applied sciences

MDPI

Article

The Interface of Privacy and Data Security in Automated City Shuttles: The GDPR Analysis

Meriem Benyahya [1,*], Sotiria Kechagia [2], Anastasija Collen [1] and Niels Alexander Nijdam [1]

[1] Centre Universitaire d'Informatique, Geneva School of Economics and Management, University of Geneva, Route de Drize 7, CH-1227 Carouge, Switzerland; anastasija.collen@unige.ch (A.C.); niels.nijdam@unige.ch (N.A.N.)
[2] Center for Digital Trust, École Polytechnique Fédérale de Lausanne, 1015 Lausanne, Switzerland; ria.kechagia@epfl.ch
* Correspondence: meriem.benyahya@unige.ch

Abstract: The fast evolution and prevalence of driverless technologies has facilitated the testing and deployment of automated city shuttles (ACSs) as a means of public transportation in smart cities. For their efficient functioning, ACSs require a real-time data compilation and exchange of information with their internal components and external environment. However, that nexus of data exchange comes with privacy concerns and data protection challenges. In particular, the technical realization of stringent data protection laws on data collection and processing are key issues to be tackled within the ACSs ecosystem. Our work provides an in-depth analysis of the GDPR requirements that should be considered by the ACSs' stakeholders during the collection, storage, use, and transmission of data to and from the vehicles. First, an analysis is performed on the data processing principles, the rights of data subjects, and the subsequent obligations for the data controllers where we highlight the mixed roles that can be assigned to the ACSs stakeholders. Secondly, the compatibility of privacy laws with security technologies focusing on the gap between the legal definitions and the technological implementation of privacy-preserving techniques are discussed. In face of the GDPR pitfalls, our work recommends a further strengthening of the data protection law. The interdisciplinary approach will ensure that the overlapping stakeholder roles and the blurring implementation of data privacy-preserving techniques within the ACSs landscape are efficiently addressed.

Keywords: automated city shuttles; connected automated vehicles; shared mobility; data privacy; privacy-preserving; GDPR

check for updates

Citation: Benyahya, M.; Kechagia, S.; Collen, A.; Nijdam, N.A. The Interface of Privacy and Data Security in Automated City Shuttles: The GDPR Analysis. *Appl. Sci.* **2022**, *12*, 4413. https://doi.org/10.3390/app12094413

Academic Editor: Juan-Carlos Cano

Received: 24 March 2022
Accepted: 20 April 2022
Published: 27 April 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

The emergence of novel technologies based on artificial intelligence (AI) and Internet of Things (IoT) in the transport sector presents substantial regulatory challenges. Since the early stages of the Internet of Vehicles (IoV) deployment through connected automated vehicles (CAVs), multifaceted problems of compliance to the European personal data protection regulation have been raised [1]. To a greater extent, with the envisioned deployment of CAVs into the public transportation through the automated city shuttles (ACSs), there are novel and daunting regulatory hurdles to overcome [2].

The ACS, such as the one deployed in a pilot-site within the Avenue project [3] and depicted in Figure 1, compiles multiple sensors and AI units' inputs to achieve a high automation driving, as per the Society of Automotive Engineering (SAE) levels 4 and 5 [4]. A highly connected ACS would use intrinsic communications and endless exchange of information through vehicle-to-vehicle (V2V), vehicle-to-infrastructure (V2I), vehicle-to-cloud (V2C), and vehicle-to-pedestrian (V2P) to broadcast traffic conditions and share the predictions within the vehicular network. Such a complex system makes the ACS able to navigate autonomously while using real-time and fine-grained data.



Figure 1. Example of an automated city shuttle (Geneva, Switzerland).

With the rise of registered information by the in-vehicle components and the vehicular external communications, data protection needs to become more significant and applicable to the different types of data generated by the ACS. The ACS's sophisticated cameras can record both eventual obstacles disrupting the autonomous driving mode and any visual personal data, including facial identities. In addition, within the vehicular network, the nodes, which can be an ACS or a road side unit (RSU), exchange beacon messages combining identity, location, and temporal properties. Such messages embed timestamps, vehicular information, authorization certificates, and location data [5,6], which is qualified as personal data. In other words, those messages' content is linked to identified or identifiable natural persons (hereinafter, referred as "data subjects") under Article 4 of the General Data Protection Regulation (GDPR) [7]. Sometimes personal location data, such as travel itineraries, can reveal sensitive information about data subjects' health condition, sexual orientation, racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership [1]. In these cases, stricter privacy protection rules are required to be applied.

An idiosyncratic element to consider when exploring the ACS ecosystem is the provision of customized services to the passengers through the integration of location-based services (LBS) such as the mobility as a service (MaaS). MaaS is a mobility platform which bridges public transport to mobility services by providing, for example, door-to-door services based on the passengers information, including their location [8]. Hence, the more LBS services are deployed, the more personal data are transferred through external digital platforms, leading to higher risk of privacy violation. Based on such assumptions, the Article 29 Working Party (WP29), the predecessor of the European Data Protection Board (EDPB) "whose purpose is to ensure consistent application of the GDPR", assessed the necessity of transparency and proportionality controls due to high risk of personal data leakage and unlawful processing of mobility data [9].

The mere identification of the collection and further processing of personal data is of vital importance to ensure protection from privacy risks, which is attempted to be provided by regulatory frameworks such as the GDPR in the European Union (EU). Any failure to comply with the GDPR requirements could potentially result in physical, material, or non-material damage to natural persons, such as loss of control over their personal data or limitation of their rights (Recital 85) [7]. Additionally, the ACS data privacy risks entail further regulations such as the ePrivacy [10] and the Network and Information Security (NIS) directives [11,12]. Those directives urge to take the appropriate and proportionate measures for a high level of information systems' security to prevent and minimize the impact of cyber incidents and attacks. For the purpose of our work, we constrain our analysis to explore only the GDPR implications.

The present article provides the following contributions:

1. An extensive analysis on how the GDPR discusses the principles of data processing, the rights of data subjects, and roles and responsibilities of the stakeholders (data controllers, data processors, sub-processors, etc.) before, during, and after the processing of personal data collected from the ACS.

2. Categorization of the main privacy-preserving techniques that are applicable to the ACS environment.
3. Presentation of the gaps between the legal definitions and technological implementation of privacy-preserving schemes recommended by the GDPR, which are mainly pseudonymization and anonymization techniques.
4. Investigation, through interdisciplinary efforts, into the shortcomings and pitfalls of the GDPR data processing principles in protecting personal data within the complex ACS context.

This article addresses the following research questions:

RQ 1. *According to the GDPR, what are the rights of data subjects that data controllers have to take into consideration before initiating any processing of collected data from the ACS ecosystem?*

RQ 2. *What is the attribution of role for each stakeholder involved in the ACS landscape? What are their respective responsibilities based on the GDPR?*

RQ 3. *What are the relevant privacy-preserving techniques to protect personal data within the ACS environment and how do each guarantee personal information protection in line with the GDPR? Would the techniques recommended by the GDPR be enough for an optimal protection?*

The remainder of this paper is structured as follows: Section 2 discusses the related work and makes a comparison of the present work with already published efforts. Section 3 delivers a thorough review of the most crucial GDPR regulatory effects while applied to the automated driving ecosystem. Section 4 overviews the gaps between technological and regulatory disciplines in terms of implementation of mitigation techniques to data privacy risks related to the ACS deployment. Section 5 offers concluding remarks and future work orientation about lawful processing of personal data within the ACS environment.

2. Related Work

With the pervasive technologies leading to driverless vehicles and their associated privacy challenges, the GDPR has been serving as the European prominent legal reference. As well as beyond the EU, the GDPR has paved the way for a global impact regarding data protection [13]. To illustrate, the Canadian Personal Information Protection and Electronic Documents Act (PIPEDA) was influenced by the GDPR to strengthen consumer rights through aligned and similar obligations [14]. Additionally, the Australian national transport authority acknowledged the GDPR while regulating the Intelligence Transport System (ITS) and automated vehicles' data processing [15]. To that end, the GDPR has been perceived as the international legislation for data governance with strict obligations to data controllers, data subjects, and engineers for consent and privacy-enhancing technology implementations [16]. Therefore, our work demonstrates efforts on deep investigation into the GDPR and its relation to the automated vehicles. To this end, we present three main avenues, namely, GDPR requirements, recognition of privacy challenges, and privacy-preserving methods.

