Benchmarking Consumer Data and Privacy Knowledge in Connected and Autonomous Vehicles

Flora Barber and Steven Furnell
School of Computer Science, University of Nottingham, U.K.

Keywords: Connected and Autonomous Vehicles, Data Privacy, User Acceptance, Consumer Behaviour.

Abstract: Connected and Autonomous Vehicles (CAVs) and their features are integrating into the conventional personal vehicle market, irrevocably transforming the definition of a vehicle. However, consumers have been routinely omitted from stakeholder research and their understanding of CAV’s data implications has been understudied. This paper addresses this through benchmarking the consumer’s current data and privacy knowledge with a survey, focus group, and analysis of privacy provisions available to consumers from manufacturers, where it found the materials insufficient. Using thematic analysis, this consultation of 168 survey respondents from 14 countries established the consumer’s need to be ‘Informed’, with further sub-themes of ‘Given Information’, ‘Information Requirements’, ‘Privacy Communications’, and ‘Privacy Control’. A follow-up focus group of 6 participants identified a further four themes of ‘Disinterest’, ‘Distrust’, ‘Impact’, and ‘Vehicle Perception’. This paper recommends industry prioritisation of consumer education and engagement with data privacy to maximise public trust, including the introduction of vehicle specific data protection legislation, government level assurance of manufacturer compliance, and use of the manufacturer’s app to control privacy. Consumers purchasing a vehicle must be made aware of its data transmission, collection, and protection technologies.

1 INTRODUCTION

As connectivity and autonomy are newer additions to vehicular design, concerns have been raised by researchers that security has become an afterthought (Karnouskos & Kerschbaum, 2017) (Strandberg, Olovsson & Jonsson, 2018). With autonomous vehicles collecting a gigabyte of data per second (Boom, 2015) and monetization of this data forecast to be worth $750 billion by 2030 (Bertoncello, Camplone, Gao, Kaas, Mohr, Moller & Wee 2016), 45% of new buyers express concern about the detriment to their privacy that these new technologies have (Dean, 2017). Consumers are already challenged to understand the data privacy options available to them on the devices they currently use. It is, therefore, vital to take a consumer-centric approach and consult the stakeholders themselves in order to ascertain their knowledge and improve the public’s confidence in Connected and Autonomous Vehicles (CAVs). In order for their deployment to be a success consumers and users of the vehicles must be allowed to make informed decisions about their data. This paper aims to address the above gaps by creating a consumer data and privacy knowledge benchmark through consumer consultation.

The study evaluates the consumer’s awareness, understanding, and recognition of data-collecting CAV features in their own vehicles, their experience of their vehicle manufacturer’s privacy materials, and what they value as important to improving consumer engagement with vehicular data privacy. The findings suggest that current privacy provisions and materials insufficiently engage and inform consumers about vehicular data use and collection. The consumer’s understanding has not kept up with the pace of innovation that is enabling once isolated vehicles to become more connected and autonomous. The participants suggested a variety of approaches to engage consumers with their vehicular privacy and to build trust in manufacturers. The findings compliment those from interviews with CAV experts about cyber security and privacy in CAVs (Liu et al. 2020), and the more generalised study by Maeng et al. (2021) into consumers’ attitudes towards CAV information security threats.

Following an overview of related work, is an examination of privacy-related materials from vehicle manufacturers. The methodology for survey
and focus group activities are described in Section 4, with results then presented in Section 5. The paper concludes with a series of related recommendations.

2 RELATED WORK

Drivers can be fingerprinted with 100% accuracy solely on 8 minutes of brake pedal data (Enev et al., 2016), purely acceleration data (Virojboonkiate et al., 2017), a combination of sensors (Pesé & Shin, 2019), or mapping the location of journeys without GPS either through fog nodes data near the vehicle’s journey (Butt, Iqbal, Salah, Aloqaily, & Jararweh, 2019) or from vehicle speed, waiting at traffic lights, and turns (Bellatti et al., 2017). This reveals that users can be identified by data that is not classed as personal under current GDPR regulations. In light of the Facebook-Cambridge Analytica data scandal and the multimillion fines against technology corporations for breaching data protection rules (Beato, 2013), consumers are wary of their data’s security, from it being sold to third parties to turning the relatively anonymous and private space of a vehicle into a means of surveillance to profile and predict their behaviour (Collingwood, 2017) (Glancy, 2012).

