Bridging Perception Gaps in Connected Vehicle Data Collection: Insights from Drivers, Passengers, and Pedestrians

Lachlan Moore[1,3,a)] Yinan Zhao[1] Rei Yamagishi[1] Tatsuya Mori[1,2,3]

Abstract: Modern vehicles collect extensive data through sensors and connected services, raising significant privacy concerns. Through an online survey, we investigate user perceptions of data collection within the connected vehicle ecosystem, examining how data type and user role influence comfort and privacy concerns. We find that biometric, audio, visual, and commercial information are perceived as highly sensitive, whereas vehicle information, general identifiers, and driver behavioral data are viewed as less intrusive. Comfort is influenced by familiarity and perceived ubiquity of collection, noting the importance of transparency and user awareness. Critically, perceptions vary across roles: drivers report higher comfort levels than passengers, and pedestrians are the most sensitive to data collection. These findings emphasize the need for role-aware communication and privacy mechanisms that align vehicle data practices with user expectations.

Keywords: Connected vehicles, vehicle companies, data collection, perception, user study

1. Introduction

As vehicles increasingly integrate connected systems, they have become complex data platforms. Connected vehicles (CV) capture a wide range of information, from telematics and drivers' behaviors to biometric, audio, and visual data, which may capture both vehicle occupants and individuals in the surrounding environment. The scope and sensitivity of the data collected raise important questions about privacy, transparency, and informed consent.

Car manufacturers emphasize the benefits of connected systems, such as improvement of safety, efficiency, and convenience. These gains come at the cost of opaque data practices [1]. Previous works indicate that users' awareness of what is collected and how it is applied remains limited [2]. In particular, comfort and expectations for data collection differ depending on one's position in this ecosystem: drivers, passengers, and pedestrians each face different privacy implications, and pedestrians are especially underrepresented in studies despite increased vulnerability [3]. This gap raises important questions about how perceptions of privacy vary across different roles within the ecosystem, and whether drivers should bear the burden of informing others about data practices inside vehicles.

To investigate these gaps, we ask the following research questions:

- **RQ1:** How do users within the vehicle data ecosystem perceive the types of data collected by CVs, how does this align with their comfort levels, and cultural background?
- **RQ2:** How does comfort with data collection vary in the vehicle ecosystem, such as drivers, passengers, and pedestrians, and how do these roles influence individuals' perceptions?
- **RQ3:** What factors influence individuals' willingness to opt out of data collection, change vehicles, and what types of data most strongly drive this intent?

We conducted a quantitative survey using the Qualtrics platform, gathering responses from 282 participants across the United States and Germany, two major regions of vehicle manufactory. The survey was designed to capture perceptions of CV data collection from individuals occupying different roles within the ecosystem, including drivers, passengers, and pedestrians. By asking participants to evaluate their comfort with various categories of data (biometric, visual, geolocation, and commercial information), as well as their expectations of data collection and willingness to opt out, we are able to generate a multidimensional understanding of how people experience CV technologies. Unlike prior studies that concentrate primarily on drivers [4], [5], our methodology intentionally broadens the scope to include passengers and pedestrians, recognizing that they too are subject to vehicle-based sensing and data transmission. This design not only provides comparative insights across roles but also highlights how

[1] Waseda University
[2] RIKEN AIP
[3] NICT
[a)] lachlan.moore@nsl.cs.waseda.ac.jp

individual perceptions may vary depending on one's position in the ecosystem.

This study contributes to the understanding of privacy perceptions in the CV ecosystem by examining how comfort with data collection varies across different user roles, including drivers, passengers, and pedestrians. It identifies the types of data that users find most sensitive and highlights the importance of transparency and role-aware communication in aligning vehicle data practices with user expectations. Additionally, by surveying participants across multiple cultural contexts, the work provides insights into how privacy perceptions may differ internationally, offering guidance for more user-centric and globally informed vehicle data policies.

2. Background

2.1 IoT Data Collection Perceptions

Prior research on the Internet of Things (IoT) provides important context for understanding privacy in connected vehicles, as these devices collect similar forms of data such as location, biometrics, audio/visual inputs, and behavioral traces. Studies consistently highlight a gap between user perceptions and actual data practices [6], as well as the persistence of the privacy paradox where users express concern yet continue adopting privacy-invasive technologies [7]. Other work has shown that users' discomfort often stems less from the raw data collected than from the potential inferences that can be drawn [8], and that attitudes vary across user groups, from privacy fundamentalists to pragmatists [9]. Similarly, research in smart home contexts shows that convenience and trust often overshadow awareness of risks [10].