2.1. GDPR in Driverless Landscape

Several works exist in the domain of data protection, aiming to provide analysis on the GDPR's proposed measures under the scope of the generic vehicular environment. For instance, Taeihagh and Lim [17,18] provided a brief overview of the GDPR with a focus on consent and penalties conditions. Similarly, Pattinson et al. [19] discussed, as a GDPR requirement, the role of the data subjects' consent while operating on level 3 and 4 CAVs. Additionally, Vallet [20] presented the GDPR's applicability within a CAV environment but with a focus on drivers and vehicles owners' rights. Moreover, Krontiris et al. [21] analyzed the data protection challenges within the CAV environment by categorizing the types of collected data, data subjects, and data controllers profiles. The authors also highlighted

the implications of the GDPR on AI by matching relevant articles to technologies used for data processing. Further multidisciplinary works presented interesting overviews on how the GDPR can be considered within the development life cycle to avoid data privacy breaches [6,13]. The genuine interest for such publications is the protection of the data subject rights. Yet, an explicit and holistic review of all the CAV's stakeholders rights and obligations is still lacking.

Shifting from legal requirements on consent, data subjects, and data controller, Bastos et al. [22] synthesized the principles of data processing and the Data Protection Impact Assessment (DPIA) concerns. However, the paper scope remains broad to all IoT devices. Despite the fact that ACS can be considered as a complex IoT device too, it has specific parameters to be taken into account while performing such analysis. Conversely, Ainsalu et al. [2] provided a more specific review to ACS discussing the legal framework of testing and deploying such vehicles in Europe, though the authors referred broadly to the GDPR as the integral law that controls the required processing within the ACS without an in-depth analysis on the GDPR implications or limitations.

2.2. Data Privacy Challenges within Vehicular Environment

Extensive efforts in identifying data privacy challenges within the CAV's ecosystem were provided in multiple research works. Collingwood [23] and Glancy [24] warned about the privacy implications of using automated vehicles as a means of public transportation. They identified three fields of concerns, which are "autonomy privacy interests, information privacy interests, and surveillance privacy interests", where the individuals acquire the total liberty and control to make independent choices about themselves, their lives, and their data. They argue that by collecting and processing data from CAVs, the passengers will lose control over their private information as the data controllers may infer the individuals' past, present, and future locations and behavior. However, these papers do not refer to or analyze the compliance to any specific law or regulation.

2.3. Data Privacy-Preserving Methods

Motivated by the potential of privacy-preserving schemes, as recommended by the GDPR, multiple literature reviews outlined clarification about the implementation of pseudonymization and anonymization. Karnouskos and Kerschbaum [25] studied the feasibility of insuring the integrity of automated vehicles while preserving the individuals' privacy. Two particular concepts are brought forth by the GDPR, namely, privacy by design (PbD) and privacy by default (PbDf). Although these concepts are not necessarily new [26], as part of the GDPR their impact can be significant. PbD starts with the implementation of security measures from the outset of data processing and extends to the implementation of technical and organizational measures during the whole lifecycle of the data involved, whereas PbDf calls for personal data, which is necessary and proportionate for each specific purpose of the processing to be accomplished. This relates to the amount of personal data collected, the extent of the processing, the retention period, and who has access to it. The authors focused on the deployment of PbD principles and technical mitigation techniques based on encryption and anonymization in the scope of CAVs. Similarly, Mulder and Vellinga [27] highlighted the difference between pseudonymization and anonymization as the key privacy-preserving techniques required by the GDPR and applied to vehicular environment. In a broader scope, Ribeiro and Nakamura [28] compared pseudonymization and anonymization techniques and how they can protect personal data within IoT systems. Unfortunately, most of these works neither provided a comprehensive categorization of the applied pseudonymization and anonymization techniques to the ACS environment, nor flagged the re-identification risks related to such techniques.

The most detailed analysis attempting to raise re-identification risks were proposed in [29–31]. Brasher [29] discussed the limitation of anonymization and encouraged its conjunction with pseudonymization to reduce the re-identification risks. Li et al. discussed the inference and de-anonymization risks within the driverless environment [30]. Löbner

et al. [31] evaluated the re-identification risks and impact within the vehicular context through a real test bed scenario, though the efforts remain limited to some of the privacy-preserving techniques without reviewing them thoroughly.

Not limited to researchers, EU institutions initiatives assessed the re-identification risks in multiple publications. The opinion 05/2014 of the WP29 [9] represents one of the first guides approaching the effectiveness of anonymization techniques. European Union Agency for Cybersecurity (ENISA) provided deeper analysis on pseudonymization and anonymization implementation, differences and limitations [32,33]. Nevertheless, such publications' scope remain very wide, yet applicable to driverless environment.

Following this presentation on the related research and the analysis from Table 1, we differentiate from the other efforts by:

- Presenting an interdisciplinary approach regarding data protection requirements in the ACS ecosystem by assessing and addressing simultaneously both regulatory and technical challenges.
- Providing an in-depth analysis of the GDPR provisions and limitations that are relevant to the ACS.
- Having a closer examination on how the legal requirements are compatible with the technologies deployed in the ACS.
- Analyzing the inconsistencies between the legal definitions on pseudonymization and anonymization in the GDPR and their technical implementation through a comprehensive categorization of the most relevant applicable techniques into the ACS landscape.
- Developing a significant reference point for academic research on ACS, public transportation operators, automobile manufacturers (OEMs), policymakers, and service providers, acquiring or looking forward to deploying ACSs within their systems.

Table 1. Related work comparison.

Related Work	Year	Scope				PC [a]	GDPR Implications					PP [d]			GDPR Pitfalls
		ACS	CAV	IoT	IT		Principles	DS Rights [b]	DC Obligations [c]	Roles	DPIA	Pseudonymization	Anonymization	Risks	
ENISA [33]	2022	✗	✗	✗	✓	✗	✗	✗	✗	✓	✓	✓	✓	✓	✓
Mulder and Vellinga [27]	2021	✗	✓	✗	✗	✓	✓	✓	✓	✓	✓	✓	✓	✗	✗
Löbner et al. [31]	2021	✗	✓	✗	✗	✓	✗	✗	✗	✗	✗	✗	✓	✓	✗
ENISA [32]	2021	✗	✓	✗	✗	✓	✗	✗	✗	✗	✗	✓	✗	✗	✗
Pattinson et al. [19]	2020	✗	✓	✗	✗	✓	✓	✓	✓	✗	✗	✗	✗	✗	✗
Krontiris et al. [21]	2020	✗	✓	✗	✗	✓	✓	✓	✓	✗	✓	✗	✗	✗	✗
Costantini et al. [13]	2020	✗	✓	✗	✗	✓	✗	✓	✗	✓	✗	✗	✗	✗	✗
Vallet [20]	2019	✗	✓	✗	✗	✓	✓	✓	✗	✗	✗	✓	✗	✗	✗
Ribeiro and Nakamura [28]	2019	✗	✗	✓	✗	✓	✗	✗	✗	✗	✗	✓	✓	✗	✗
Li et al. [30]	2019	✗	✓	✗	✗	✓	✗	✗	✗	✗	✗	✗	✓	✓	✗
Taeihagh and Lim [17,34]	2018	✗	✓	✗	✗	✓	✓	✓	✗	✗	✗	✗	✗	✗	✗
Veitas and Delaere [6]	2018	✗	✓	✗	✗	✓	✓	✗	✗	✓	✗	✗	✗	✗	✗
Bastos et al. [22]	2018	✗	✗	✓	✗	✗	✓	✓	✓	✗	✓	✓	✓	✗	✗
Ainsalu et al. [2]	2018	✓	✗	✗	✗	✓	✗	✗	✗	✗	✗	✗	✗	✗	✗
Karnouskos and Kerschbaum [25]	2018	✗	✓	✗	✗	✓	✗	✗	✗	✗	✗	✓	✓	✗	✗
Brasher [29]	2018	✗	✗	✗	✓	✓	✗	✗	✗	✗	✗	✓	✓	✓	✓
Collingwood [23]	2017	✗	✓	✗	✗	✓	✗	✗	✗	✗	✗	✓	✓	✗	✗
WP29 [35]	2014	✗	✗	✗	✓	✗	✗	✗	✗	✗	✗	✓	✓	✓	✗
Glancy [24]	2012	✗	✓	✗	✗	✓	✗	✗	✗	✗	✗	✓	✓	✓	✗
This work		✓	✗	✗	✗	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

[a] Privacy challenges. [b] Data subjects. [c] Data controllers. [d] Privacy-preserving.

3. GDPR Implications

While the level of privacy protection for ACSs has been effective with a number of technical solutions offered by pseudonymization and anonymization techniques, further discussed in Section 4, it is noteworthy that the respective legal framework as set forth in GDPR faces new regulatory challenges [27]. In the present section, we overview how the data protection regulation covers the principles of data processing and the rights of data subjects. This section sheds the light on the difficulties regarding the application of the data processing principles within the ACS ecosystem, and the risks to data subjects' rights during the data processing operations. In addition, we identify all the stakeholders who manage personal data, their interaction, their exact roles in the ACSs ecosystem based on the GDPR terminology, and their compliance requirements. Furthermore, by analyzing whether it is justifiable to incorporate DPIA in relation to the purposes of personal data collected, generated, and stored by ACS, and its value [9,15,36].