Current automotive manufacturer privacy policies fail to define the “legitimate business purposes” used as a reason for collecting data (Booz Allen Hamilton, 2019). Further research has found that no original equipment manufacturer (OEM) details the data it is collecting, who has access to or uses it, the security in place to protect it, or that real time querying may occur unknown to the consumer, despite researchers discovering that this data could be accessed via the vehicle’s VIN at a car dealer (Frassinelli et al., 2020).

The importance of CAV consumer training has been identified, but not prioritised, by the United Nations Economic and Social Council (ECOSOC) World Forum for Harmonization of Vehicle Regulations (UNECE, 2019). The Society of Motor Manufacturers and Traders has called on the UK Government to provide consumers with materials to increase public confidence in industry, data privacy, and the safety provisions of CAVs (SMMT, 2017).

Consumer trust, readiness, and acceptability is one of 10 priority areas that has been identified by researchers as imperative to the success of CAVs (Nikitas, 2020). The proposed assurance framework for assessing a CAV’s cyber security level, known as the 5StarS initiative, is designed to support consumers and insurers in understanding the cyber security risk for vehicles that have been independently tested under the framework, yet omits consumers from its stakeholder research (5StarS, 2019). Consumers are at risk of their data being targeted by hackers for purposes of extortion, increasing the credibility of targeted social engineering attacks, burglary, and exploitation as a back door into companies for intellectual property or data theft (Kam, 2016).

As the average vehicle life span is 13.9 years, a figure exceeding that of many operating systems, new vehicle specific security systems must be flexible to change and work consistently to protect the vehicle user’s data (SMMT, 2016). Researchers propose vehicle specific solutions such as a Differentially Private Data Streaming (DPDS) system to address privacy weakness in distributed edge computing, guaranteeing privacy levels over time as well as when vehicles dynamically move over time (Ghane et al., 2020), a start, predict, mitigate, and test (SPMT) system to predict and mitigate vulnerabilities systematically (Strandberg et al., 2018), and an architecture (CARAMEL) that detects attacks, provides in-vehicle anti-hacking measures, and real-time validation of the integrity of the vehicle’s data transmissions (Vitale et al., 2020).

Such solutions are part of a number of tools that need to be considered. It is crucial that regulation is brought up-to-date to reassure consumers and demonstrate respect for user privacy, ensuring that the consumer and users of CAVs have control over all aspects of their data (Collingwood, 2017) (Karnouskos & Kerschbaum, 2017).

3 DATA PRIVACY INFORMATION AVAILABLE TO CONSUMERS

It is vital to understand the resources currently available to consumers in order to contextualise their knowledge as benchmarked by this study. Six manufacturers (namely Audi, BMW, Ford, Tesla, Toyota and Volvo) were selected to represent a range of vehicles in production. These represent a selection of manufacturing groups from the top 15 ‘Most innovative Automotive OEMs of 2021’, as ranked by the Center of Automotive Management (CAM, 2021), and from the top 15 manufacturers by market capitalisation (Ghosh, 2021). The owner’s manual and privacy policies for these manufacturers were evaluated from a consumer’s perspective for their ease of use when locating privacy information, as well as the details covered in the material. All
documentation was manually evaluated using document analysis by one researcher. The vehicles chosen for analysis were:
- Audi A6 – 2021, Executive (Audi, 2021)
- BMW i3 Electric – 2015, Small Family Car (BMW, 2015)
- Ford Focus – 2021, Small Family Car (Ford, 2021)
- Tesla Model 3 – 2021, Large Family Car (Tesla, 2021)
- Toyota Corolla - 2020, Small Family Car (Toyota, 2020)
- Volvo XC40 – 2021, Small Off-Road (Volvo, 2021)

