to a recent AAA survey indicated that Americans might be less willing to share the information generated by their cars.

This chapter of the report considered whether and how data anonymization can contribute to protecting consumers’ data privacy, as well as a recent MIT study that indicates that data anonymization techniques may not actually protect PII. Researchers explored how automakers are dealing with the challenge of protecting drivers’ privacy, given the advent of new automated and connected vehicle technologies. The subject matter expert interviews provided insights about how Americans are currently trading their personal data for perceived benefits, via their smartphones and their automobile navigation and telemetry applications. Despite the reported widespread exchange of personal data for perceived benefits, recent AAA and Auto Alliance surveys show that many individuals are concerned about who can obtain the information generated by their car, and a large majority of those surveyed by AAA support the idea of laws and policies to protect consumers’ rights to the information generated and captured by their vehicles. For this reason, the next chapter focuses on transportation-related PII in Texas as a data privacy case study.

---

Chapter Five: Case Study of Transportation-Related PII in Texas

The 2015 Texas legislative session highlighted Texas policy makers’ concerns about protecting citizens’ PII. Texas legislators introduced seven measures that would address the protection of PII, including protection of the data privacy of certain populations (e.g., juveniles) or protection of citizens’ data privacy in specific circumstances (e.g., the use of location data in search warrants). Three proposed bills addressed security breach notification, removal of PII from car crash information the state maintains, and a requirement for a search warrant to obtain location information.

The following summary of research concerning transportation-related data and PII provides a case study about policy issues concerning new technologies. First, researchers summarize the results of interviews with stakeholders who represent different perspectives on data collection and use by states currently, as well as the use of data related to automated and connected vehicles as these technologies develop. These stakeholders include OEMs, the insurance industry, Texas transportation industry experts, toll authority representatives, privacy experts, privacy foundations, automobile industry associations, telematics developers and suppliers, state department of transportation (DOT) experts, and infotainment industry experts.

The summary addresses the following questions:
- How are vehicular data being collected in Texas?
- How are transportation data being used?
- Who owns the data?
- What is the value of vehicular data?

Next, the chapter considers policy issues relating to automated and connected vehicles, specifically how the federal government intends to protect automated and connected vehicle data privacy. Researchers assess state legislation and state efforts to address automated and connected vehicle data privacy following the discussion of federal actions. Finally, researchers identify key policy considerations Texas can draw from this case study.

How Are Vehicular Data Being Collected in Texas?

The research team learned from the interviews that states currently collect data from vehicles in a variety of ways, including transponders, Bluetooth readers, cameras, and sensors at the side of the roadway. Toll authorities obtain data from radio-frequency identification tags and cameras that snap photographs of...
license plates as vehicles pass through toll gates. State and local police and agencies may obtain data from automated license plate readers.

A transportation expert discussed data collection activities across several hundred miles of I-35 in Texas right now, and Bluetooth sensors have been installed throughout the region. Wavetronic detector cameras have been installed at some locations. Bluetooth devices can collect the media access control (MAC) address from digital equipment, like a cell phone, GPS device, smartphone, or laptop, that passes by the Bluetooth reader. The Bluetooth reader grabs the MAC address wirelessly and detects where the vehicle is located. Another Bluetooth reader down the road detects that same signal and computes the time traveled and speed and puts them together. The transportation expert noted that MAC addresses are purposefully either dropped or truncated in the field so that they cannot be reconstructed.

The data obtained from passing vehicles are aggregated. The expert said that “hundreds of jurisdictions are collecting information.” He added that there are many people selling these devices and collecting data. When asked if there is a state protocol about how to handle this vehicular data, the expert replied, “Everyone is doing that based on their own guidelines.” He commented, “The PII never makes it out of the field and into the central database.” When asked if commercial companies might be doing otherwise, he said, “Possibly.”

A state transportation expert noted that the equipment for a Bluetooth reader can be “cobbled together for about $200,” and is much more open and not proprietary. He added that state DOTs want to collect vehicular data, either because they think they can do it better or they do not want to purchase the data from the private sector.

A Texas Department of Transportation (TxDOT) expert explained that data are collected via cameras that read the license plates, which are then linked to Texas Department of Motor Vehicle billing records. He stated that TxDOT maintains the data indefinitely.