2.2 Vehicle Data Collection Transparency

Extending the broader lens of IoT privacy into vehicles, prior research has demonstrated that automaker practices are marked by a lack of transparency. Our earlier analysis of privacy policies from 17 major manufacturers revealed widespread opacity in how both personal and sensor-derived data are collected and shared, exposing significant shortcomings in user awareness and policy clarity [11]. These findings align with broader concerns raised by watchdog organizations. For example, the Mozilla Foundation's *Privacy Not Included* review identified cars as the worst product category ever evaluated for privacy, with all 25 surveyed automakers failing to meet baseline standards of transparency and data protection [12]. Reports based on this analysis describe troubling practices, including the collection of deeply personal information such as sexual activity, facial expressions, and biometric or genetic data, often with limited user control or meaningful consent [13], [14].

Given vehicles' dual role as both essential transportation tools and sensor-rich, data-intensive IoT platforms, these shortcomings raise critical questions about how individuals perceive and evaluate data collection in this context. Unlike many IoT devices, vehicles capture information not only about their owners but also about passengers and even bystanders, expanding the privacy implications to a broader ecosystem. Investigating user perceptions across these roles represents a necessary step toward understanding public trust and informing more transparent, privacy-preserving data practices in the connected vehicle domain.

2.3 Vehicle Data Collection Perceptions

Public perceptions of connected and autonomous vehicles (CAVs) have been explored in a range of prior studies. Matin et al. [4] investigated Australian attitudes toward CAVs and found that safety concerns were a primary barrier to adoption, with nearly 70% of respondents reporting discomfort when driving alongside CAVs. Trust in technology emerged as a significant factor affecting acceptance. Similarly, Schoettle et al. [5] surveyed populations in the U.S., the U.K., and Australia, finding that roughly half of respondents held a generally positive view of autonomous vehicles, suggesting a mixed but cautiously optimistic public perception.

Research specifically examining perceptions of vehicle data collection is more limited. Acharya et al. [15] studied U.S. participants' views on connected vehicle data, revealing that acceptance decreased when respondents considered security and privacy issues, though this work primarily addressed vehicle-related data rather than personal or sensitive information. Schmidt et al. [16] examined public attitudes toward V2X technology, showing that participants were generally unwilling to share driver-related data, while vehicle-related information was considered less critical. Bloom et al. [17] investigated users' comfort with self-driving vehicle sensing and data collection, finding that a substantial proportion (54%) were willing to spend significant time opting out of identifiable data collection.

While prior research has explored cultural differences in privacy perceptions as well as role-based perspectives in vehicle or IoT contexts, these dimensions are often studied separately. Our work advances this line of inquiry by examining both simultaneously: we consider multiple roles within the vehicle data ecosystem, drivers, passengers, and pedestrians, across two countries with differing privacy cultures, the United States and Germany. By integrating insights from both vehicle and IoT privacy research, our study offers a more comprehensive understanding of comfort, expectations, and behavioral intentions related to CV data collection, highlighting cross-role and cross-cultural differences in a unified framework.

3. Methods

We conducted an online survey to quantitatively and systematically investigate the perceptions of users with regards to the data collection of vehicle companies. We explain the design, recruitment, and participants of the survey.

Fig. 1: An overview of the structure of the survey design.

3.1 Survey Design

Figure 1 gives a visual representation of the structure of the survey which consists of six parts, the informed consent, vehicle usage, general data collection comfort, vehicle related data collection: comfort and likelihood, privacy beliefs, and demographics.

Part–0: Informed Consent. The survey began with an informed consent process, which included a clear explanation of the study's purpose, the amount of compensation, the estimated time required for completion, and the procedures for data handling. Only participants who provided informed consent were permitted to proceed with the survey. Participants were also reminded that they could withdraw from the study at any time without penalty.

Part–1: Vehicle Usage. Participants were asked a series of questions regarding their use of and experience with vehicles, particularly connected vehicles. Specifically, we inquired about their familiarity with CV technologies, driving frequency, ownership status of the vehicle they use, experience with vehicle purchases, and how often they travel as passengers. These questions were designed to contextualize participants' experiences with vehicles, which may influence their perceptions of vehicle-related data collection practices.

Part–2: General Data Collection Comfort. We assessed participants' comfort levels with data collection across seven data categories using a five-point Likert scale:
(1) General identifiers (e.g., name, email, phone number)
(2) Commercial information (e.g., payment methods such as credit cards)
(3) Biometric information (e.g., fingerprints, facial or voice recognition)
(4) Location information (e.g., GPS data)
(5) Network and device activity (e.g., app usage, browsing history)
(6) Visual information (e.g., video recordings)
(7) Audio information (e.g., audio recordings)
Participants were also asked about their comfort with data being shared with third parties and the use of data to make inferences about them. The selection of data categories was informed by previous studies [17] and our own analysis, with a focus on categories commonly found in privacy policies [11] and easily interpretable by participants. Finally, we asked participants to identify which data category they would be least willing to share.