3.1. Data Processing Principles

Any processing of personal data should occur in the light of the legal principles as set in the body of the GDPR subsequently illustrated in Figure 2. According to Article 83 Section 4 [7] of GDPR, any failure of the entities involved in the processing to comply with the data processing principles may result in administrative fines and other sanctions. Certainly, the adherence to these principles by the controllers seems to be a fundamental requirement for the assessment of their compliance to the GDPR. To that end, data should be processed with respect to the following principles [7]:

1. Lawfulness, transparency, and fairness: where data collection practices are conducted based on a thorough understanding of the GDPR law and without hiding the type of collected data and the reason for its processing from the data subjects (Article 5, Sections 1.a and 6).
2. Purpose limitation: where the processing is approached based on the specified, explicit, and legitimate purposes with no further processing in a manner that is incompatible with those agreed on purposes (Article 5 Section 1.d).
3. Storage limitation: calling for data storage no longer than it is necessary for the purpose for which the personal data is processed (Article 5 Section 1.e).
4. Accuracy: where controllers should take necessary measures to process only correct data (Article 5 Section 1.b).
5. Data minimization: aiming to limit the amount of processed data to the lowest level and requiring data destruction once the purpose of the processing is completed (Article 5 Section 1.e).
6. Security: requiring data controllers to employ the appropriate technical and organizational measures designed to effectively implement integrity and confidentiality through PbD and PbDf principles (Article 25 Sections 1 & 2).
7. Accountability: requiring data controllers to put in place appropriate privacy-preserving measures that are able to demonstrate compliance to the regulation at any stage (Article 5 Section 2).

To support continuous automated navigation of ACS, data are used permanently. This perpetual data usage poses multiple challenges to the aforementioned principles. First of all, collecting large datasets of personal data to train the AI models jeopardizes the data minimization principle [21]. Secondly, further collection, use, transmission, or storage of personal data may exceed the purpose limitation principle from the beginning of these data processing operations. Additionally, innovative autonomous and connected technologies applied in the ACS complicate the implementation and monitoring of the proper security measures in line with as the data protection by PbD and PbDf principles. Finally, while being part of the public transport system, the ACS have multiple stakeholders involved in data processing. These stakeholders act as data controllers or data processors or their

roles change. For this reason, they have to provide personal data guarantees and mitigate potential privacy-related risks [2,18] based on the accountability principle.

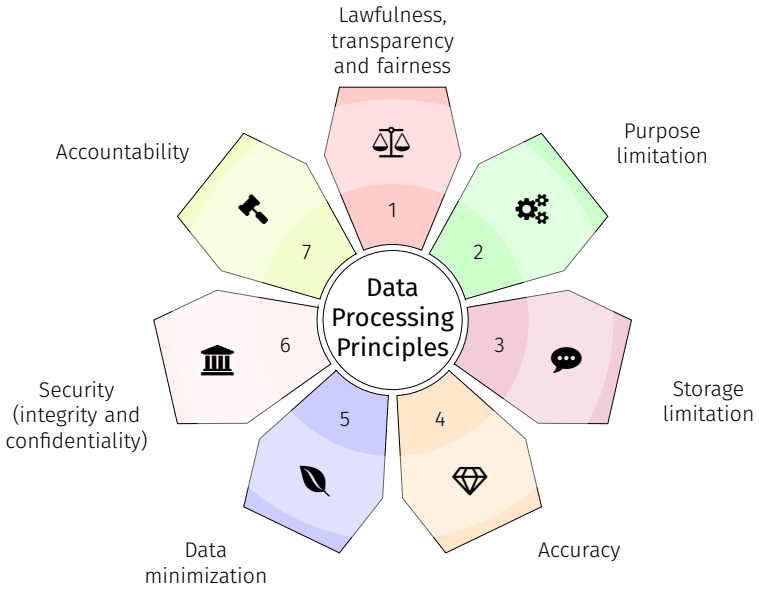

Figure 2. Data processing principles summary.

The subsequent sections are the result of our research to find the answers to the target research questions. Section 3.2 sets the basis for understanding the scope of RQ1 and provides the relevant findings. Then, Section 3.3 outlines our analysis in the form of a comprehensive study focused on the ACS's environment to address the RQ2.

3.2. Data Subjects Rights

To ensure transparency and fairness of data processing, the GDPR grants the data subjects specific control rights [20]. It should be mentioned that, at every stage of data processing, data subjects remain the owners of their personal data as verified by the right of access to personal data by virtue of Article 15 of the GDPR. By providing individuals access rights, the GDPR imposes a number of obligations to the entities that collect and process data, as well as allows the data protection authorities (DPAs) to ask for demonstrations of accountability or impose fines if data subjects' rights are not secured. Data controllers provide specific practices and technologies to the data subjects to control and exercise their rights during the entire data processing. For instance, the information about the exercise of rights is available in the privacy policy at the controller's website. Controllers can facilitate, specifically, the access, the deletion, the transfer, or the removal of personal data by providing modification settings [37].

One of the most crucial rights is the *right to be informed* (Articles 12–14). Prior to the processing of personal data by the controllers, the data subjects shall be informed, in a transparent way, of the identity of the data controller, the purpose of processing, the data recipients, the data retention period, and the data subjects' rights. In relation to the right to be informed, it is noted that data subjects should be informed in clear and plain language about any data breaches where their personal information is leaked and this leakage is likely to result in a high risk to their rights and freedoms as per Article 34 of the GDPR. For exercise of the right to be informed when the collection and use of data is intended for the vehicular automated decision-making and profiling purposes, the data subjects should receive "meaningful information about the logic involved" as well as the significance and the envisaged consequences of such processing for the data subject as per Article 13 Section 2.f [7]. The WP29 in its revised guidelines [38] clarified that the complexity of the technologies should not justify the lack of information. In the light of this clarification, the OEMs, other equipment manufacturers, and service providers, qualifying as data

controllers, should explain clearly to the ACS' data subjects the automated processing methods and their objectives.

Other rights include the *right to rectification* (Article 16). This right allows the data subjects to correct their personal data when it is inaccurate. *The right to erasure* (or right to be forgotten) allows the individuals to ask for their personal data to be deleted (Article 17). For instance, if an ACS operator (i.e., data subject) has consented to test the efficiency of a newly deployed system, he or she can withdraw such consent at any time and require the controller to erase all the information that was processed for the validation of this system. However, in following our example, if the OEM (i.e., data controller) demonstrates that he or she collects and processes personal data to detect shortages in the ACS with a view to improve safety of the vehicles and, subsequently, the safety of the public (overriding legitimate interests), the right to be forgotten may not be invoked pursuant to Article 17 Section 1.c of the GDPR. Further, the *right to restriction of processing* gives the data subjects the power to limit the processing of their personal data with several rules and exceptions (Article 18) [7]. Article 20 of the GDPR foresees the *right to portability*, permitting the data subjects to receive a copy of their personal data in a structured, commonly used, and machine-readable format and to transmit those data to another controller without impediment from the initial controller. An illustrative example of the exercise of data portability takes place when ACS passengers may require an ACS technical service provider to give all the information collected during a specific period and share it with a third party, such as an insurance company, in case of a car accident. As per Article 21, the data subjects also have the *right to object* to any processing of their personal data that they do not consent to. This latter right could require special attention and enforcement due to the large amounts of data collected and analyzed from the ACS. Meaning, continuous data from sensors, onboard processing, and big-data increase the complexity and impose additional requirements. Furthermore, the necessity for safe operating implies the capture of potential (inadvertently) subject data, where an objection would compromise its safe operation. From the above analysis, it is clear that the exercise of these rights is a vital part of the privacy interests of the data subjects, offering many possibilities to perform them and demanding the necessary guarantees from the controllers. The existence of several controllers in the ACS ecosystem means that common controllers should specify whose responsibility is the protection of the rights as per Article 26 Section 1 of the GDPR.