A state transportation expert said that with the advent of connected vehicles, the data coming out of vehicles “would communicate wirelessly” to roadside equipment. He added that data collection for connected vehicles is in the test deployment stage, and federal officials, state officials, and vehicle manufacturers are working together to deploy the systems for data collection. He stated, “They have specific applications they wrote to gather and use the information.”

A privacy foundation expert explained, “We are looking at ways that governments are increasingly collecting data through things like tolls, garages, MAC addresses, Bluetooth signals, and so forth.” He noted that these types of data are often bundled together with other applications. He cautioned, “They provide one service but could reveal other sensitive information.”
How Are Transportation Data Being Used?
The research team learned from interviews that state and local transportation-related governmental entities currently gather many types of data for a variety of purposes, such as traffic management. Some of the data state and local entities collect include:

- Individual and aggregate origin and destination data.
- Individual and aggregate vehicular travel times.
- Individual and aggregate vehicular location information.
- License plate photographs.

Transportation data can come from a variety of sources. Some state and local governments install their own equipment to capture vehicular data with systems like roadside Bluetooth readers. Toll authorities gather information about the travel patterns of roadway users and have a link to PII through credit card billing information. Other states purchase vehicular data from commercial third parties, such as INRIX. Some of these data include location data, which may be considered PII. Other data have been anonymized and aggregated.

State officials noted that they look forward to obtaining data from connected and automated vehicles in order to know more about the occupants of each vehicle for tolling and high-occupancy vehicle (HOV) enforcement purposes. Several transportation experts explained that states would be unable to handle the deluge of data that would become available once connected and automated vehicles and vehicle-to-infrastructure systems have been widely deployed. According to the respondent, a state like Texas would potentially need to contract with third parties to collect, manage, analyze, store, and provide access to Texas drivers’ vehicular data, some of which may include PII. The following excerpts illustrate some of the current uses states have for vehicular data.

A state transportation expert noted that public agencies install Bluetooth readers on roadways and get real-time traffic data. Another state transportation expert stated that a lot of metropolitan planning organizations (MPOs) in Texas, as well as on the east and west coasts, are buying data from companies like INRIX. He explained that when INRIX provides data to the state or MPO, it provides aggregate data in five-minute or two-minute intervals. According to this expert, the INRIX data do not contain PII. He said, “INRIX collects a lot of data from truckers, school buses, taxis, and so forth. They know which one it is coming from, but when it is provided to the state, that is not important or included.” Another state transportation data expert commented, “Right now, states have sold the rights to collecting transportation data to third-party companies. Private companies are laying down air tubes to measure traffic flow on roads...states go back and ask for the data they collected to do traffic management.” He added that traffic management centers have a pipe of data from INRIX. He concluded, “If states are going to sell the rights of owning the data to a third-party company, it [comes] full circle back to the question of anonymizing data and making meaning out of it.”
Who Owns the Data?
The research team learned from interviews that there are widely divergent ideas about who owns, or can have access and control of, vehicular data. This is potentially an important issue for Texas policy makers. Texas drivers may believe that all data associated with their vehicle belong to them, but many transportation industry stakeholders believe they have the right to use—or own—that same data.

The data that will be generated by connected and automated vehicles may include PII such as location information and links to a vehicle identification number (VIN) or billing information. Many transportation industry stakeholders have plans to use and monetize the data, which may create consternation and concern among consumers. There may be calls for legislation about this issue as connected and automated vehicles deploy in the market. As the representative of an automobile association stated, consumers may demand control and choice about the uses of their vehicular data. The following excerpts from the interviews illustrate the diversity of positions about this important consumer issue.

A toll authority representative stated that in his view, there is customer-specific data, vehicle-generated data, and occupant-generated data. He compared these data sources to a cloud of data being generated on roadways. He agreed that it is fair to ask to what extent highway authorities can access the data. He stated that if onboard data about vehicle condition and speed are important for traffic management and safety, toll authorities should be allowed to tap into that data. He said, “If your car is on our road, we [should be able to] know about your system.” He added that states should assess if they want to tap into this model of data access and how they could do so. When asked about the ownership of transponder data that toll authorities collect, including additional information that may be available, he stated, “We view the information we collect as our data. We own it.”

The debate isn’t about who owns the data but what rights or limits should be placed on people who want to use the data.