The findings are summarised in Table 1 and the parameters are grouped into the three main outlets of privacy information that vehicle manufacturers provide: the in-vehicle infotainment system, the vehicle handbook/owner’s manual, and the manufacturer’s website. The results for each parameter are based on the joint findings from the selected owner’s manuals and privacy policies. The infotainment system has three main parameters, for which the results were based on the information provided in the owner’s manuals. Access to an electronic copy of the owner’s manual was determined to establish the ways in which consumers can find privacy information in-vehicle. Access to privacy information and settings from the infotainment system determined if the consumer could control the data transmitted from their vehicle. The infotainment and manufacturer’s websites were jointly checked for software release notes availability. These notes are an important method of engaging consumers with their vehicle and with their data privacy by understanding the functions and abilities their vehicle possess and the cyber security protections in place. The manufacturer’s websites were judged for their signposting and ease of navigating the privacy policy. Lack of these factors may dissuade consumers from engaging with privacy information and weaken their privacy knowledge. The websites were also analysed for material that emphasised the importance of removing personal data from a vehicle before sale, thus protecting the consumer’s data. The owner’s manual was checked for the same emphasis as well as how to complete this procedure. The selected owner’s manuals were analysed for the inclusion of information about vehicular privacy, Event Data Recorders, and how to update the vehicle, including references to full copies of the manufacturer’s privacy policy. It is important that all data collecting and recording features in the vehicle are clearly explained to the consumer, as well as where they can access further privacy information. Vehicle software update procedures are important in maintaining the cyber security protections of the vehicle, protecting the consumer’s privacy and data.

Table 1: Summary of the privacy information available to the consumer from selected manufacturers.

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Manufacturer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Access privacy info and settings in-vehicle?</td>
<td>Y, P, Y, Y, N, Y</td>
</tr>
<tr>
<td>Software release notes available</td>
<td>P, N, N, Y, N, Y</td>
</tr>
<tr>
<td>Privacy information included</td>
<td>Y, P, Y, Y, P, Y</td>
</tr>
<tr>
<td>Dedicated chapter on data protection</td>
<td>Y, N, Y, N, N, Y</td>
</tr>
<tr>
<td>References on where to find full privacy policy</td>
<td>Y, N, Y, Y, N, Y</td>
</tr>
<tr>
<td>Event Data Recorder information</td>
<td>Y, Y, Y, P, P, Y</td>
</tr>
<tr>
<td>Includes how to remove personal data stored in vehicle</td>
<td>N, Y, Y, Y, N, Y</td>
</tr>
<tr>
<td>Information on how to update vehicle</td>
<td>Y, N, Y, Y, N, Y</td>
</tr>
<tr>
<td>Clear who is responsible for updating the vehicle</td>
<td>N, N, N, Y, N, N</td>
</tr>
<tr>
<td>Ease of privacy policy navigation</td>
<td>N, Y, N, Y, Y, Y</td>
</tr>
<tr>
<td>Emphasis on personal data removal</td>
<td>N, N, Y, N, Y, N</td>
</tr>
<tr>
<td>Software release notes available</td>
<td>N, Y, N, N, N, Y</td>
</tr>
</tbody>
</table>

Key: Y=Yes, N=No, P=Partially available depending on regions or vehicle, and manufacturers (A=Audi, B=BMW, F=Ford, Te=Tesla, To=Toyota, V=Volvo).

Owner’s manuals were checked for clear signposting to privacy information through the use of dedicated chapters detailing the vehicle’s data protections.
Audi’s owner’s manual contained multiple prompts to remove personal data before sale, and the privacy information was generally well written. However, the advice about software update responsibility was conflicting and the privacy policy was very difficult to find. BMW’s online offerings were much easier to navigate with hyperlinked buttons and subdivided sections. The owner’s manual was devoid of privacy information despite having ‘ConnectedDrive’ features. Ford’s manual contained a ‘Data Privacy’ chapter which was thorough and detailed. Only software updating responsibility was omitted. All of Ford’s online provisions are available from their one-stop resource ‘Terms & Privacy Policy Hub’ (Ford, 2021). Whilst very clear, the density of the documentation could be better subdivided with the use of hyperlinked sections. Only Tesla specified who is responsible for software updates, but they lacked a dedicated data privacy section in the manual. Tesla’s online provisions were extremely clear and organised to minimise information fatigue. Toyota’s website placed significant focus on deleting personal data before selling your vehicle, but this information, and some of the privacy policies inferred, were not easily found. Toyota’s owner’s manual provided the least amount of privacy information of those compared. In contrast, Volvo’s materials were very comprehensive throughout, including provision of a software release notes finder. However, Volvo did not make clear who should be responsible for updating the vehicle.