Part–3: Vehicle Related Data Collection Comfort and Likelihood. We asked participants to evaluate how likely they believed vehicle companies were to collect each of the previously listed data categories. Two additional categories were included:
(1) Vehicle information or Vehicle maintenance data.
(2) Driver behavioural information.
These categories are specific to vehicle use and were thus treated separately from general data collection. For each category, participants indicated both the perceived likelihood of the data being collected and their comfort with its collection. Importantly, we differentiated between data collection by the CV itself and by the vehicle company, reflecting the common separation of privacy policies between in-vehicle systems and broader corporate practices [11].

Participants were also asked to assess their comfort with data collection in three distinct scenarios: when acting as a driver, a passenger, and a pedestrian. In addition, we further probed their comfort with third-party data sharing.

Finally, participants were asked about their willingness to opt out of data collection services and whether data collection practices might influence their decision to change vehicles. Follow-up questions explored how long they would take to opt out and what types of data collection might prompt them to switch vehicle brands.

Part–4: Privacy Beliefs (IUIPC). Participants' technical privacy beliefs were measured using the Internet Users' Information Privacy Concerns (IUIPC) framework [18]. Following this framework we assessed participants with nine questions.

Part–5: Demographics. We collected demographic information including age, gender identity, education level, IT literacy, and country of residence. At the conclusion of the survey, participants were asked to confirm whether they had read and understood all questions to the best of their ability, following established practices for self-reported honesty checks [19], [20], [21].

3.2 Recruitment and Participants

We recruited participants through Prolific in July 2025. Participants were recruited from among residents in the US and Germany, two major regions for vehicle manufacturing. To maintain consistency across countries, we used Prolific's general sampling tool rather than the representative sample based on U.S. Census data, as a representative sampling option was not available for Germany.

We excluded 18 participants who answered 'no' to the honesty check, completed the survey in under 250 seconds, and/or provided incoherent responses. After these exclusion criteria, we obtained a final sample of 282 valid responses. Participants in the United States were compensated $2.37, while those in Germany received approximately €2.77. The median completion time was 740 seconds (approximately 12 minutes), corresponding to an hourly rate of $11.53 in the U.S. and €13.47 in Germany

Table 1: Participant demographics (N = 282)

		N	%
Age	18–29	84	29.8%
	30–39	91	32.3%
	40–49	54	19.1%
	50–59	36	12.8%
	60–69	15	5.3%
	70+	2	0.7%
Gender	Male	175	62.1%
	Female	104	36.9%
	Other / Prefer not to say	3	1.0%
Education	High school	63	22.3%
	College	19	6.7%
	Trade / Technical	19	6.7%
	Undergraduate	88	31.2%
	Post-graduate	87	30.9%
	Other / Prefer not to say	6	2.1%
IT knowledge	Yes	76	27.0%
	No	206	73.0%

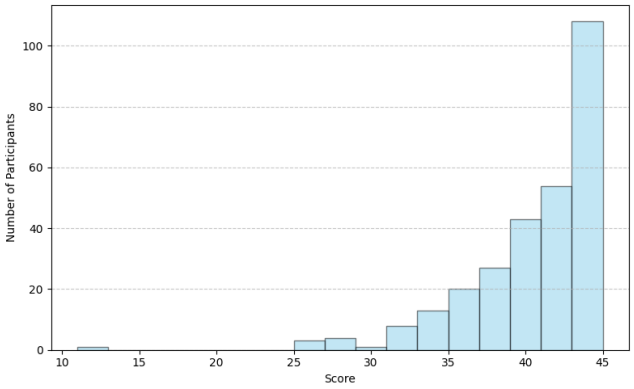

Fig. 2: Distribution of our participants privacy beliefs (IUIPC).

—both exceeding the minimum wage in their respective countries.

Table 1 presents the demographic characteristics of the final sample. Participants ranged in age from 18 to 72 years (mean = 38.0, SD = 11.9). A majority (62.1%) identified as male, while 1.0% identified as non-binary/third gender or selected "prefer not to say." Regarding professional background, 27.0% reported working in IT or a related field, while 73.0% reported they were not.

Figure 2 shows the distribution of IUIPC privacy concern scores. The mean IUIPC score among participants was 40.4 (SD = 4.7). These scores imply that the majority of our participants highly value privacy.

4. Results

4.1 Vehicle Use

Table 2 presents participants' reported experiences and interactions with vehicles. The majority indicated that they drive a vehicle regularly (41.5%), occasionally (22.3%), or often (16.7%). A smaller portion reported rarely (6.4%) or never (3.5%) driving, while 8.9% stated that they had driven in the past but no longer do so.

Most participants reported riding in vehicles as passengers either occasionally (41.8%), often (22.0%), or regularly (15.2%). In contrast, 19.9% indicated they rarely ride

Table 2: How participants use vehicles.