When asked who will own the data automated and connected vehicles may generate, a member of a privacy organization responded, “The consumer!” He added that other people may have legitimate interests to use it. He stated, “The debate isn’t about who owns the data but what rights or limits should be placed on people who want to use the data.” Similarly, a representative of an automobile association said that owning data is secondary in comparison to other questions. Specifically, she stated, “What is more important is access and control.” She explained that what is important is that a vehicle owner knows what data are being generated and has the ability to control the generated data. She added that in a digital environment, ownership implies that “there is a piece you can keep.” In contrast, she explained, “You don’t necessarily want to own bits of data on different servers but be able to control the use of it on different servers.”

An automobile insurance industry representative stated, “Unless there is a signed agreement otherwise, the customer owns the data.” He added, “We have an end-user agreement, and the user has to agree that they know how the information is being used.” He explained that
it has to be clear that people are giving consent to how their vehicular data are going to be used. He said that the company requests a broad form of consent from customers that will allow it to develop products without having to go back repeatedly to customers to obtain consent for specific data uses. An insurance executive concluded, “We want the freedom to potentially monetize the data our customers provide similar to the way other market leaders do.”

**What Is the Value of Vehicular Data?**

The individuals interviewed told the research team that there is a rush to monetize vehicular data, and that entities as diverse as insurance companies, OEMs, telematics companies, and state governments are all interested in monetizing automotive data. These stakeholders will all have strong views about any proposed state legislation concerning the management and use of vehicular data.

A state transportation expert said that “there are two low-hanging fruit: one is insurance, and the other is location-based services.” He added that there are many companies moving in that direction, primarily telematics companies.

One state transportation expert explained that sometimes data are used for planning, but primarily they are used for traffic management. He observed, “That is how connected vehicle information is going to be used on the private side.” He added, “They already have deals with auto manufacturers to get data off the cars.” He explained that companies are looking for new revenue sources and are trying to figure out how to spin this product. He said that some companies offer individual but anonymized data, such as origin/destination data. Another company is refining the average speed calculations that companies are already producing. He added, “Businesses are using this data to drive their own applications and reselling to other providers.” He said that businesses also resell these data to governments, and both companies and governments are doing this. He said that the data come from cars that talk to government-owned infrastructure sensors.

Another state transportation expert pointed out that most readers of data are public agencies, but some are private companies. He observed, “Generally, there is some sort of public-private partnership going on.” He said he was unsure whether the agreements between the parties protect PII. He stated, “I don’t think agencies are even thinking about this. What you do with the data could become a revenue stream for the private toll operator.”
Chapter Six: State and Federal Approaches to Automated and Connected Vehicle Data Privacy

This chapter presents a case study concerning automated and connected vehicle technologies. It first describes the actions the federal government has taken to address data privacy in automated and connected vehicles, and then describes actions states have taken to address these concerns.

Federal Efforts Concerning Automated and Connected Vehicle Data Privacy

**Automated Vehicles**

Since fully automated vehicles are not yet widely deployed, their likely effects on society are not well understood. Because the technology is not yet well developed, the federal government has taken a relatively cautious and slow approach to regulating the vehicles—choosing to allow the technology to develop before implementing regulations. This choice was intentional, and the National Highway Traffic Safety Administration (NHTSA) chose to do so because it did not wish to slow the development of the technology (15). NHTSA has provided guidance on state regulations concerning automated vehicles, although the agency does not address data privacy concerns in its brief report.

**Connected Vehicles**

The federal government led the development of the connected vehicle program and took several measures to protect users’ data privacy. In a technical report describing the readiness of the vehicle-to-vehicle (V2V) system, NHTSA discussed a variety of concerns related to data privacy (16). From the outset, NHTSA emphasized a few key points about the connected vehicle (CV) system envisioned by the program:

- The CV system will “not collect or store any data on individuals or individual vehicles.”
- Neither law enforcement nor private entities would be able to use the data in the basic safety message to identify a speeding or erratic driver.
- The CV system does not allow for tracking specific individual vehicles through space and time.
- The CV system does not collect information linked to individuals, including PII or personal communications.
- Vehicles are automatically enrolled, and enrollment does not require users to provide PII about themselves or their vehicle.
- The CV system does not “provide a pipe into the vehicle for extracting data,” but does enable OEMs and NHTSA to find “lots of production runs of potentially defective V2V equipment without use of VIN or other information that could identify specific drivers or vehicles.”
NHTSA worked with two groups representing a broad coalition of automobile manufacturers—the Crash Avoidance Metrics Partnership and the Vehicle-Infrastructure Integration Consortium—to address privacy concerns.

