Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 280 (2026) 398–405

Procedia Computer Science

www.elsevier.com/locate/procedia

The 17th International Conference on Ambient Systems, Networks and Technologies (ANT)
April 14-16, 2026, Istanbul, Türkiye

Conceptual Context-based Access Control Framework for Improving Privacy in Autonomous Vehicles

Naz Nebile Karataş*, Ethiopia Nigussie, Seppo Virtanen, Jouni Isoaho

Department of Computing, Faculty of Technology, University of Turku, Turku and Finland

Abstract

Autonomous vehicles (AVs) collect large amounts of raw data from various sensors. Often the data collection, processing and communication raises privacy concerns. These sensor data are integrated with contextual information. This transforms the data into usable metadata that enables real-time processing and analysis. Metadata is privacy sensitive, which requires General Data Protection Regulation (GDPR) compliance and comes with a set of compliance obligations for data controllers and processors. In this research, we propose a conceptual context-based access control framework that enhances the privacy of data collected and processed in autonomous vehicles. The proposed access control framework introduces a layered and modularized architecture. The use of a layered architecture allows the handling of differing privacy risk levels between real-time and historical data. As the potential risk of revealing a specific person in combined data is higher in historical data compared to real-time data, a containment method is proposed for real-time data and sandboxing is proposed for historical data. The data that are processed by AVs consists of personal and non-personal data. While personal data are already a privacy-breaching risk solely, there is also a risk of identifying a specific person by the combination of non-personal data. The proposed access control framework classifies the data, containerizes the data through sandboxing and containment methods, and applies specific access control according to the data context and its privacy implications.

Keywords: autonomous vehicles; privacy, access control; privacy-preserving technologies; sandboxing, data protection

1. Introduction

The Autonomous vehicles (AV) market is projected to reach over 2.3 trillion dollars worldwide by 2030 [15], it is crucial to prepare both legal regulations and privacy-preserving technologies while we still have time before the destruction of individual privacy, which is defined as the queen of social harms [22, p.20], becomes more severe. Current legal regulations are not fully and directly sufficient or prepared to solve possible privacy issues related to AVs.

* Corresponding author. Tel.: +0-000-000-0000 ; fax: +0-000-000-0000.
E-mail address: nnkara@utu.fi

AVs are equipped with many sensors, such as LiDAR, radar, and cameras, and these sensors are needed to ensure safe driving. These sensors collect the raw data, and the onboard systems in the vehicle add contextual information like timestamps and coordinates before processing and storing. This process transfers the raw data into metadata, which is structured, analyzable information. Metadata can be a mix of personal and non-personal datasets, which is hard to separate. When personal and non-personal data are stored together and the separation of personal and non-personal data is either impossible or not feasible in the dataset, the entire dataset becomes legally classified as personal data, referred to as inextricably linked [4, p.10]. Therefore, they need to be collected and processed in accordance with GDPR unless they are fully anonymised data. These requirements add many challenges for manufacturers and other developer participants. But from a technical standpoint, full anonymisation is really hard to achieve [7]. Even if data are anonymised, deanonymisation is still possible. According to [3], it was possible to identify individuals using different AV sensor data, and researchers could identify drivers with very high accuracy solely based on brake pedal usage. In [11], it is proposed to detect the thief of an auto-theft according to driving profiling, which proves driving habits differ from person to person, and they can be used for identifying individuals.

This mixed dataset of non-personal and personal data is not the only problem in the privacy aspect. Non-personal data, in principle, do not identify a specific person on their own, so, therefore, they fall outside of the GDPR's scope. But when they are combined with other similar non-personal data, the person becomes identifiable [18], blurring the line between personal and non-personal data. This combination of non-personal datasets creates the risk of revealing a specific individual [12]. In that sense, it is important to look for ways of storing these personal and non-personal data separately, and also, these non-personal data, which can be linked to a specific person when combined together, are never stored together in the same dataset. This approach reflects and aligns with "*privacy by design*", which is described under Article 25 of the GDPR.