3.3. Data Controllers' and Data Processors' Compliance

The data controllers play a crucial role in the course of processing as they decide its purposes and means. A data controller can be a natural or legal person, a public authority, an agency, or other bodies. The core responsibilities of the data controllers involve [7]:

- The implementation of the data processing principles (Articles 5–11).
- To inform data subjects as elaborated in Section 3.1 and secure their rights (Articles 12–23).
- The implementation of security measures such as the deployment of privacy-preserving techniques discussed in Section 4 (Articles 5, 25, and 32).
- The arrangements with the joint controller (if any) (Article 26).
- The engagement of processors (Article 28).
- The notification of personal data breach to the relevant data protection authority (Article 33).
- The communication of personal data breach to the data subjects (Article 34).
- The realization of DPIA (Article 35).
- The designation of the Data Protection Officer (Article 37).
- The transfer of data to third countries (Chapter V, Articles 44–50).
- The communication with the DPAs (Articles 31, 36, and 37).
- The compliance with specific processing situations (Articles 85–91).
- Retain all the necessary documentation and records (as listed throughout GDPR articles).

The existence of two or more data controllers makes the involved parties "joint controllers" by virtue of Article 26 of the GDPR. Based on this article, two entities meet the requirements of joint controllers when they "jointly" determine the purposes and means of processing. If the criterion of joint decision is missing, then the stakeholders are not "joint" but sole data controllers. The EDPB with its recent guidelines on data controllers and data processors shed light on the ambiguity regarding the respective roles of joint controllers. Under Article 26, the joint controllers need to clearly define their respective compliance obligations, especially with regard to the exercising of data subjects' rights and the provision of information under Articles 13 and 14 of the GDPR. With regards to the legal relationship among the joint data controllers, the GDPR does not require a specific legal binding act. Nevertheless, for reasons of legal certainty, the EDPB recommends a contract or other legal act under EU law to identify the specific obligations that (joint) controllers have. In general, the principles of transparency and accountability of the data controller are applicable to all the joint controllers. Therefore, the responsibilities enlisted above extend to them. This practically means that joint controllers are responsible for demonstrating compliance to the GDPR based on their specific obligations as described in the contract or any other legal act, and noncompliance may result in administrative fines under Article 83 [36].

The Court of Justice of the European Union (CJEU) has adopted a broad definition regarding the notion of joint controllers and the allocation of their responsibilities. In a 2019 decision, in the Facebook Fan Pages case, the Court held that the administrators of Facebook fan pages are joint controllers together with Facebook. Although they had access only to anonymized statistical data and not to any personal data by creating the fan webpage, the administrators allowed to Facebook to collect data and made it joint controller. This judgment also stated that "the existence of joint responsibility does not necessarily imply equal responsibility of the various operators involved in the processing of personal data" [39]. Further, in another notable case, the Facebook Fashion ID case [40] it was held that the "existence of joint liability does not necessarily imply equal responsibility of the various operators engaged in the processing of personal data. On the contrary, those operators may be involved at different stages of that processing of personal data and to different degrees, with the result that the level of liability of each of them must be assessed with regard to all the relevant circumstances of the particular case". With these two decisions, the CJEU has validated a broad scope of joint controllers and has taken a clear position on the level of involvement in the data processing, the extent of responsibilities of joint controllers at every processing stage, and the degree of their liability.

The analysis of the joint controllership is relevant to the ACS because many service providers may qualify as joint controllers, as Figure 3 shows. The public transport operators (PTOs), the OEMs, and cloud service providers may determine the means and data processing purposes of data collection and usage by the ACS. In actuality, in the light of the Facebook ID case, it was held that the notion of joint controllership exists even if the data are not personal. This statement extends the possibilities of joint controllership to other actors who process even anonymized data in the ACS. As far as the obligations and responsibilities, a clear arrangement among the data controllers should arrange the compliance to the regulation and the data subject's rights. It should be noted that the decision about joint controllership may reveal that the very same of the above stakeholders may act as joint controllers in the event of commonly determining the means and purposes of processing. In other cases, though, this interaction concerns a party that processes personal information on behalf of another while both parties may have been involved in joint operations that precede or are subsequent in the overall chain of processing [41].

When the data processing is carried out on behalf of the data controller, the entity performing the processing acts as a data processor. Under Article 4, a data processor can be a natural or legal person, public authority, agency, or other body. Article 28(1) of the GDPR necessitates that only processors providing sufficient guarantees to implement appropriate technical and organizational measures should be engaged by a controller.

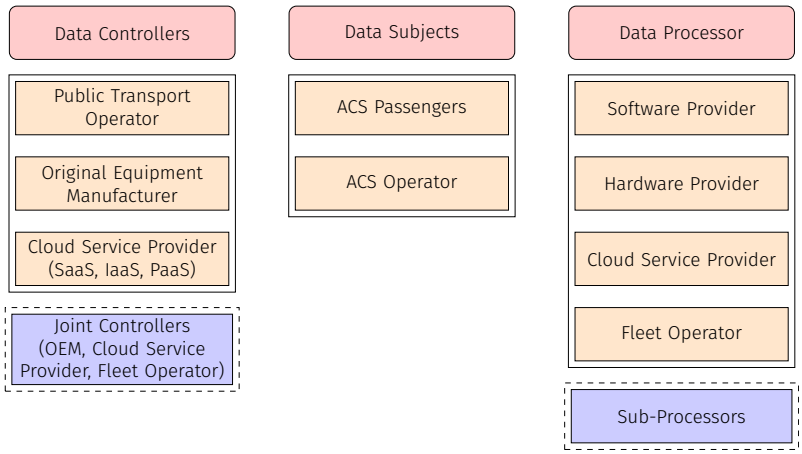

Figure 3. ACS's stakeholders classification per the GDPR terminology.

Although the definitions of the data controllers and data processors are clear by law, there is perplexity around the identification of the data controllers and the data processors in the ACS ecosystem [42]. This perplexity can be justified due to the fact that the roles of data controllers and data processors change rapidly. This may also mean that when a processor acts in a way that infringes the contract or another legal act or makes decisions in an autonomous way about the purpose and the the means of a specific processing operation, it may qualify as a controller (or a joint controller). When a processor acts as a delegate of the data controller with a mandate to perform specific tasks following specific instructions, this entity remains a processor as long as the duties do not deviate from these responsibilities. It is noted that data processors can engage in the data processing of other entities. The latter entities qualify as sub-processors. Pursuant to Article 29 Section 2, the data controller should provide a prior specific or general written authorization.

The GDPR has a detailed description of the legal relationship among controllers and processors. The basis is a contract or other legal act under Union or Member State law that is binding. This legal document sets out the subject matter and duration of the processing, the nature and purpose of the processing, the type of personal data and categories of data subjects, and the obligations and rights of the controller. Article 28 Section 3 of the GDPR enlists the specific processing duties that this legal act describes [7]. Inter alia, data processors should take all security measures required under Article 32, such as a data pseudonymization and encryption that will be reviewed in the section hereafter. Furthermore, they should assist controllers in fulfilling their obligations to respond to requests for exercising the data subject's rights. Finally, data processors should assist controllers' compliance with the obligations pursuant to Articles 32 to 36.

More specifically, as shown in Figure 3, in the essence of data processing in the ACSs we have the data subjects. The OEMs can be regarded as data controllers since they determine the means and purposes of processing. A fleet operator can be seen as joint data controller, under the conditions of joint controllership mentioned above, as they process data on behalf of the controllers [27]. Further examples of data processors include the distributors who perform legitimate remote monitoring, auto repair shops, navigation software providers and navigation apps developers, telematic service providers, or mobile network operators (MNOs) [21]. Per the required V2C communication within the ACS environment, the cloud service providers have a duplicated role to be considered, which can be a data processor or a data controller. This differentiation of the cloud service providers' roles derives from the fact that depending on the specific "service" they offer, i.e., software as a service (SaaS), platform as a service (PaaS) and/or infrastructure as a Sservice (IaaS), different roles and respective responsibilities are assigned.

3.4. Data Protection Impact Assessment

Pursuant to Article 35 of the GDPR, the data controllers are required to undertake a DPIA prior to data processing, especially when this processing is likely to result in a high risk to the rights and freedoms of natural persons. In particular, the GDPR enlists a non-exhaustive list of risk factors to be taken into consideration, and once assessed, a DPIA must be performed:

- A systematic and extensive evaluation of automated processing, including profiling and similar activities that have legal effects or affect the data subjects.
- Processing on a large scale of special categories of sensitive data such as racial or ethnic origin, political opinion, and of personal data relating to criminal convictions and offenses.
- A systematic monitoring of a publicly accessible area on a large scale.

Data processing at a large scale seems more relevant to the processing of personal data by and within the ACSs. Recital 91 provides information about what large scale means. It is the processing by making reference to the number of data subjects concerned, the volume of data processed, and the duration and the geographical extent of the data processing activity. It is understood that under these circumstances a DPIA is mandatory.