The federal government led the development of the connected vehicle program and took several measures to protect users’ data privacy.

The industry groups recommended that the federal government should play a central role in protecting individual privacy, but that the eventual security system should conform to an industry privacy policy framework that requires a variety of protections. These include protections such as:

- End-to-end anonymity for privately owned/leased vehicles and occupants for all mandatory V2V technologies, including security system processes and mandatory applications and services.
- Protection from attacks on system integrity, including from hackers and system administrators (i.e., insiders), by:
  - Providing secure, end-to-end encryption of vulnerable communications.
  - Changing short-term security certificates and vehicle identification every few minutes to prevent location tracking.
  - Assigning certificate signing requests—now called enrollment certificates or long-term certificates—in an anonymous fashion.
  - Separating security, legal, and administrative systems to prevent misuse or abuse of information.
  - Securing data to prevent inappropriate access to PII or identifying vehicle data (like VINs) (16).

NHTSA stated that the industry-recommended framework provided “an initial framework and useful starting point for development of privacy-protective V2V technologies,” but that recent “mission-critical and system-specific safety information needs that affect system privacy...have necessitated development of various additional controls to mitigate adverse privacy impacts” (16).

These controls have made some tenets of the original privacy framework infeasible, like the requirement for complete data anonymity. NHTSA pointed out that to identify and recall defective devices, there needs to be a mechanism to identify such defects and the vehicles on which they occur.

Future assessments of the privacy controls built into the connected vehicle system will be based on a set of principles, entitled the Fair Information Practice Principles, which have roots in the Privacy Act of 1974. The eight principles include:

- Transparency.
- Individual participation and redress.
- Purpose specification.
- Data minimization.
- Use limitation.
- Data quality and integrity.
- Security.
- Accountability and auditing.

These principles, and questions relating to each, will be a component of a technical assessment (known as a Privacy Impact Assessment) of how the connected vehicle system will affect privacy. The Privacy Impact Assessment will be published alongside the agency’s eventual Notice of Proposed Rulemaking concerning V2V communications technology.
What Information Is Collected, Shared, and Stored through the Connected Vehicle System?

Connected vehicle system data collection, transmission, storage, and sharing can be broken down into two categories: data used for system safety and data used for system security. The safety functions require the transmission of a basic safety message (BSM), which contains information about (16):

- Vehicle position, heading, speed, and other information relating to vehicle state and predicted path. The BSM, however, contains no personally identifying information (PII) and is broadcast in a very limited geographical range, typically less than 1 km. Nearby motor vehicles will use that information to warn drivers of crash-imminent situations. Except in the case of malfunction, the system will not collect and motor vehicles will not store the messages sent or received data sent/received by V2V devices.

The security system also requires the transmission of information through a security certificate management system (SCMS). The SCMS allows vehicles and messages to be certified as authentic by checking the vehicle and messages' certificates against the SCMS bank of verified messages. Additionally, the communications between system actors “are encrypted and subject to additional security measures designed to prevent SCMS insiders and others from unauthorized access to information that might enable linkage of BSM data or security credentials to specific motor vehicles” (16).

While the federal government is taking steps to guard citizens’ privacy, as automated and connected vehicles begin to deploy in the U.S. vehicle fleet, states may consider further steps to guard citizens’ privacy in a highly connected, ubiquitous vehicular data environment.

State Actions to Address Automated and Connected Vehicle Data Privacy

The connected vehicular program is primarily an effort of the federal government. The U.S. Department of Transportation (USDOT) spearheaded the research and development of this program, and states have thus far chosen a relatively limited role in addressing both the technical and policy goals. The research team is unaware of any specific legislation that states have passed to mitigate privacy concerns surrounding the USDOT connected vehicle program.12

As of this writing, four states and the District of Columbia have passed laws regulating the testing of autonomous vehicles, but there are very few policies specifically relating to data privacy protections.13 California, Nevada, Michigan, Florida, and Washington, D.C., have all passed legislation authorizing autonomous vehicle use (for testing, operation, or both), but only California and Nevada have addressed issues related to data collection and use.