In this paper, we propose an access control framework as a contributor towards "*privacy by design*" principle. The access control utilizes containment and sandboxing to ensure the privacy of data processed by AVs. Thus, the contributions of this work are:

- Explanation of the data types and flow in AVs
- Providing privacy implication of data processing in AV within the GDPR context
- Proposing a conceptual context-based layered access control framework that relies on containment and sandboxing according to privacy sensitivity of the data.

The paper is organized as follows. In Section 2, a brief discussion on privacy from legal framework and privacy-preserving technologies in the context of AVs is presented. The classification of the different types of data processed in AVs into four main categories is presented in Section 3 and the data flow and processing brief explanation follows in Section 4. The proposed conceptual context-based access control framework is presented in Section 5. Finally, the conclusions of the work are provided in Section 6 .

2. Privacy and GDPR in AV Context

GDPR is the fundamental legal regulation in the context of privacy in the EU. In GDPR, the core principles of data processing are set [6]. According to Article 5 of GDPR, these principles are: lawfulness, fairness, and transparency, purpose limitation, data minimisation, accuracy, storage limitation, integrity and confidentiality, and accountability. In this sense, according to GDPR, processing is an umbrella term for any operation that involves, among others, personal data, such as collection, recording, organisation, structuring and storage. While data subjects have the right to control their personal data, controllers and processors must comply with several obligations [2, p.17]. On the other hand, GDPR does not apply to truly anonymised data. But from a technical standpoint, full anonymisation is really hard to achieve [8].

According to GDPR Article 4(1), personal data is any information relating to an identified or identifiable natural person. The identifiability in this sense means the possibility of identification, which states for being identified in the future [16, p.164]. One way of identification is individuation, which makes a person distinctive in a group. The most basic form of individualization is targeting, which is also called t-identification (targeted identification). Targeting takes place in real-time and in a moment of contact. Even without a rich dataset, targeting can be achieved [17, p.252]. To achieve targeting, parameters are used, and more parameters mean narrowing down the circle of characterization

to get closer to the person [16, p.172] being identified. One of the aims of our layered context-based access control framework is to prevent this individualization in AVs.

According to the European Data Protection Board 01/2020 guideline on the processing of personal data in the context of connected vehicles and mobility related applications [5], it is suggested to implement a profile management system that enables users to control their privacy settings anytime. Furthermore, it is stated that companies should not collect all data just because they can or they might need it at some point, and according to the type of data, the consent of the user is needed. In addition, users need to understand the meaning of consent and have control over both their consent and the data that are processed.

The consent mechanism is complicated in AVs because most of the time, the subject of the consent is not a single fixed person. The passengers as well as the driver might change. Also, the context of the consent is hard to frame because there is so much diverse information that is processed via the Autonomous Driving System (ADS). Even if framing the consent is achieved, the data subject should understand what the consent is about and what the purpose is. These issues make the consent mechanism more difficult to implement in the context of AVs. Having control over the data that are processed is also hard to fully achieve. The user should be able to erase the data that are processed by the system, but from a technical perspective, full deletion of the data might not be possible in every scenario. For example, the data may have been stored in a cloud/edge device where the user has no permission to access the cloud/edge device.

In GDPR, pseudonymization and anonymisation methods are proposed as a way to prevent revelation of a specific person from data. While anonymised data fall outside the scope of GDPR, pseudonymized data can still be considered as personal data because they can still be used to match specific person with additional information [19, p.369].

Under GDPR Article 25, privacy by design and default is explained. Privacy by design is a principle that entails the proactive integration of privacy at the onset of system design. To ensure privacy by design and default, privacy-preserving technologies might help, but are not enough by themselves. In addition, these technologies should be up-to-date and evolve along with the technology landscape [10, p.118,119].

3. Data Classification in Autonomous Vehicles

Different types of data are processed by AVs through their various subsystems [21]. These data can be grouped in four main categories: sensor-generated data, user input data, telematics data, and connectivity data [14]. These data types are listed in detail in Fig. 1.