Nonetheless, even in cases where a DPIA is not legally mandated, data controllers should consider evaluating the data processing. Such practice would allow them to make a thorough assessment of the envisaged processing operations and to mitigate the risks as detailed in Article 35 [7]. In general, even in the cases where a DPIA is not required, it would be useful to carry out one as early as possible in the design process [37] under Articles 35(1) and 35(10) in combination with recitals 90 and 93 and with prior consultation of the DPAs under Article 36 of the GDPR. The DPIA should be publicly available and "continuously reviewed and regularly reassessed" [43].

Finally, conducting a DPIA is a tool of controllers' accountability. As such, it is part of the data controller's responsibilities to which they must show compliance, and, as mentioned, inability to be compliant to Articles 35 and 36 holds them accountable. However, due to the rapidly evolving nature of the autonomous technologies, the time for the realization of a DPIA, and the requirement of periodical assessment, a DPIA, in practice, might not fulfill its purpose and this challenges the controller's accountability.

4. The Interface of Privacy and Data Security in ACSs

Our next challenge that we would like to tackle uncovers how legal requirements, as derived from the GDPR, meet and coexist with technical solutions for preserving privacy within the ACS ecosystem. More specifically, we seek for an answer to the RQ3 to advance the ACS specific knowledge on the appropriateness of the existing technological solutions for the data privacy preservation. As stated in Section 3, PbD and PbDf require the implementation of several technological measures, such as pseudonymization and anonymization. Multiple scientific research reviews and law provisions [27,32,33] discussed those techniques to meet the integrity and confidentiality in addition to data minimization as part of the GDPR data processing principles (Figure 2). In addition, such techniques embed further reverse engineering risks, leading to re-identification of personal information as assessed by opinion 05/2014 of the WP29 [9]. The present section elevates the discrepancy of the required mitigation solutions by evaluating legal recommendations and technical approaches.

4.1. Privacy-Preserving Techniques Overview

The GDPR introduces pseudonymization and anonymization as prominent countermeasures to protect personal data since they lower the risk of linking personal data to their related data subjects. In legal terms, such schemes are forethought differently and independently, while, technically, they offer incommensurable levels of privacy-preserving.

We observed an ongoing debate regarding the effectiveness of privacy-preserving approaches and the inherent risk to re-identify data by using specific reverse engineering

technologies or by combining anonymized data with other information [44,45]. The opinion 05/2014 flagged three main risks related to non-robust privacy-preserving methods [35]:

- "Singling out": when the data subject's data are isolated to identify the natural person attributes or track their localization.
- "Linkability": when correlation of multiple records, at least two, of the same individual leads to the identification of the person.
- "Inference": when a data subject's data are deducted from a dataset or through additional information leading to their identification [30].

By extending the opinion 05/2014 risks, we present a clear categorization and highlight the limitations of the most relevant pseudonymization and anonymization schemes for the ACS environment.

4.1.1. Pseudonymization

The pseudonymization technique is introduced in Article 4(5) of the GDPR [7] as a privacy-preserving measure and a safeguard that supports data processors and controllers to meet the data protection obligations. Pseudonymization does not remove all identifying information from the personal data but merely reduces the linkability by hiding the identity of the data subjects from third parties [32,33,35]. Technically, by using pseudonymization, the data subjects identifiers are substituted by a code, hiding the sensitive data, which can be re-identified using a key [46].

Encryption illustrates perfectly the pseudonymization technique as it uses secret keys that can reversely be decrypted and hence make personal data readable. Within the automated driving context, multiple researchers [47,48] introduced several privacy-preserving encryption schemes such as conditional privacy-preserving authentication (CPPA) and group signature and identity-based signature (GSIS) where only predefined and trusted authorities are legitimated to decrypt the keys within the vehicular communication system. To that end, encryption increases the confidentiality and lowers the risk of misusing personal data. However, such protection depends on the strength of the algorithm that can be built either using public/private keys or hash functions which guarantee different levels of security [28].

Tokenization is another pseudonymization scheme that can be considered within the vehicular environment. It is the process of replacing sensitive characters by other random non-sensitive values that cannot be mathematically computed from the data source length [49]. Although, when applied to location data, even if the identities are replaced by pseudonyms, individuals can still be singled out and tracked, as demonstrated by De Montjoye et al. [50].

Zero-knowledge proof (ZKP) is a promising cryptographic protocol that is based on an exchange of messages between a prover and a verifier where the prover has a secret but does not reveal information about the secret, per se. However, the prover should provide more information about their secret to establish the trust with the verifier [51]. Gabay et al. [52] demonstrated how ZKP can be used for electric vehicle authentication within the vehicle-to-everything (V2X) communication. ENISA [32,33] presented the ZKP as a pseudonymization technique implementing authentication and which increases confidentiality and data minimization required by the GDPR, although, the additional collected information by the verifier presents a risk for a future data leakage and again impacts the individual's privacy.

Consequently, not all pseudonymization techniques are suitable to any sensitive data type within the ACS environment. In addition, data controllers and data processors have to keep track of such techniques to make the processing part of the GDPR scope. Furthermore, the combination of pseudonymization with further anonymization-based schemes is much recommended and would definitely strengthen the privacy within the ACS ecosystem [33,35,45].

4.1.2. Anonymization

The anonymization technique is where both identifiers and keys are removed and the link between individuals and their locations is concealed for identity perturbation purposes [53]. It is considered as a solution to permanently remove personal data leading to identification of a natural person [7]. Recital 26 excludes anonymous information from the GDPR scope as it is assumed to be a long-lasting and irreversible solution. Similarly, the ePrivacy Directive [10] emphasizes erasing or anonymizing traffic and location data within the vehicular communications (Article 6(1) and Article 9(1)).Despite the best efforts on removing identifiers or even the keys, privacy can still be compromised [46].

There is a wide spectrum of anonymization techniques that can be applied to protect vehicular data, but their definitions are noticeably overlapping and commonly mistaken [49]. Data randomization (also called noising) and data generalization are the core anonymization families according to opinion 05/2014 of the WP29 [35]. Researchers discussed multiple *randomization* approaches that vary from noise addition, differential privacy, and swapping, to masking.

Noise addition occurs when an appropriate and proportionate noise is added to randomly modify the sensitive data [31]. Applied to location data, noise addition would result in falsifying, for example, the exact differential Global Positioning Systems (GPS) coordinates with an additional 10 km. Although, with some knowledge about the other dataset's attributes, an attacker can filter out the noise or link it to another database to regenerate the missing information [35]. Differential privacy is a different approach of noise addition that can be applied to larger datasets with the promise of learning useful information about a population while the individual's privacy is conserved. By adding random noise to a set of data before learning from it, private inputs are hidden without impacting the result accuracy [54]. However, simple efforts would lead to inferring or predicting individuals and/or location attributes [55]. Swapping, also called permutation, is a randomization technique that leads to falsifying a data entry by shuffling personal information within a dataset. Although, if a dataset has redundant attributes, the permutation will not be efficient and will lead to potential failure of the intended anonymization goals [35].

Data masking is frequently considered as a further randomization technique that refers to the process of hiding true values to make the sensitive information inconceivable, such as by replacing characters with asterisks [49]. Nevertheless, data masking remains a generic term that refers to the process of hiding true values [33], which at the end can refer to pseudonymization or anonymization [56]. The key measure to classify data masking as either an anonymization or a pseudonymization scheme is the eventual reversibility property. According to the GDPR (Recitals 28 and 29, Article 4(5)) [7], a reverted process is considered as a pseudonymization technique which, unlike anonymization, makes it subject to different legal provisions and additional technical and organization efforts from data controllers. Therefore, in our view, we consider data masking to be a parent node for pseudonymization and anonymization, as depicted in Figure 4. This reflects how easily misunderstandings about the anonymization techniques can lead to unintentional unlawful processing [45].

Generalization is when the attribute's value is substituted with a broad but semantically logical value [57]. Applied to personal data within the automated driving system, generalization can replace the specific location data with a broad city or country name. Such techniques are implemented by generalizing or diluting the identifiers in a way that replaces the week instead of the day and the country instead of a city location [35]. The main limitation of the generalization theories is that it cannot be applied to all types of data such as names or data entry identifiers, yet they remain efficient for location data. Meanwhile, even with the most refined generalization technique, the risk of inference depends on the adversary's knowledge and the selected anonymization parameters [58]. The key generalization techniques, pointed out by the WP29 [35] and ENISA [33], and which can be applied to the ACS personal data, are:

- k-anonymity/aggregation: hides a data subject in the crowd (called also equivalent class) by grouping their attributes with k-1 other individuals. The scheme provides a convenient protection from being singled out, though the inference risk remains important [28].
- l-diversity: handles the k-anonymity limitation by ensuring that in every crowd there are l-different values. Such difference reduces the inference risk but does not completely eliminate it [59].
- t-closeness: is a refinement of the l-diversity theory that aims to set a t-threshold by computing the resemblance of a sensitive value distribution within the equivalent class in comparison to the attribute distribution in the whole dataset [49].