As of this writing, four states and the District of Columbia have passed laws regulating the testing of autonomous vehicles, but there are very few policies specifically relating to data privacy protections.

California's legislation (SB 1298) mandates that automated vehicle manufacturers disclose and describe the information collected by autonomous vehicle technology. California and Nevada both require autonomous test vehicles to have a separate device that functions similarly to, but is distinct from, an event data recorder; the device must record the autonomous system information for 30 seconds preceding a crash and preserve it for three years. Nevada's legislation (AB 511) requires that data must be maintained for three years following a collision. The Oregon draft legislation (HB 2428) introduced in January 2013 established requirements (similar to California’s) for data recording and disclosure.14 Other than these limited measures, there does not appear to be any legislative provisions that directly address data collection or privacy at the state level in regard to automated or connected vehicles.

12If USDOT’s plans for connected vehicles become agency regulations, state governments or other entities may use their authority to take action to directly ensure that connected vehicle users on state roads have properly maintained their vehicles so that they can participate in the security infrastructure for V2V communications. See S. P. Wood, J. Chang, T. Healy, and J. Wood, “The Potential Regulatory Challenges of Increasingly Autonomous Motor Vehicles,” Santa Clara Law Review, Vol. 52, 2012, p. 1423.
13Nevada, California, and Florida refer specifically to autonomous vehicles.
14Texas legislation (HB 2932) introduced in March 2013 did not include any provisions concerning data privacy.
Chapter Seven: Policy Considerations for Texas

The research team interviewed a cross section of experts in Texas and other states to acquire insights into the emerging policy landscape as connected and automated vehicles are deployed.

This chapter of the report explores two important issues:

- Whether there is a need for new policies to manage vehicular and transportation data collected and used by the state as connected and automated vehicles are deployed.
- Whether there should be a state, federal, or combined approach to managing vehicular data issues from connected and automated vehicles.

Is There a Need for New Policies to Manage Vehicular and Transportation Data Collected and Used by the State as Connected and Automated Vehicles Are Deployed?

The interviewees all thought that new guidelines and policies would be needed to manage the data related to connected and automated vehicles. The respondents had different perspectives about what would be needed, from managing the security certificates associated with connected vehicles to establishing public-private partnerships to manage the amount of data collected. The majority of respondents indicated that states need to be thinking about both technical and policy issues related to data from automated and connected vehicles, as well as the need for uniform policies from state to state. There was no consensus about what policies are needed, but they ranged from policies concerning how to manage large quantities of data to protecting PII that may be embedded in data used to make roads more efficient.

Several transportation experts noted that states would be unable to handle the deluge of data that would become available once connected and automated vehicles have been widely deployed. According to the respondents, a state like Texas would potentially need to contract with third parties to collect, manage, analyze, store, and provide access to Texas drivers’ vehicular and transportation data, some of which may include PII. New state policies about how Texans’ vehicular and transportation data may be collected, managed, and stored by third parties—perhaps a public-private partnership—may need to be developed.

There also may be a need for a new state policy concerning how to handle and maintain security certificates for connected vehicles. Some of the experts interviewed thought it should be a state function but were concerned that if states managed the certificates, one state’s certificate might not be compatible with other states’ certificates. A state transportation expert commented, “If states own the data, they will have to come up with guidelines, rules, and regulations to govern the use of data.” He explained that although the federal government can provide guidelines or policy recommendations, given state
autonomy, it is likely that states will be able to set up their own ways of handling the data, including setting up policy and guidelines.

Another potentially needed policy initiative concerns how Texas manages traffic data containing PII. For example, state officials said that they look forward to obtaining data from connected and automated vehicles in order to identify the occupants of each vehicle for tolling and HOV enforcement purposes (16). A state transportation expert explained that the amount of data obtainable in the near future from automated and connected vehicles will grow quickly. Another state transportation expert cautioned, “I think we need to establish and maintain trust with the public by not doing foolish things, by having established procedures, and by being up-front with how we collect and use [their] data.” When asked about how his organization handled destruction or transmission of transportation data, another state transportation official replied, “It is informal at this point.” He added, “We need to be careful, or it will come back to bite us.”