4 ASSESSING CONSUMER AWARENESS

Following on from ascertaining the information available to the public, this section details the survey and focus group consumer consultations. A thematic analysis approach was chosen to evaluate the resulting qualitative data, allowing for rich thematic discussions of consumer knowledge (Braun & Clarke, 2006). A primarily inductive analysis method was used to allow for data-driven results without a pre-existing coding framework, although it is acknowledged that aspects of deductive analysis were required to ensure the themes’ relevance. (Byrne, 2021). The analysis performed combines semantic and latent approaches to identifying meanings in data, recognising both the levels of explicit meaning and underlying assumptions that the respondents hold (Braun & Clarke, 2006). This approach is important to ascertaining how consumers understand privacy in the context of their vehicles and if the current provisions identified are effective or influential.

The wider contextual influences expressed on a latent level are important to establishing the reasoning behind the quantitative results of the survey. As only a single researcher coded and analysed the resulting data there was a significant risk of bias being introduced. This has been minimised through using Braun and Clarke’s six ‘Phases of Thematic Analysis’ to structure the process of coding and analysis (2006). The manual coding method was replaced by the use of NVivo 12 Pro as the themes became more numerous and more difficult to track. The software enabled a more flexible and detailed hierarchical organisation of themes, as well as a better adhesion to the six ‘Phases of Thematic Analysis’ (Braun & Clarke, 2006). This method also preserved responses that were divergent from themes with a greater number of coded references, which is imperative to creating a comprehensive benchmark that accurately reflects the market CAVs are entering.

All responses were anonymous, and no personal data was collected from respondents. Participants were recruited from social media, where the survey link and focus group were advertised from the researcher’s account. Convenience sampling was primarily used alongside snowballing sampling. Participant’s consent was obtained before both the survey and the focus group, and a pilot survey was conducted prior to the primary version.

4.1 Consumer Survey

The online consumer survey, titled ‘Surveying Vehicular Data Privacy and the Consumer’, consisted of seven sections: the participant information sheet, demographic details, the participant’s primary vehicle, privacy in relation to the primary vehicle, general privacy questions, improving current privacy provisions, and contact information for joining the virtual focus group. These sections aimed to evaluate consumer awareness of connected and autonomous features in their own vehicles, their current understanding and recognition of vehicular data collection and privacy, their experience of current privacy provisions and materials from manufacturers, and what is important to improving the consumer’s engagement with their data privacy.

Question branching was used to ensure the survey was asking suitable questions (e.g. not asking about experiences of manufacturer’s privacy policies if they had answered ‘No’ or ‘I am unsure of what that is’ to the question ‘Are you aware of what a privacy policy
Multiple choice questions with option shuffling to minimise bias was the primary question type used. The primary consumer survey of 28 questions was conducted from 5th – 19th August 2021, receiving 168 responses from 14 countries.

4.2 Focus Group

A small focus group was used to expand upon and investigate further the identified themes, generating a more detailed insight into consumer’s knowledge, differing from the interviews of CAV experts conducted by Liu et al (2020). The content of the focus group was semi-structured around key questions, developed from the data of the initial survey results, and a supporting presentation. The focus group began with more open questions and gradually increased the level of structure whilst allowing for spontaneous pursuit of any points raised of interest. The questions concluded with a highly structured scenario based question where two vehicles, a vehicle with and without CAV features, were compared under given circumstances.

The focus group was conducted in August 2021 with 6 participants and lasted 1 hour 15 minutes. The majority of participants identified as male, with only one participant identifying as female. All were from different undergraduate backgrounds and professions, including areas such as business, the humanities and sciences, environmental science, and the vehicle manufacturing industry. None of the participants were experts in the area of CAVs.

5 RESULTS

This section details and discusses the results of the consumer survey and the focus group, which incorporate quantitative survey results to support the primary qualitative thematic analysis.