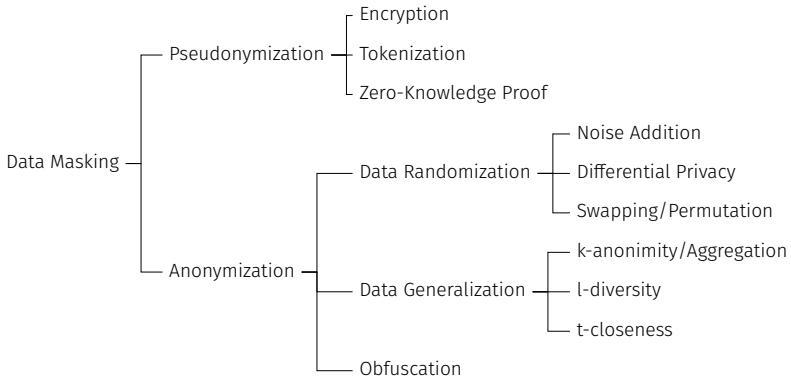

Figure 4. Classification of privacy-preserving techniques.

More specific to location data, *obfuscation* techniques are meant to purposely return incorrect location information to make the unauthorized tracking difficult [60]. Although the obfuscation approach is considered as a multifaceted scheme that does not have a straightforward definition in the literature, multiple researchers interpret obfuscation as a type of data masking or noise addition [31,49,61], while others [34,53] define it as a segregated principle from anonymization and pseudonymization. As depicted in Figure 4, the obfuscation has been classified as an anonymization scheme for its theoretical irreversibly. Despite the definition, the obfuscation techniques come with the promise to definitively break the linkability between the individuals' data and their location in combination with time, though it is true that obfuscation techniques offer geo-indistinguishability and protect location data. However, such mechanisms remain as vulnerable to re-identification attacks as other anonymization theories, as per Kawamoto and Murakami's attack simulation [60,62].

To that end, there is no "one-size-fits-all" [31], and privacy-preserving is a process of combining accurate schemes depending on the nature of personal data, the acceptable level of re-identification with regard to data protection obligations, and the intended data usability.

4.2. Privacy-Preserving Pitfalls

The aforementioned discussion highlights the discrepancy between regulator's efforts and scientific works. From the legal perspective, we summarize below our key findings:

- Pseudonymization is indicated as an appropriate "technical and organizational measure" for data protection (Article 25(1)) [7] without proposing the mixed use of pseudonymization and anonymization schemes to make the privacy-preserving level even higher.
- The GDPR considers the pseudonymization to be a reversible process and anonymization to be permanent without highlighting the de-anonymization risk over the time.
- The opinion 05/2014 WP29 introduced the main risks and mitigation solutions against the re-identification risk of anonymization; though the recommendations do not cope

with the rapid evolving technologies as more risks and countermeasures are worthy to be extended by the WP29, which is currently represented by EDPB.

- There is no precision from legal provisions about which anonymization technique should be applied to each context, even if the likelihood of re-identifying personal information might vary from one technique to another [35,46].

Nonetheless, from the technical perspective:

- The re-identification likelihood can never be zero.
- Anonymization is not everlasting, as it can be reverted in the future. Instead of a one-time operation, it should be assessed continuously.
- The choice of privacy-preserving scheme depends on the nature of the attribute of the private data itself. To illustrate, techniques applied to anonymize a data subject's name or data entry identifier might not be suitable for location data.
- There is a common misunderstanding, as pointed out in Section 4.1.2, in defining data masking as a subcategory of anonymization techniques. However, this technique is broad enough to embed both pseudonymization and anonymization, which would apply to different data protection obligations.
- Some researchers wrongly discussed encryption as an anonymization scheme, though it should be considered as a powerful pseudonymization technique.
- There is no unique solution that fits all processing, but the privacy-preserving technique should be selected on a case-by-case basis and depending on technologies involved within the ACS.

To that end, the implementation of the GDPR principles may fall short even without the compliance violation. Pseudonymization and anonymization help to comply with the data protection obligations. However, they just assist in reducing the risks and not to completely preserve privacy within the automated vehicle ecosystem. Hence, the GDPR-ordained implementation requires further guidance from DPAs and the EDPB to keep pace with the rapidly evolving technologies embedded within the ACS environment.

5. Conclusions and Future Work

Many ACS manufacturers envisage the deployment of ACS for public transport in the coming years. According to The Global Market Insights report [63], its value will rise from USD 1 billion in 2021 to USD 4 billion in 2028. Addressing the privacy concerns is crucial and a major challenge for the adoption of the ACS. The GDPR EU law provides an unprecedented level of data protection. Nevertheless, the collection and further processing of personal data raise crucial privacy concerns. We have identified the challenges for the principles of purpose limitation and data minimization resulting from the immense amount of data processed by the ACS. These principles, encapsulated within the concepts of PbD and PbDf, aim to mitigate any risks to the fundamental rights and freedoms of data subjects by the implementation of appropriate technical and organizational measures. These risks are also associated with the multiple data controllers and data processors, and their roles in the lifecycle of data processing as stakeholders of ACS ecosystems. The multiplicity of stakeholders who are involved in collection, transfer, and storage of data complicates the transparency of processing, the adoption of the appropriate and sufficient security measures, and, finally, their compliance on the basis of the accountability principle.

With regard to the responsibility of the data processors and sub-processors, new regulatory efforts have been initiated at the European level [64] with a view to reinforce the responsibilities throughout the supply chain. Similar efforts could be initiated for the management of the automated and connected vehicles supply chain. It should be highlighted that the efforts to address the regulatory complexity of the innovative ecosystems should be approached in an interdisciplinary way by examining all the existing regulatory aspects in the domains of AI, privacy, and human rights, together with the involving technologies per sector. For instance, in 2021, the draft of the AI act was circulated by the EU [65], destined to impact the critical infrastructures, the transport sector being one of them.

In this work, we raise the complexity of the stakeholders' duties and how they may accumulate multiple roles and obligations. Our recommendation to the stakeholders, who act as controllers and processors, is to encompass all those factors to protect fundamental rights and prevent potential data breaches within the ACS environment. We advise to push for deploying specific use cases with constant security control, monitoring, and assessment overtime. Additionally, the present paper discusses the discrepancy between the privacy-preserving techniques and the way the same techniques are advocated by the GDPR. As the existing data protection laws leave much uncertainty about the effectiveness of pseudonymization and anonymization as major countermeasures, the present work recommends a combination of legal provisions, more fine-grained definitions and a description of specific technologies preventing from re-identifying personal data within the automated driving ecosystem. Additionally, we endorse continuous control and aggregated countermeasures to exchange data and knowledge leading towards lawful processing, secure interoperability, and safe integration of ACS to the new mobility concepts. As a matter of fact, the present findings can be considered as the foundation for a potential upgrade of the European personal data protection regulations to provide more granular insights on the ACS's stakeholders' roles and reconsider the de-anonymization risks for an optimal data privacy protection within the IoV environment.

As an ongoing effort, and under the umbrella of the AVENUE project, a data privacy assessment will be implemented to rate, audit, and quantify the technical and organization mitigation techniques in place with regard to the GDPR requirements. Then, by elevating the findings from the present work, we intend to provide a risk management plan to the potential privacy threats through a real testbed on a predefined testing site, and over a vehicle of SAE automation level four.

As a future work regarding the security and, more specifically, the implementation of PbD within the ACS landscape, there is a need to have more clarity as there is no consensus nor directive on the exact definition or methodology of what this implementation should look like. As an example, collecting data from non-users (data subjects that are not using the ACS) [21] should be looked at as a future work; as the ACSs' cameras can collect images of the outside environment of the vehicle, personal video surveillance concerns should be raised to protect not only the users but also non-users of the ACS. Moreover, another open issue requiring further research is the insurance regimes in case of hacked ACS or personal data breach. The classic vehicular insurance model is more likely to evolve with the widespread integration of ACS. The new transition from driver-controlled public transportation vehicles to ACS operating without human interaction will obviously impact the insurance fees and policies and their implication to the public transport sector.

Author Contributions: Conceptualization, M.B. and S.K.; methodology, M.B. and S.K.; investigation, M.B.; writing—original draft preparation, M.B. and S.K.; writing—review and editing, M.B., A.C., S.K. and N.A.N.; visualization, N.A.N.; supervision, N.A.N. and A.C. All authors have read and agreed to the published version of the manuscript.