		N	%
Frequency of driving	I drive often	47	16.7%
	I drive regularly	117	41.5%
	I drive occasionally	63	22.3%
	I have rarely driven	18	6.4%
	I have never driven	10	3.5%
	I used to drive but do not currently	25	8.9%
	Other / Prefer not to say	2	0.7%
Vehicle passenger use	I ride regularly	43	15.2%
	I ride often	62	22.0%
	I ride occasionally	118	41.8%
	I rarely ride	56	19.9%
	I never ride	3	1.1%
Vehicle ownership	I own the vehicle myself	185	65.6%
	A family member owns it	66	23.4%
	A friend owns it	6	1.0%
	I use a car sharing or rental service	12	4.3%
	Other / Prefer not to say	13	4.6%
Purchased a vehicle	Connected	115	40.8%
	Non-connected	99	35.1%
	Never	64	22.7%
	Prefer not to say	4	1.4%
Familiar with connected vehicles	Yes*	208	73.8%
	No	74	26.2%

*For simplicity, we show the percentage of participants who selected "Very familiar" or "Somewhat familiar" on a 5-point Likert scale in this table.

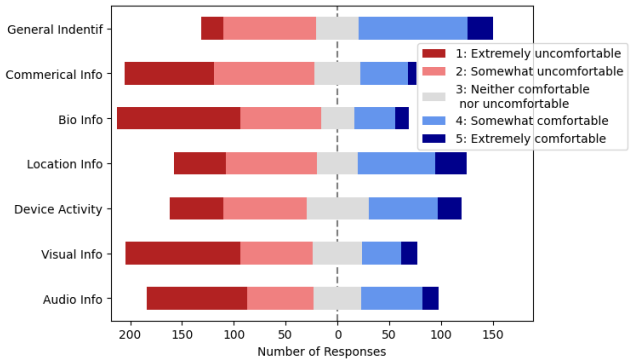

Fig. 3: The distribution of general comfort levels reported.

as passengers, and 1.1% reported never doing so. These results suggest that participants had substantial exposure to vehicles in diverse roles, whether as drivers, passengers, or pedestrians.

A large majority reported either owning a vehicle themselves (65.6%) or using a vehicle owned by a family member (23.4%). Regarding vehicle purchases, 40.8% had purchased a CV, and 35.1% had purchased a non-CV. Only 22.7% reported never having purchased a vehicle. Indicating that most participants have had direct purchasing experience, likely providing them with some understanding of CV features and data practices.

Finally, 73.8% of participants reported being familiar with connected vehicles. This suggests a high level of familiarity with the features and functions associated with modern CV systems.

4.2 General Data Collection Perceptions

We first examined participants' comfort levels of general data collection. Figure 3 reports the mean comfort scores for each data type. Participants expressed the least comfort with sharing biometric information (mean = 2.11), fol-

lowed by visual information (mean = 2.21) and commercial information (mean = 2.27). In contrast, participants were most comfortable sharing general identifiers and location information. These results suggest a generally low level of comfort with most forms of data collection, particularly for sensitive categories such as biometric information.

In addition to Likert-scale responses, we asked participants to identify which type of data they were least willing to share in the context of general data collection. These responses are shown in the first row of Figure 3. The most frequently selected categories were biometric and commercial information. Other categories, including general identifiers, location, device activity/network, visual, and audio information, were selected less frequently and were relatively evenly distributed across the remaining participants. These findings highlight that while discomfort spans multiple data categories, biometric and financial information data stand out as particular areas of concern for users.

4.3 Vehicle Data Collection

Figure 4 shows how likely users believe that vehicle companies and CVs are collecting data of each type. The majority of participants reported being familiar with CV features (73.8%), the overall likelihood ratings for the types of data CVs are capable of collecting were relatively low. Among the data types, participants perceived vehicle information and location information as the most likely to be collected by both vehicle companies and CVs. In contrast, audio, visual, and commercial information were considered some of the least likely to be collected. Notably, participants distinguished between the two entities: commercial information and general identifiers were perceived as more likely to be collected by vehicle companies, whereas these same data types were considered less likely to be collected by CVs. Interestingly, only the likelihood of commercial information was found to significantly correlate with cultural background (χ^2 test, $\chi^2 = 330.96$, $p = 0.005$).

RQ1: These findings suggest that users do not accurately perceive the full range of data collected within the vehicle ecosystem. Even though most participants reported familiarity with CV features, their likelihood estimates for many sensitive data types, such as biometric, audio, or visual information, were notably low. This misalignment suggests that public understanding of vehicle data practices remains incomplete [15], [17]. Additionally, we see that cultural contexts do not influence participants expectations for most data types, expect for commercial information. Suggesting differing levels of sensitivity to corporate data practices, varying norms of trust in commercial institutions, and diverse understandings of the value and ownership of personal information.