- *Sensor-generated data* are the data that are generated by physical in-vehicle/onboard sensors. While environmental sensors help the perception of the vehicle, internal sensors monitor vehicles motion and internal health. Other onboard sensors the current situation of the driver and passengers.
- *User input data* are the information that is directly fed by the user or the information that is generated by human-vehicle interaction.
- *Telematics data* are related to the reports of the monitoring vehicle's performance, system errors, or any related information related to the system's safety and the vehicle's health.
- *Connectivity data* are about the interaction with the external actors, such as infrastructure, other vehicles, and the network.

According to GDPR, data controllers (who determine the purposes and means of the processing of personal data) must ensure the implementation of privacy by design [1, p.260] in the data processing. If a data processing system is able to restrict the combination of data that may risk revealing of a specific person, then this system implemented privacy by design successfully. This is important because the mixed dataset is considered personal data as a whole, which means GDPR requirements should be applied [12]. These requirements add many challenges for manufacturers and system developers. This research proposes a context-based access control framework that enforces restriction of data combination that could risk privacy violation.

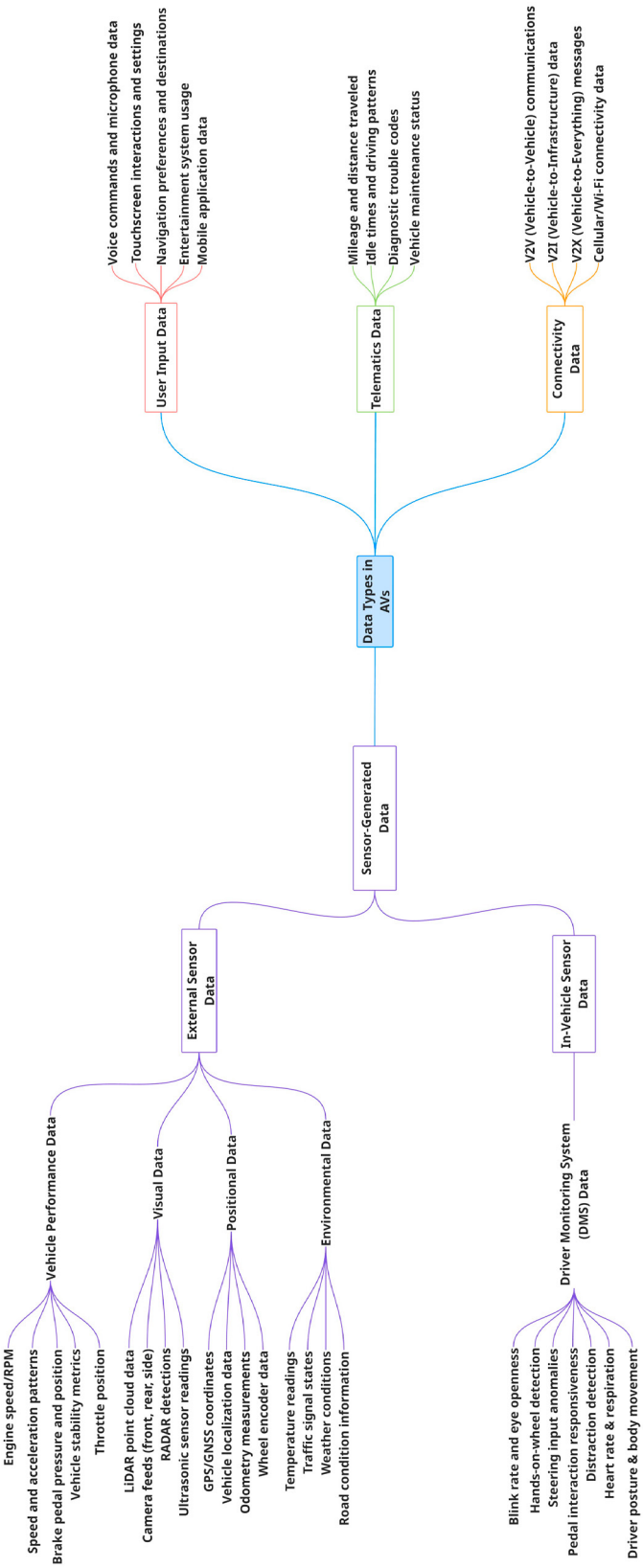

Fig. 1. Data Types in AVs

4. Data Flow and Processing in AVs

AVs need to collect diversified types of data to perceive their environment. This process is one of the core elements for their safe operation. The autonomous driving functionality of AVs is realized through an autonomous driving system (ADS), which consists of perception, localization, planning, control, and communication modules [2]. These modules need to exchange information. This information exchange is done through inter-module connections such as message buses, shared data structures, and network links. They carry the intermediate results of detected objects, lane-level maps, and predicted trajectories. Some of the data are communicated through wireless technologies [13]. This data sharing process increases concerns about data in the GDPR aspect. This is why it is important to embed data classification into the data flow process.

Each module has a responsibility and processes a specific kind of data. The raw data come into the ADS from the vehicle's hardware, which is called *external input*. It can be said that these raw sensor streams feed the car's brain. After collecting the raw data via sensors, this is enriched with metadata, which makes the raw stream identifiable and synchronized for downstream use. Those enriched data are sent to the *perception module*, which helps to detect objects, semantic segmentation, and localization. Then the *planning module* creates a plan for the vehicle on how to move with the behavior and trajectory planner. The *control module* helps to control the vehicle's longitudinal and lateral movement according to the data that come from the planning module. Then the *actuator