Funding: This work has been funded and supported by the European Union's Horizon 2020 Research and Innovation Programme through AVENUE project (https://h2020-avenue.eu/ accessed on 24 March 2022) under grant agreement No. 769033, nIoVe project (https://www.niove.eu/ accessed on 24 March 2022) under grant agreement No. 833742 and SHOW project (https://show-project.eu/ accessed on 24 March 2022) under grant agreement No. 875530.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The photo in Figure 1 was provided by the Geneva Public Transport, Switzerland.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

ACS	Automated city shuttle
AI	Artificial intelligence
CAV	Connected automated vehicle
CJEU	Court of Justice of the European Union
CPPA	Conditional privacy-preserving authentication
DPA	Data protection authorities
DPIA	Data protection impact assessment
EDPB	European Data Protection Board
ENISA	European Union Agency for Cybersecurity
EU	European Union
GDPR	General Data Protection Regulation
GPS	Differential Global Positioning Systems
GSIS	Group signature and identity-based signature
IaaS	Infrastructure as a service
IoT	Internet of Things
IoV	Internet of Vehicles
ITS	Intelligence transport system
LBS	Location-based services
MaaS	Mobility as a service
MNO	Mobile network operator
NIS	Network and information security
OEM	Automobile manufacturer
PaaS	Platform as a service
PbD	Privacy by design
PbDf	Privacy by default
PIPEDA	Personal Information Protection and Electronic Documents Act
PTO	Public transport operator
RSU	Roadside unit
SaaS	Software as a service
SAE	Society of Automotive Engineering
V2C	Vehicle-to-cloud
V2I	Vehicle-to-infrastructure
V2P	Vehicle-to-pedestrian
V2V	Vehicle-to-vehicle
V2X	Vehicle-to-everything
WP29	Article 29Working Party
ZKP	Zero-knowledge proof

References

1. Balboni, P.; Botsi, A.; Francis, K.; Barata, M.T. Designing Connected and Automated Vehicles around Legal and Ethical Concerns: Data Protection as a Corporate Social Responsibility. In Proceedings of the WAIEL2020, Athens, Greece, 3 September 2020.
2. Ainsalu, J.; Arffman, V.; Bellone, M.; Ellner, M.; Haapamäki, T.; Haavisto, N.; Josefson, E.; Ismailogullari, A.; Lee, B.; Madland, O.; et al. State of the art of automated buses. *Sustainability* **2018**, *10*, 3118. [CrossRef]
3. Konstantas, D. From Demonstrator to Public Service: The AVENUE Experience. In *The Robomobility Revolution of Urban Public Transport*; Mira-Bonnardel, S., Antonialli, F., Attias, D., Eds.; Springler: Cham, Switzerland, 2021; pp. 107–130. [CrossRef]
4. Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles. SAE Technical Report J3016_202104. 2018. Available online: https://www.sae.org/standards/content/j3016_202104 (accessed on 23 March 2022).
5. Elliott, D.; Keen, W.; Miao, L. Recent advances in connected and automated vehicles. *J. Traffic Transp. Eng. (Engl. Ed.)* **2019**, *6*, 109–131. [CrossRef]
6. Veitas, V.K.; Delaere, S. In-vehicle data recording, storage and access management in autonomous vehicles. *arXiv* **2018**, arXiv:1806.03243.
7. European Union. Regulation (EU) 2016/679: The European Parliament and of the Council of 27 April 2016 on the Protection of Natural Persons with Regard to the Processing of Personal Data and on the Free Movement of Such Data. *Off. J. Eur. Communities* **2016**, *L119*, 1–88.
8. Smith, G.; Smith, G. *Making Mobility-as-a-Service*; Chalmers University of Technology: Gothenburg, Sweden, 2020.

9. Article 29 Data Protection Working Party. *Opinion 03/2017 on Processing Personal Data in the Context of Cooperative Intelligent Transport Systems (C-ITS)-217/EN-WP 252*; Technical Report October; European Commission: Brussels, Belgium, 2017.
10. European Union. Directive 2002/58/EC: The European Parliament and of the Council of 12 July 2002 Concerning the Processing of Personal Data and the Protection of Privacy in the Electronic Communications Sector (Directive on Privacy and Electronic Communications) (L 201). *Off. J. Eur. Communities* **2002**, *L201*, 37–47.
11. The European Parliament and the Council of the European Union. *Directive (EU) 2016/1148 The European Parliament and of The Council—NIS Dircetive 1*; Technical Report; European Commission: Brussels, Belgium, 2016.
12. The European Parliament and the Council of the European Union. *Proposal for a Directive Directive (EU) 2016/1148 of the European Parliament and of the Council—NIS Dircetive 2*; Technical Report; European Commission: Brussels, Belgium, 2020.
13. Costantini, F.; Thomopoulos, N.; Steibel, F.; Curl, A.; Lugano, G.; Kováčiková, T. Autonomous vehicles in a GDPR era: An international comparison. *Adv. Transp. Policy Plan.* **2020**, *5*, 191–213. [CrossRef]
14. OneTrust Data Guidance. Comparing Privacy Laws: GDPR vs. PIPEDA. Technical Report. 2020. Available online: https://www.dataguidance.com/sites/default/files/gdpr_v_pipeda.pdf (accessed on 15 April 2022).
15. Australia, N. *Regulating Government Access to C-ITS and Automated Vehicle Data*; Technical Report September; National Transport Commission: Melbourne, VIC, Australia, 2018.
16. George, D.; Reutimann, K.; Larrieux, A.T. GDPR bypass by design? Transient processing of data under the GDPR. *Int. Data Priv. Law* **2019**, *9*, 285–298. [CrossRef]
17. Taeihagh, A.; Lim, H.S.M. Governing autonomous vehicles: Emerging responses for safety, liability, privacy, cybersecurity, and industry risks. *Transp. Rev.* **2019**, *39*, 103–128. [CrossRef]
18. Lim, H.S.M.; Taeihagh, A. Autonomous vehicles for smart and sustainable cities: An in-depth exploration of privacy and cybersecurity implications. *Energies* **2018**, *11*, 1062. [CrossRef]
19. Pattinson, J.A.; Chen, H.; Basu, S. Legal issues in automated vehicles: Critically considering the potential role of consent and interactive digital interfaces. *Humanit. Soc. Sci. Commun.* **2020**, *7*, 1–10. [CrossRef]
20. Vallet, F. The GDPR and Its Application in Connected Vehicles—Compliance and Good Practices. In *Electronic Components and Systems for Automotive Applications*; Springer: Cham, Switzerland, 2019; pp. 245–254. [CrossRef]
21. Krontiris, I.; Grammenou, K.; Terzidou, K.; Zacharopoulou, M.; Tsikintikou, M.; Baladima, F.; Sakellari, C.; Kaouras, K. Autonomous Vehicles: Data Protection and Ethical Considerations. In Proceedings of the CSCS 2020: ACM Computer Science in Cars Symposium, Feldkirchen, Germany, 2 December 2020. [CrossRef]
22. Bastos, D.; El-Mousa, F.; Giubilo, F. GDPR Privacy Implications for the Internet of Things. In Proceedings of the 4th Annual IoT Security Foundation Conference, London, UK, 4 December 2018.
23. Collingwood, L. Privacy implications and liability issues of autonomous vehicles. *Inf. Commun. Technol. Law* **2017**, *26*, 32–45. [CrossRef]
24. Glancy, D.J. Santa Clara Law Review Privacy in Autonomous Vehicles. *Number Artic.* **2012**, *52*, 12–14.
25. Karnouskos, S.; Kerschbaum, F. Privacy and integrity considerations in hyperconnected autonomous vehicles. *Proc. IEEE* **2018**, *106*, 160–170. [CrossRef]
26. Hes, R.L.; Borking, J.J. *Privacy-Enhancing Technologies: The Path to Anonymity*; Registratiekamer: The Hague, The Netherlands, 1988.
27. Mulder, T.; Vellinga, N.E. Exploring data protection challenges of automated driving. *Comput. Law Secur. Rev.* **2021**, *40*, 105530. [CrossRef]
28. Ribeiro, S.L.; Nakamura, E.T. Privacy Protection with Pseudonymization and Anonymization in a Health IoT System: Results from OCARIoT. In Proceedings of the 2019 IEEE 19th International Conference on Bioinformatics and Bioengineering (BIBE), Athens, Greece, 28–30 October 2019; pp. 904–908. [CrossRef]
29. Brasher, E.A. Addressing the Failure of Anonymization: Guidance from the European Union's General Data Protection Regulation. *Columbia Bus. Law Rev.* **2018**, *2018*, 209–253.
30. Li, H.; Ma, D.; Medjahed, B.; Kim, Y.S.; Mitra, P. Analyzing and Preventing Data Privacy Leakage in Connected Vehicle Services. *Sae Int. J. Adv. Curr. Pract. Mobil.* **2019**, *1*, 1035–1045. [CrossRef]
31. Löbner, S.; Tronnier, F.; Pape, S.; Rannenberg, K. Comparison of De-Identification Techniques for privacy-preserving Data Analysis in Vehicular Data Sharing. In *Computer Science in Cars Symposium*; ACM: New York, NY, USA, 2021; pp. 1–11. [CrossRef]
32. ENISA. *Data Pseudonymisation: Advanced Techniques & Use Cases*; Technical Report; ENISA: Athens, Greece, 2021. [CrossRef]
33. European Union Agency for Cybersecurity. *Data Protection Engineering*; Technical Report; ENISA: Athens, Greece, 2022. [CrossRef]
34. Lim, J.; Yu, H.; Kim, K.; Kim, M.; Lee, S.B. Preserving Location Privacy of Connected Vehicles with Highly Accurate Location Updates. *IEEE Commun. Lett.* **2017**, *21*, 540–543. [CrossRef]
35. Article 29 Protection Working Party. *Opinion 05/2014 on Anonymisation Techniques*; Technical Report April; European Commission: Bussels, Belgium, 2014.
36. EDPB. *Guidelines 07/2020 on the Concepts of Controller and Processor in the GDPR*; Technical Report; EDPB: Brussels, Belgium, 2020.
37. EDPB. *Guidelines 1/2020 on Processing Personal Data in the Context of Connected Vehicles and Mobility Related Applications*; Technical Report March; European Data protection Board: Brussels, Belgium, 2021.