Should There Be a State or Federal Approach to Managing Vehicular Data Issues from Connected and Automated Vehicles?

Respondents held differing views about whether the states or the federal government should take the lead in establishing minimum requirements, guidelines, or regulations concerning connected and automated vehicle data. A recurring theme, however, was the need for uniformity of state regulation to avoid the problems created by inconsistencies of different state legislation, both for OEMs and state governments. Several respondents concluded that action needs to be taken at both the state and federal levels to address data issues from connected and automated vehicles.

A state transportation official suggested that whatever regulations or standards are put in place should directly specify the purpose they are designed to serve. He said that it makes sense that some connected vehicle data would be used to make roads safer and more efficient, but the purpose should be very specific about how the data are to be used. He added, “I think having a federal standard would be beneficial over time because tolling and other transportation solutions providers need consistency.” He observed that if there is a call for interoperability for tolling, all the different vendors that support states and regional vendors have guidelines, and having it mandated would make interoperability easier.

The same official explained that the minimum requirements should be federally mandated, but if the states want additional requirements for their area, that could be reasonable as well. He described having standards so that “if you take your vehicle from state X to state Y, your vehicle will be compatible.” He added that best practices for data collection and storage might be needed.

“I think we need to establish and maintain trust with the public by not doing foolish things, by having established procedures, and by being up-front with how we collect and use [their] data.”

Another state transportation expert noted, “It has to be a federal perspective.” He explained that “the rules have to be applied across all 50 states.” He pointed out that a vehicle originating in Colorado has to be able to request data in Texas the same way as in Colorado, and that requires creating universal federal standards. OEMs, the insurance industry, and national automobile association interviewees echoed concerns about the negative impact of having a patchwork of state rules and regulations for managing data from automated and connected vehicles (17).

15According to NHTSA, it does not appear that such information will actually be provided to law enforcement through the connected vehicle system.
16Simultaneously, some stakeholders with whom researchers spoke expressed concerns that premature regulation on the state level may result in a “crazy quilt of different, and perhaps incompatible, requirements. Attempting to meet state regulations could increase costs and make the technology uneconomical.” Similarly, NHTSA recently expressed caution about regulations on the state level, noting that “Particularly in light of the rapid evolution and wide variations in self-driving technologies, we do not believe that detailed regulation of these technologies is feasible at this time at the federal or state level” (18).
Chapter Eight: Findings

The protection of PII is a current topic of interest for Texas, as well as federal policy makers. This is illustrated by the wide variety of PII-related proposed legislation in the Texas 2015 legislative session, which included bills about security breach notification and prosecution of the offense of breach of computer security, removal of PII from car crash information maintained by the state, removal of PII from state juvenile court records, removal of PII from announcements about lottery winners, handling of biometric information by state agencies, handling of medical data with PII by state agencies, and a requirement for search warrants to obtain location information. In addition, in January 2015, the White House announced several new data privacy legislative proposals. The research detailed in this report addressed specific concerns about the definition of PII in Texas, especially in the context of emerging technologies. One example is the capture and use of data about an individual’s location by a smartphone or by automated and connected vehicle technology.

The research team found that privacy protections for PII are siloed at the federal and state levels. In addition, there are different definitions of PII in Texas, as well as nationally and internationally, which can lead to legal confusion and discord. This research revealed several areas concerning PII that Texas might address to prevent problems arising from new technologies in the near future.

There Is Not a Consistent Statutory Definition of “Personal Identifying Information”; Section 521.002 of the Texas Business and Commerce Code Appears to Be the Most Encompassing Definition of PII for Texas