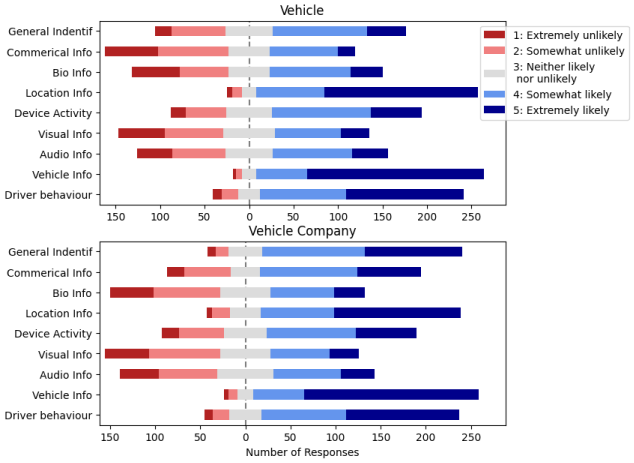

Fig. 4: How likely participants think vehicles and vehicle companies are at collected each type of data.

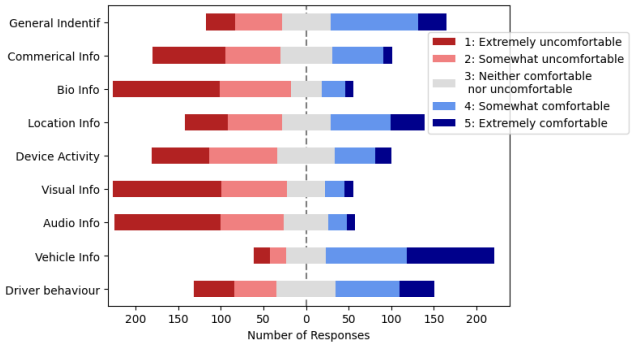

Fig. 5: How comfortable participants are with vehicle companies collecting each type of data.

Next, we examine the comfort levels of participants with various data types collected by vehicle companies, as illustrated in Figure 5. The findings reveal a notable inverse relationship: data perceived as less likely to be collected, such as audio, visual, commercial, and particularly biometric information, elicited the greatest discomfort. In contrast, participants reported higher comfort levels for data types believed to be more frequently collected, including general identifiers, vehicle information, and driver behavioral data. Furthermore, chi-square tests reveal a strong correlation between participants likelihood rating and comfort for each type of data.

This pattern aligns with the mere-exposure effect, which posits that familiarity increases comfort and likability by reducing uncertainty and cognitive load [22]. It also echoes findings from IoT research that indicate that users are generally more comfortable with data practices they perceive as common, while uncommon or intrusive data types provoke greater concern [7]. These insights suggest the importance of enhancing transparency around less familiar data collection practices.

Furthermore, the second row of Table 3 demonstrates that the types of data participants are least willing to share with vehicle companies closely align with both their general sharing preferences and their reported comfort levels.

Table 3: The types of data participants are least willing to share in general and with vehicle specific companies.

Comfort of Data Collection	General Identif	Cmmrical Info	Bio Info	Location Info	Device Activity	Visual Info	Audio Info	Driver Behav	Vehicle Info	Other
Least willing to share (General)	24 8.5%	87 30.9%	101 35.8%	14 5.0%	15 5.3%	28 9.9%	8 2.8%			
Least willing to share (Vehicle Specific)	20 7.1%	79 28.0%	101 35.8%	8 2.8%	10 3.5%	34 12.1%	12 4.3%	13 4.6%	2 0.7%	3 1.1%

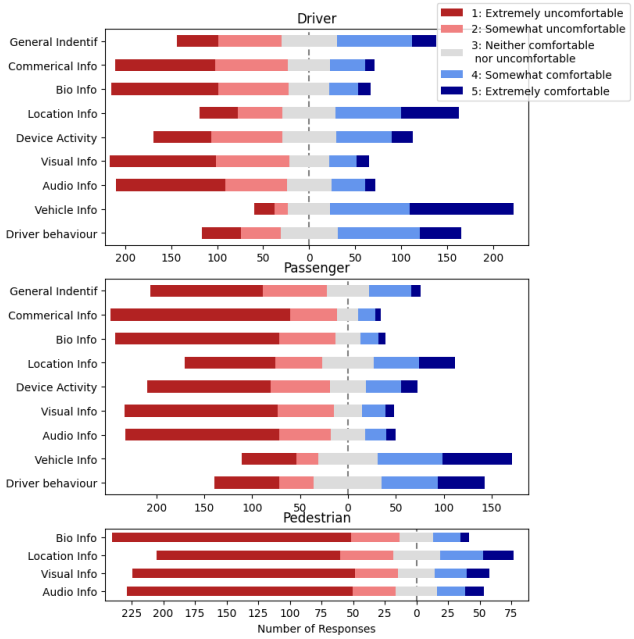

Fig. 6: How comfortable participants are with vehicles collecting each type of data in different roles.