38. Article 29 Data Protection Working Party. *Guidelines on Automated Individual Decision-Making and Profiling for the Purposes of Regulation 2016/679*; Technical Report; Article 29 WP; European Commission: Brussels, Belgium, 2018.
39. Curia Caselaw. Judgment of The Court. 2018. Available online: https://curia.europa.eu/juris/document/document.jsf?docid=202543&doclang=EN (accessed on 25 February 2022).
40. Curia Caselaw. Judgment of the Court on Facebook Ireland Ltd. 2019. Available online: https://curia.europa.eu/juris/document/document.jsf?docid=216555&mode=req&pageIndex=1&dir=&occ=first&part=1&text=&doclang=EN&cid=4232790 (accessed on 10 March 2022).
41. European Data Protection Supervisor. *EDPS Guidelines on the Concepts of Controller, Processor and Joint Controllership under Regulation (EU) 2018/1725*; Technical Report; EDPS: Brussels, Belgium, 2019.
42. Mulder, T.; Vellinga, N. Handing over the Wheel, Giving up Your Privacy? In Proceedings of the 13th ITS Europe Congress, Eindhoven, The Netherlands, 3–6 June 2019.
43. Article 29 Data Protection Working Party. *Guidelines on Data Protection Impact Assessment (DPIA) and Determining Whether Processing Is "Likely to Result in a High Risk" for the Purposes of Regulation 2016/679*; Technical Report; European Commission: Bussels, Belgium, 2017.
44. Bu-Pasha, S. Location Data, Personal Data Protection and Privacy in Mobile Device Usage: An EU Law Perspective. Ph.D. Thesis, Faculty of Law, Helsinki, Finland, 2018.
45. AEPD. *Ten Misunderstandings Related to Anonymisation*; Technical Report 1; AEPD: Madrid, Spain, 2019.
46. Vokinger, K.N.; Stekhoven, D.J.; Krauthammer, M. Lost in Anonymization—A Data Anonymization Reference Classification Merging Legal and Technical Considerations. *J. Law Med. Ethics* **2020**, *48*, 228–231. [CrossRef]
47. Manivannan, D.; Moni, S.S.; Zeadally, S. Secure authentication and privacy-preserving techniques in Vehicular Ad-hoc NETworks (VANETs). *Veh. Commun.* **2020**, *25*, 100247. [CrossRef]
48. Dibaei, M.; Zheng, X.; Jiang, K.; Abbas, R.; Liu, S.; Zhang, Y.; Xiang, Y.; Yu, S. Attacks and defences on intelligent connected vehicles: A survey. *Digit. Commun. Netw.* **2020**, *6*, 399–421. [CrossRef]
49. Ouazzani, Z.E.; Bakkali, H.E. A Classification of non-Cryptographic Anonymization Techniques ensuring Privacy in Big Data. *Int. J. Commun. Netw. Inf. Secur. (IJCNIS)* **2020**, *12*, 142–152.
50. De Montjoye, Y.A.; Hidalgo, C.A.; Verleysen, M.; Blondel, V.D. Unique in the Crowd: The privacy bounds of human mobility. *Sci. Rep.* **2013**, *3*, 1376. [CrossRef] [PubMed]
51. Wan, Z.; Guan, Z.; Zhou, Y.; Ren, K. Zk-AuthFeed: How to feed authenticated data into smart contract with zero knowledge. In Proceedings of the 2019 2nd IEEE International Conference on Blockchain, Blockchain 2019, Atlanta, GA, USA, 14–17 July 2019; pp. 83–90. [CrossRef]
52. Gabay, D.; Akkaya, K.; Cebe, M. Privacy-Preserving Authentication Scheme for Connected Electric Vehicles Using Blockchain and Zero Knowledge Proofs. *IEEE Trans. Veh. Technol.* **2020**, *69*, 5760–5772. [CrossRef]
53. Takbiri, N.; Houmansadr, A.; Goeckel, D.L.; Pishro-Nik, H. Limits of location privacy under anonymization and obfuscation. In Proceedings of the 2017 IEEE International Symposium on Information Theory (ISIT), Aachen, Germany, 25–30 June 2017; pp. 764–768. [CrossRef]
54. Dwork, C.; Kohli, N.; Mulligan, D. Differential Privacy in Practice: Expose your Epsilons! *J. Priv. Confidentiality* **2019**, *9*. [CrossRef]
55. Ha, T.; Dang, T.K.; Dang, T.T.; Truong, T.A.; Nguyen, M.T. Differential Privacy in Deep Learning: An Overview. In Proceedings of the 2019 International Conference on Advanced Computing and Applications (ACOMP), Nha Trang, Vietnam, 26–28 November 2019; pp. 97–102. [CrossRef]
56. Tachepun, C.; Thammaboosadee, S. A Data Masking Guideline for Optimizing Insights and Privacy Under GDPR Compliance. In Proceedings of the 11th International Conference on Advances in Information Technology, Bangkok, Thailand, 1–3 July 2020; ACM: New York, NY, USA, 2020; pp. 1–9. [CrossRef]
57. Murthy, S.; Abu Bakar, A.; Abdul Rahim, F.; Ramli, R. A Comparative Study of Data Anonymization Techniques. In Proceedings of the 2019 IEEE 5th International Conference on Big Data Security on Cloud (BigDataSecurity), IEEE International Conference on High Performance and Smart Computing, (HPSC) and IEEE International Conference on Intelligent Data and Security (IDS), Washington, DC, USA, 27–29 May 2019; pp. 306–309. [CrossRef]
58. Wang, J.; Cai, Z.; Yu, J. Achieving Personalized k-Anonymity-Based Content Privacy for Autonomous Vehicles in CPS. *IEEE Trans. Ind. Inform.* **2020**, *16*, 4242–4251. [CrossRef]
59. Sangeetha, S.; Sudha Sadasivam, G. Privacy of Big Data: A Review. In *Handbook of Big Data and IoT Security*; Springer: Cham, Switzerland, 2019; pp. 5–23. [CrossRef]
60. Kawamoto, Y.; Murakami, T. On the Anonymization of Differentially Private Location Obfuscation. In Proceedings of the 2018 International Symposium on Information Theory and Its Applications (ISITA), Singapore, 28–31 October 2018.
61. Lu, Z.; Qu, G.; Liu, Z. A Survey on Recent Advances in Vehicular Network Security, Trust, and Privacy. *IEEE Trans. Intell. Transp. Syst.* **2019**, *20*, 760–776. [CrossRef]
62. Murakami, T. A Succinct Model for Re-identification of Mobility Traces Based on Small Training Data; A Succinct Model for Re-identification of Mobility Traces Based on Small Training Data. In Proceedings of the 2018 International Symposium on Information Theory and Its Applications (ISITA), Singapore, 28–31 October 2018.
63. Wadhwani, P.; Saha, P. *Autonomous Bus Market Trends 2022–2028. Size Analysis Report*; Technical Report; Global Market Insights: Selbyville, DE, USA, 2021.

64. Center for Strategic and International Studies. *European Union Releases Draft Mandatory Human Rights and Environmental Due Diligence Directive*; Center for Strategic and International Studies: Washington, DC, USA, 2022.
65. Evas, T.; Heflich, A. *Artificial Intelligence in Road Transport*; Technical Report; European Parliament: Strasbourg, France, 2021.