interface* commands that the system sends out to the vehicle's actuators, which are used for converting the control signal into mechanical movement, so that the car actually moves. Perception, planning, and control modules share data bilaterally with the *communication module*, which enables communication with the environment of the vehicle, such as infrastructure, and other vehicles. Besides, all modules in AVs periodically transmit information with telemetry, which monitors the system and reports the functions to support safe driving. The data flow in AVs is explained in Fig. 2.

5. Context-based Access Control Through Sandboxing and Containment

Access control determines who can access what data, when, and for what purpose. All data access requests are processed through the access control. Different access control models are proposed in the literature, and for this study, the Attribute-Based Access Control (ABAC) is referred to, because of its flexibility. ABAC is not just about approving who is trying to get access to the data; it also checks the current situation and acts accordingly, which means it is dynamic, not static [20]. There are four core attributes that are defined in ABAC: subject (who requests the access?), object (what data are requested?), action (what is the aim?), and environment (what are the conditions?) [9]. Context-Based Control (CBC) emphasizes the role of the environment attribute, which checks the context. Furthermore, all these sets of rules that are used in the decision-making process of the access request, are often referred to collectively as access control policies.

The proposed access control framework uses sandboxing and containment methods according to the context-based policies. Sandboxing allows controlled, and monitored movement of data with specific boundaries through defined channels. The access to the data is explicitly controlled based on the context based access control policies, and the movement is tracked and limited. Containment prevents the data from moving or being combined, which represents passive isolation. In this method, data stay where they are; there is no transmission or access to data that should not be. While sandboxing is for controlled data sharing, containment is about preventing the data from being shared.

The risk of revealing a specific person in a non-personal data mix has a higher potential risk in historical data than in real-time data. Historical data contain relatively long-term observations compared to real-time data, which facilitates the possibility of extracting data that can be easily correlated and produce identifying parameters. For real-time data, fast data flow and decision-making are also important. An instant decision-making process is crucial for AVs, and if this process is interrupted, the safe and reliable operation of the vehicle would be endangered, which may cause crashes. This is why, for real-time data, containment would be a better option than sandboxing. The mix of non-personal data becomes risky for historical data, which are mostly used for analysis purposes, such as training the system for better operation. In this sense, sandboxing is a more suitable and safer option for historical data.