5.1 Consumer Survey

Of the 168 respondents of the survey 69% drive a vehicle. Despite only 8% of drivers reporting that current privacy provisions are sufficient, only 5% of respondents who say they drive a vehicle with privacy settings have changed their in-vehicle settings, whilst the remaining 95% of respondents report never having changed or looked at such settings. 29% of respondents did not know if their vehicle had this optionality. As only 14% of drivers had read and 52% partially read their vehicle handbook, many may be unaware that such privacy controls exist. Groups with particularly low engagement with their vehicle handbook included drivers who neither own nor lease the primary vehicle they drive and those who drive monthly or less frequently than monthly. Despite the lack of engagement with the owner’s manual, it was the second most popular place (33%) respondents aware of what a privacy policy is said they would look for privacy information. Those who owned their primary vehicle and those who drive weekly were more likely to read a vehicle handbook.

The primary overarching theme of the question ‘What would help you feel in control of your data?’, was that respondents needed to be ‘Informed’. 85% of responses relate to the themes of ‘Given Information’, ‘Information Requirements’, ‘Privacy Communications’, and ‘Privacy Control’, which are summarised in Figure 1. 35% of all respondents wanted the information provided to specifically address how their data is being used, where it is stored, who has access to it, why it is being collected, and what is being collected from their vehicle. When asked which data types the respondents thought vehicle manufacturers collect from modern vehicles, the most chosen type was location data (77%). All the
data types listed are collected by manufacturers. The question’s full results are presented in Figure 2.

A similar question that asked respondents who drive about the features of their primary vehicle also highlighted this uncertainty. All the features listed in the question collected data from the vehicle. The ‘built-in SIM’ feature had the highest level of uncertainty (43% responded with “I don’t know”) concerning whether the primary vehicle had such a feature, with an average of 17% of respondents being unsure about any of the listed features. Drivers with a vehicle aged 5 years old or newer had a particularly high rate of uncertainty about the data collecting features of their vehicle, answering with “I don’t know” if their vehicle had the listed features to 23% of the listed features. There were no respondents who reported having a primary vehicle with all the features listed. These levels of uncertainty about the data collecting features of the respondent’s primary vehicle correlates with the 20% of respondents who wanted to know exactly what data was being collected in order to feel in control of their data. Whilst 28% of the primary vehicles reported in this survey were 11 years old or older and may currently only include few of the listed features, the average age of a vehicle at scrappage is only 13.9 years and therefore these drivers may soon be replacing their vehicle with one that may have such features (SMMT, 2016).

Privacy information provided should follow key guidelines, ensuring that the information is more visible and accessible to the consumer from the manufacturer’s website, written clearly and concisely in ‘layman’s terms’ using ‘simple language and expression’, and is without the use of ‘jargon’. Despite these responses requiring more concise and clear privacy information, ‘Brevity of policies’ was the least chosen factor regarding data use by manufacturers in the survey with only 20% of the 168 respondents regarding it as one of the most important factors to them. Transparency from the manufacturer at every stage was prioritised by as one of the most important factors to them. Privacy information being available from the website, and this being the preferred communication method of a third of respondents, only 3 had actually checked this source.

Participants also expressed a need to be able to control their data privacy from different places, such as in the vehicle, from a mobile device, and/or from an online account. 26% of respondents specifically noted the need for opt-in or opt-out options to enable them to control the data collected from their vehicles. These results correspond with 70% of respondents prioritising ‘Clear opt-out information’ as the second most important factor of their data use by vehicle manufacturers.

5.2 Focus Group

The four key themes, summarised in Figure 3, were identified from the focus group as: ‘Disinterest’, ‘Distrust’, ‘Impact’, and ‘Vehicle Perception’.

Participants expressed that disinterest forms from two distinct branches – uninterest in data privacy due to the technical wording and length of the current policies, and disinterest from not experiencing data misuse or that they are not at risk of harm if their data was misused. This correlates with the survey results, where only 57% of respondents who drive know what a privacy policy is, and of those only 3 respondents had read their vehicle manufacturer’s privacy policy. Of the 116 respondents who drive, only 2 had changed their in-vehicle privacy settings.

One participant described their reluctance towards CAVs as stemming from having ‘grown up knowing what the risks are with cyber’. Another participant suggested ‘some sort of tiered system’ for privacy settings, similar to that used for cookies on website, where each tier relates according to personal attitudes towards risk, as an approach to increasing engagement with data privacy. The group also
discussed using the vehicle manufacturer’s app as a more native environment for vehicular privacy controls, as one participant noted that drivers may not associate a vehicle with data privacy. The disjointed relationship between modern, connected vehicles and those that participants have grown up with again contributes to the belief that privacy concerns are not relevant for vehicles and their users.