In particular, biometric information and commercial information emerge as particularly sensitive categories, and participants express a strong reluctance to disclose them. This consistency in willingness and comfort measures displays the increased importance of these data types in shaping user trust and perceived risk.

Interestingly, we find that the type of data participants are least willing to share in both contexts (General and Vehicle) has no significant corelation to cultural backgrounds.

Next, we examine the comfort levels of participants with data collection in different roles within the vehicle ecosystem, as shown in Figure 6. Not all data types are applicable to each role, for example, pedestrians are realistically subject only to external sensor or camera data collection. Accordingly, we restricted pedestrian-related questions to audio, visual, biometric, and location information.

As drivers, participants reported being comfortable only with the collection of vehicle-related and driver behavioral information, while expressing moderate comfort with location information and general identifiers. Consistent with the broader comfort trends observed in Figure 5, drivers were most uncomfortable with the collection of audio, visual, commercial, and biometric data.

When considering the role of passengers, responses closely mirrored those of drivers. However, discomfort was even more pronounced, particularly regarding audio, visual, commercial, and biometric information.

This pattern intensified when participants assumed the perspective of pedestrians. Overall, pedestrians expressed discomfort with all types of data that might reasonably be collected about them. Location information elicited comfort levels most comparable to those reported for drivers and passengers, whereas biometric, audio, and visual data triggered even higher levels of unease in this context.

Table 4: How long participants would take to opt out of data collection by their vehicles and if they would still opt out if it meant reduced functionality.

		N	%
Likely to opt-out	Yes*	225	79.8%
	No	57	20.3%
How long(Minutes)	Less than 5	25	11.1%
	5-10	41	18.2%
	11-15	45	20.0%
	16-20	28	12.4%
	21-25	6	2.6%
	26-30	14	6.2%
	More than 30	66	29.3%
Limited features	Yes*	141	62.7%
	No	84	37.3%

*For simplicity, we show the percentage of participants who selected “Extremely likely” or “Somewhat likely” on a 5-point Likert scale in this table.

RQ2: These results demonstrate that privacy perceptions in the CV ecosystem are strongly role-dependent. Data types already viewed as sensitive become increasingly uncomfortable for participants as their role moves further from the vehicle, shifting from driver to passenger to pedestrian.

4.4 Opt-out intentions

We asked participants whether they would opt out of vehicle data collection, how much time they would be willing to spend to complete the opt-out process, and whether they would still choose to opt out if doing so limited certain vehicle features. Results are presented in Table 4.

The majority of participants (79.8%) indicated that they would choose to opt out of data collection. Among these, 49.3% reported a willingness to spend less than 15 minutes completing the process. In contrast, and consistent with previous research [17], 50.5% were willing to spend more than 15 minutes opting out, and 29. 3% of the participants were even willing to spend more than 30 minutes. Of those willing to opt out (79.8%), 62.7% reported that they would still choose to do so even if it meant losing access to certain vehicle features.

We also asked participants whether they would consider changing their vehicle brand due to the types of information it collects. Among respondents, 58.2% indicated that they would consider switching brands. Table 5 details the specific types of data that participants reported would influence such a decision. Consistent with previous findings, biometric, commercial, visual, and audio information emerged as the most influential factors driving participants' consideration of brand change. In contrast, vehicle information, general identifiers, and driver behavioral data were perceived as less critical in shaping participants' brand preferences. Unsurprisingly, those participants who frequently drove had a high correlation (χ^2 test, $\chi^2 = 52.17$, $p = 0.004$) to changing their vehicle based on its data collection.

RQ3: Overall, our findings suggest that users place high value on privacy and are willing to invest both time and potential convenience to retain control over their personal data. The collection of biometric, commercial, visual, and audio information is the most influential in participants considering changing their vehicle brand.

5. Discussion

Our study provides several key insights into user perceptions of data collection within the connected vehicle ecosystem. First, consistent with prior IoT and vehicle studies [15], [17], [23], certain types of data, particularly biometric, audio, visual, and commercial information, are perceived as highly sensitive. Participants expressed heightened discomfort and reluctance to share these types of data, even when other, more commonly collected information such as vehicle data, driver behavior, or general identifiers was considered acceptable. This pattern was also reflected in participants' willingness to opt out or change vehicle brands, suggesting that intrusive, personally identifiable data are the strongest drivers of user privacy concerns.

5.1 Implications

Our results further reveal that familiarity and perceived ubiquity of data collection influence comfort: **participants were generally more accepting of data types they believed were routinely collected.** This suggests that one way to mitigate privacy concerns is to increase transparency and user awareness, particularly for sensitive data categories. Practical strategies could include clear and concise in-vehicle notifications, interactive dashboards that show what data is collected and why, and contextual explanations highlighting the utility of data for safety or performance. These recommendations align with prior work emphasizing the role of transparency and user control in IoT systems [23].