Currently, different definitions are used in segmented approaches to privacy protection in Texas business, state agency, and criminal law contexts. For example, Section 521.002 of the Texas Business and Commerce Code defines personal identifying information as information that “alone or in conjunction with other information identifies an individual.” The Texas Business and Commerce Code also defines many of the same data elements as sensitive personal information. The Texas Penal Code criminalizes the fraudulent use or possession of PII, which is referred to in Section 32.51 of Title 7 as “identifying information.” The potential result of inconsistent definitions of PII might be a failure to protect important PII, such as location data, which is not specifically addressed by any of these provisions. Similarly, it is unclear how the fraudulent use or possession of location data might be prosecuted under the Texas Penal Code.
This research provided specific examples of why a consistent definition of PII is important. Researchers learned that many different businesses are currently collecting and selling vehicular transportation data, which may contain PII. In addition, many businesses—OEMs, insurance companies, and telematics providers—are planning on gathering and combining vehicular data concerning automated and connected vehicles, which also may contain PII. The PII may include the location information from an individual’s vehicle, a vehicle’s license tag, or the name and address of the vehicle’s owner. Toll authorities have the capability of matching vehicle origin and destination information with an individual’s credit card information, concerning the collection, use, management, retention, and destruction of vehicular and transportation data that may contain PII.

**The Adequacy of Data Anonymization as the Basis for Privacy Protection of PII Gathered by State Agencies, for Transportation or Other Uses, May Need to Be Reconsidered**

The debate over whether and how to anonymize consumer data like PII has been complicated by a recent MIT research study that found that four pieces of spatiotemporal financial metadata (data showing the time and location of a financial transaction) could identify 90 percent of individuals, despite the data set lacking any names or other traditional identifying characteristics (12). Transportation data frequently include spatiotemporal data, which can be used to identify individuals with relative ease. Policy makers may need to revise their assumption that anonymized data do not require the same high level of controls as information with personally identifying characteristics. If location and time are, by themselves, intrinsically identifying, then many sources of transportation data may need stricter protections than previously thought.

**Stakeholders Are Uncertain about Who Owns the Data from Vehicles**

The research team also identified several important findings about automated and connected vehicle data. For instance, researchers learned from interviews that there are widely divergent ideas about who owns, or can have access and control of, vehicular data. This is potentially an important issue for Texas policy makers. Texas drivers may believe that all data associated with their vehicle belong to them, but many transportation industry stakeholders believe they have the right to use—or own—that same data.

**There Is a Rush by Stakeholders to Monetize Vehicular Data from Automated and Connected Vehicles**

The individuals interviewed noted that entities as diverse as insurance companies, OEMs, and telematics companies are all interested in monetizing vehicular data from automated and connected vehicle technologies. These
stakeholders are likely to have strong views about any proposed state legislation concerning the management and use of vehicular data.

State Agencies Have Made Few Preparations to Deal with the Extremely Large Volume of Vehicular Data That Experts Predict Will Be Produced by Automated and Connected Vehicle Technologies

The state transportation experts interviewed were unanimous in observing that state agencies have made few preparations to deal with the enormous amount of data that will be produced by automated and connected vehicle technologies. A state transportation expert interviewee explained that the data that will be obtained in the near future from automated and connected vehicles will grow quickly. Other state transportation experts explained the need to establish and maintain trust with the public by having established procedures and by being transparent about how state agencies collect and use drivers’ data.

An OEM executive revealed that remote APIs and in-vehicle APIs will soon permit a host of applications in a vehicle, much like a smartphone. Like apps on smart-phones, however, those apps will deliver information about their users to API developers. As with smartphone apps, these data may be sold to targeted marketing companies. Drivers may be profiled and the data used to market to them while they are driving.

Once vehicle-to-infrastructure technology is deployed, states may be overwhelmed by the amount of data they receive and may need to seek third-party solutions to manage the data. One transportation expert interviewed estimated that the amount of data that states would receive “would be in the terabytes.”

There may be a call to standardize the process for anonymizing toll reader data. One of the state transportation experts noted, “The upside is obviously interoperability; the downside is that you are forcing everyone to conform to a universal standard, which might impose costs on toll authorities trying to comply with a standard.” He added that standards make it difficult in the early stages of a product because the industry wants innovation and does not want to hamstring development. He explained that once a technology becomes mature, as with MAC address readers, it makes sense to start applying standards.

The research team learned from two of the respondents that automobile insurance companies and trucking fleets are very interested in vehicular data. Auto insurance companies are already using vehicular data to assess driver risk, and trucking fleet management may be using vehicular data in its operations soon. In several interviews, respondents noted that consumers are more willing to trust private companies than the government. They suggested that consumers are willing to give up their personal data in return for benefits and services. One of the respondents commented, “From my perspective, people are giving up their civil liberties and haven’t necessarily thought through what the consequences are.”
References