Participants discussed the need for transparency, honesty, and frankness about the potential risks involved as a way of motivating consumers, as well as requiring that vehicle manufacturers to show what is being done to protect consumers’ data. Participants expressed that their distrust in manufacturers and third parties could be remedied through government assurance and through manufacturers using their apps to demonstrate the benefits of sharing data, such as ‘early intervention’ system for mechanical issues or the eCall system. One participant was concerned about the interdependence of the data transmitted used in other systems, such as the collection of voice command data and voice ID authentication for banking. Another was concerned about the impact of data theft or tracking for CAV users who are in witness protection or are being stalked. Participants who had experienced or grown up with CAVs were more comfortable with vehicular data collection.

The understanding of what a vehicle is carried a significant amount of uncertainty about the data that may be collected. What was an isolated system is now able to connect with other vehicles, infrastructure, and/or manufacturers as part of a wide range of services and features. The consumer’s understanding of this has not caught up with the fast pace of vehicle development, with associations of CAVs being limited to futuristic, expensive, or ‘flashy’ vehicles.

CAVs are representative of a scale of disruptive, pervasive, and integrated technologies that are present in vehicles both on the market and on the road, as well as those in concept. Vehicle manufacturers must ensure their privacy information is clearly visible, accessible, written with simple expression, provides examples, is transparent, and easily navigated through. This information must be accessible from multiple places, with recommendations for use in their mobile applications. Manufacturers should use their app to actively engage consumers and show the consumer how their data is being used. An opt-in, tiered system of privacy controls based on risk levels is recommended. Information about privacy and the data collection activities of a vehicle must be available at the time of the vehicle’s purchase. The manufacturer must make the consumer aware of how to remove personal data from their vehicle, how to change the privacy settings in their vehicle, how to find privacy information, and told who to contact regarding privacy questions or concerns. Extra support is recommended for those unfamiliar with connected vehicles. An in-vehicle and/or in-app walkthrough of the data transmitting features and privacy settings is recommended for all consumers purchasing a vehicle with CAV features.

Future research may further consider the data stored in-vehicle and on applications in the vehicle infotainment system, as well as the privacy issues that may be additionally added by the use of the vehicle manufacturer’s associated mobile applications. Future work may consider examining how other fields are attempting to engage the public with their cyber security and if any approaches may address the barriers respondents raised. Future research may also be conducted into mapping the changing data and privacy knowledge of consumers through repeating the survey and focus group at periodic intervals, especially as CAVs become more commonplace.

6 CONCLUSIONS AND RECOMMENDATIONS

CAVs are representative of a scale of disruptive, pervasive, and integrated technologies that are present in vehicles both on the market and on the road, as well as those in concept. Vehicle manufacturers must ensure their privacy information is clearly visible, accessible, written with simple expression, provides examples, is transparent, and easily navigated through. This information must be accessible from multiple places, with recommendations for use in their mobile applications. Manufacturers should use their app to actively engage consumers and show the consumer how their data is being used. An opt-in, tiered system of privacy controls based on risk levels is recommended. Information about privacy and the data collection activities of a vehicle must be available at the time of the vehicle’s purchase. The manufacturer must make the consumer aware of how to remove personal data from their vehicle, how to change the privacy settings in their vehicle, how to find privacy information, and told who to contact regarding privacy questions or concerns. Extra support is recommended for those unfamiliar with connected vehicles. An in-vehicle and/or in-app walkthrough of the data transmitting features and privacy settings is recommended for all consumers purchasing a vehicle with CAV features.

Future research may further consider the data stored in-vehicle and on applications in the vehicle infotainment system, as well as the privacy issues that may be additionally added by the use of the vehicle manufacturer’s associated mobile applications. Future work may consider examining how other fields are attempting to engage the public with their cyber security and if any approaches may address the barriers respondents raised. Future research may also be conducted into mapping the changing data and privacy knowledge of consumers through repeating the survey and focus group at periodic intervals, especially as CAVs become more commonplace.

REFERENCES