A second major contribution of our study is the demonstration that **perceptions of data collection are highly role-dependent**. Drivers reported higher comfort levels across most data types, whereas passengers exhibited more discomfort, and pedestrians were the most sensitive to data collection. This distinction highlights the need to design privacy and transparency mechanisms that account for the full ecosystem of vehicle users and non-users. For example, communication strategies should not rely solely on the vehicle owner but should also inform passengers and pedestrians when sensors or cameras may capture their information, in line with recommendations from Windl et al. [23].

Taken together, our findings highlight the dual importance of addressing both the sensitivity of specific data types and the diversity of roles within the vehicle ecosystem. By combining transparency, role-aware communication, and user-centric opt-out mechanisms, manufacturers can better align data practices with user expectations and promote trust in increasingly sensor-rich CVs.

5.2 Limitations

These findings should be interpreted in light of several limitations. First, we rely on self-reported data, which may be subject to biases such as social desirability, recall errors, or misinterpretation of questions. Second, our sample, though drawn from the United States and Germany, may not fully represent all demographics within these countries, and certain subpopulations, such as older adults or individuals with limited vehicle experience, may be underrepresented. Third, the survey focuses on perceptions of data collection rather than directly measuring behavioral responses, such as actual opt-out behavior or engagement with privacy controls. Fourth, while our survey includes multiple roles (drivers, passengers, pedestrians), participants' responses may still be influenced by their primary experiences, which could bias role-based comparisons. Finally, as this study is cross-sectional, it captures attitudes at a single point in time and may not reflect how perceptions evolve as connected vehicle technologies, regulations, and public awareness change.

5.3 Future Work

While the present study offers valuable insights into how drivers, passengers, and pedestrians perceive data collection in connected vehicles, several avenues remain for exploration. One key question is: where do gaps exist between public expectations and current vehicle data practices? To address this, future work should involve structured discussions with vehicle manufacturers to better understand how data is actually collected, processed, and shared. Comparing public perceptions with manufacturer-reported practices would help identify specific areas where transparency or privacy protections could be improved.

In addition, expanding the survey to include participants from countries with different privacy regulations and cul-

Table 5: The type of information being collected that participants (58.2%, N=164) would consider changing their vehicle.

Type of Information	General Identif	Cmmrical Info	Bio Info	Location Info	Device Activity	Visual Info	Audio Info	Driver Behav	Vehicle Info	Other
Information to consider change	46 7.2%	103 16.2%	116 18.2%	52 8.2%	61 9.6%	102 16.0%	95 15.0%	48 7.5%	12 1.9%	2 0.3%

*Participants could select more than one type of data. Hence N is greater than 164.

tural norms, such as China, would provide valuable cross-cultural perspectives. Such an extension would allow exploration of how regulatory frameworks, societal attitudes toward privacy, and cultural expectations shape comfort levels, perceived risks, and willingness to opt out of vehicle data collection. Together, these efforts will build a more comprehensive understanding of the vehicle data ecosystem, informing both policy recommendations and the design of privacy-preserving technologies for CVs.

6. Conclusion

This study examined user perceptions of data collection in the connected vehicle ecosystem, focusing on how different types of data and user roles influence comfort and privacy concerns. Our findings indicate that certain data types, particularly biometric, audio, visual, and commercial information, are viewed as highly sensitive, while more commonly collected data, such as vehicle information, general identifiers, and driver behavioral data, are perceived as less intrusive. Comfort with data collection is influenced by both familiarity and perceived ubiquity, highlighting the need for greater transparency and user awareness around less familiar or more sensitive data types.

Importantly, our study demonstrates that privacy perceptions vary across roles within the vehicle ecosystem: drivers generally report higher comfort levels than passengers, and pedestrians are the most sensitive to data collection. These insights highlight need for privacy mechanisms and communication strategies that address all CV stakeholders.

Acknowledgments. This work was partially supported by JST CREST JPMJCR23M4 and JST SPRING JPMJSP2128.

References

[1] Abdelkader, G., Elgazzar, K. and Khamis, A.: Connected Vehicles: Technology Review, State of the Art, Challenges and Opportunities, *Sensors*, Vol. 21, No. 22 (2021).

[2] Znidi, F., Morsy, M. and Rathore, H.: Future on Wheels: Safeguarding Privacy in Tomorrow ' s Connected Vehicles-FUTURE-SP, *IEEE Access*, Vol. 12, pp. 179857–179878 (2024).

[3] Islami, L., Fischer-Hübner, S. and Papadimitratos, P.: Capturing drivers' privacy preferences for intelligent transportation systems: An intercultural perspective, *Computers Security*, Vol. 123, p. 102913 (2022).

[4] Matin, A. and Dia, H.: Public perception of connected and automated vehicles: Benefits, concerns, and barriers from an Australian perspective, *Journal of Intelligent and Connected Vehicles*, Vol. 7, No. 2, pp. 108–128 (2024).

[5] Schoettle, B. and Sivak, M.: A survey of public opinion about connected vehicles in the U.S., the U.K., and Australia, *2014 International Conference on Connected Vehicles and Expo (ICCVE)*, pp. 687–692 (2014).

[6] Al-Ameen, M. N., Chauhan, A., Ahsan, M. M. and Kocabas, H.: A look into user ' s privacy perceptions and data practices of IoT devices, *Information and Computer Security*, Vol. 29, No. 4, pp. 573–588 (2021).

[7] Williams, M., Nurse, J. R. C. and Creese, S.: Privacy is the Boring Bit: User Perceptions and Behaviour in the Internet-of-Things, *2017 15th Annual Conference on Privacy, Security and Trust (PST)*, pp. 181–18109 (2017).

[8] Singh, A. D., Wang, B., Garcia, L., Chen, X. and Srivastava, M.: Understanding factors behind IoT privacy – A user's perspective on RF sensors (2024).

[9] Ponciano, L., Barbosa, P., Brasileiro, F., Brito, A. and Andrade, N.: Designing for Pragmatists and Fundamentalists: Privacy Concerns and Attitudes on the Internet of Things, *Proceedings of the XVI Brazilian Symposium on Human Factors in Computing Systems* (2017).

[10] Zheng, S., Apthorpe, N., Chetty, M. and Feamster, N.: User Perceptions of Smart Home IoT Privacy, *Proc. ACM Hum.-Comput. Interact.*, Vol. 2 (2018).

[11] Moore, L., Yamagishi, R., Sawada, K. and Mori, T.: What Are Cars Collecting? A Study of Privacy Policies in the Automotive Industry, 3rd USENIX Symposium on Vehicle Security and Privacy (VehicleSec ' 25) (2025).

[12] Caltrider, J., Rykov, M. and MacDonald, Z.: It ' s Official: Cars Are the Worst Product Category We Have Ever Reviewed for Privacy (2023). [Online; posted 6-September-2023].

[13] Ziegler, B.: How to Keep Your Car From Spying on You (2024). [Online; posted 5-June-2024].

[14] Sorrel, C.: Why Your Car Is a 'Privacy Nightmare,' and What You Can Do About It (2023). [Online; posted 8-September-2023].

[15] Acharya, S. and Mekker, M.: Public Perception of the Collection and Use of Connected Vehicle Data, Final Report MPC-621, MPC 21-439, U.S. Department of Transportation (2021).

[16] Schmidt, T., Philipsen, R., Themann, P. and Ziefle, M.: Public perception of V2X-technology - evaluation of general advantages, disadvantages and reasons for data sharing with connected vehicles, *2016 IEEE Intelligent Vehicles Symposium (IV)*, pp. 1344–1349 (2016).

[17] Bloom, C., Tan, J., Ramjohn, J. and Bauer, L.: Self-Driving Cars and Data Collection: Privacy Perceptions of Networked Autonomous Vehicles, *Proceedings of the Thirteenth Symposium on Usable Privacy and Security*, pp. 357–375 (2017).

[18] Malhotra, N. K., Kim, S. S. and Agarwal, J.: Internet users ' information privacy concerns (IUIPC): The construct, the scale, and a causal model., *Information systems research*, pp. 336—-355 (2004).

[19] Markert, P., Bailey, D. V., Golla, M., Durmuth, M. and Aviv, A. J.: On the security of smartphone unlock pins., *ACM Transactions on Privacy and Security (TOPS)*, pp. 1–36 (2021).

[20] Bailey, D. V., Munyendo, C. W., Dyer, H. A., Grant, M., Markert, P. and Aviv, A. J.: "Someone Definitely Used 0000": Strategies, Performance, and User Perception of Novice Smartphone-Unlock PIN-Guessers, *Proceedings of the 2023 European Symposium on Usable Security*, p. 158–174 (2023).

[21] Moore, L., Mori, T. and Hasegawa, A. A.: Negative Effects of Social Triggers on User Security and Privacy Behaviors, *Twentieth Symposium on Usable Privacy and Security (SOUPS 2024)*, pp. 605–622 (2024).

[22] Zajonc, R.: Mere Exposure: A Gateway to the Subliminal, *Current Directions in Psychological Science*, Vol. 10, No. 6, pp. 224–228 (2001).

[23] Windl, M., Akgul, O., Malkin, N. and Cranor, L. F.: Privacy Solution or Menace? Investigating Perceptions of Radio Frequency Sensing, *USENIX Security 2025* (2025).