In this research, the proposed conceptual access control framework uses the containment and sandboxing in a layered manner as shown in Fig. 3. In layer 1, which is called "*pre-access control*", a containment method for the real-time data is proposed. With this layer, the data are protected where it originated. This layer has two sub-layers: 1A is

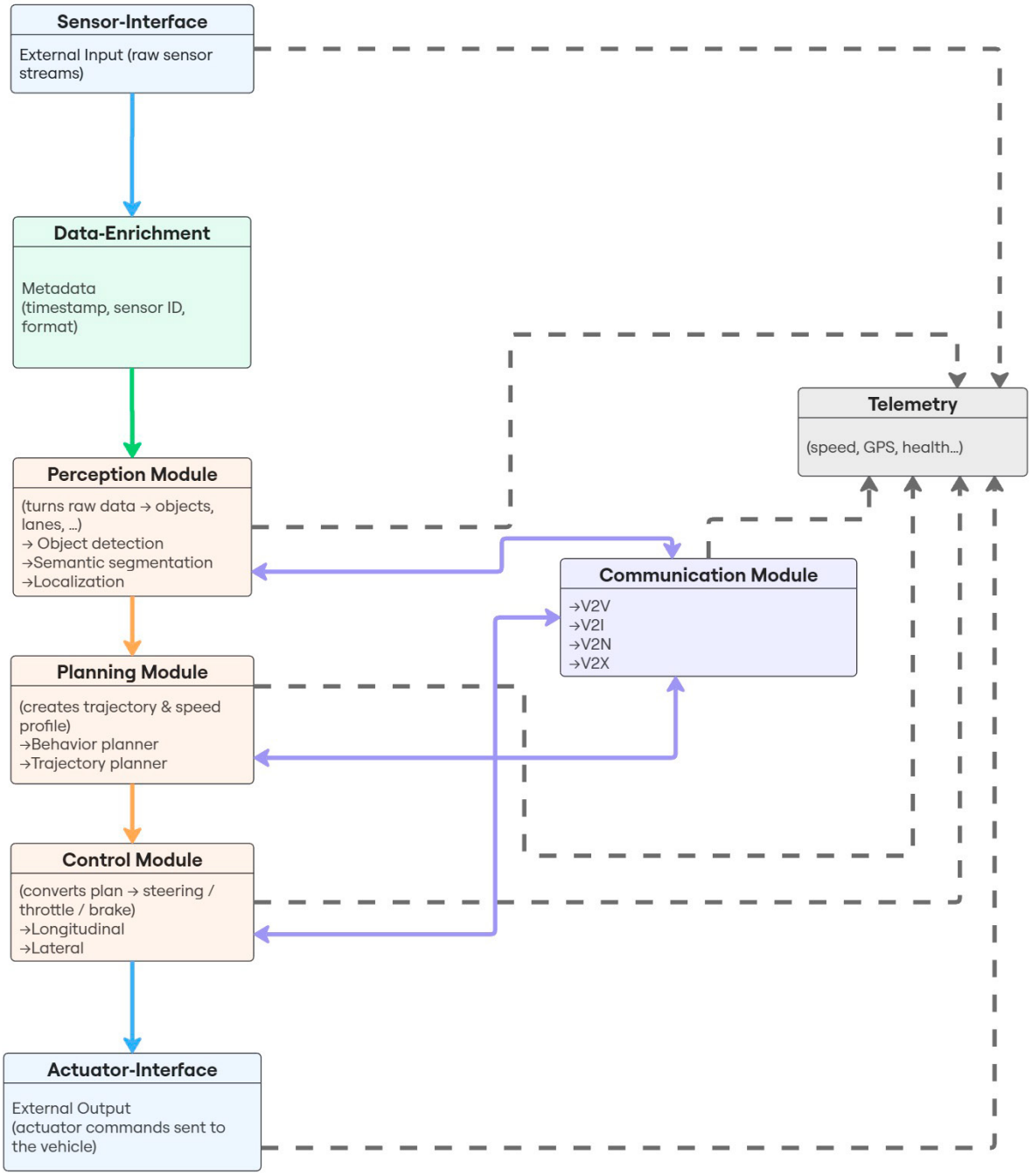

Fig. 2. Data Flow in AVs

"*modularization of processing*" and 1B is "*raw data to metadata boundary control*". As mentioned, ADS is composed of different modules that exchange data. In the sub-layer 1A, since all modules have different responsibilities, the data flow between them is limited to the purpose of the data, and the sharing process is only accepted when needed. Raw data are enriched with metadata, as explained, which makes the data meaningful. In the sub-layer 1B, the aim is to

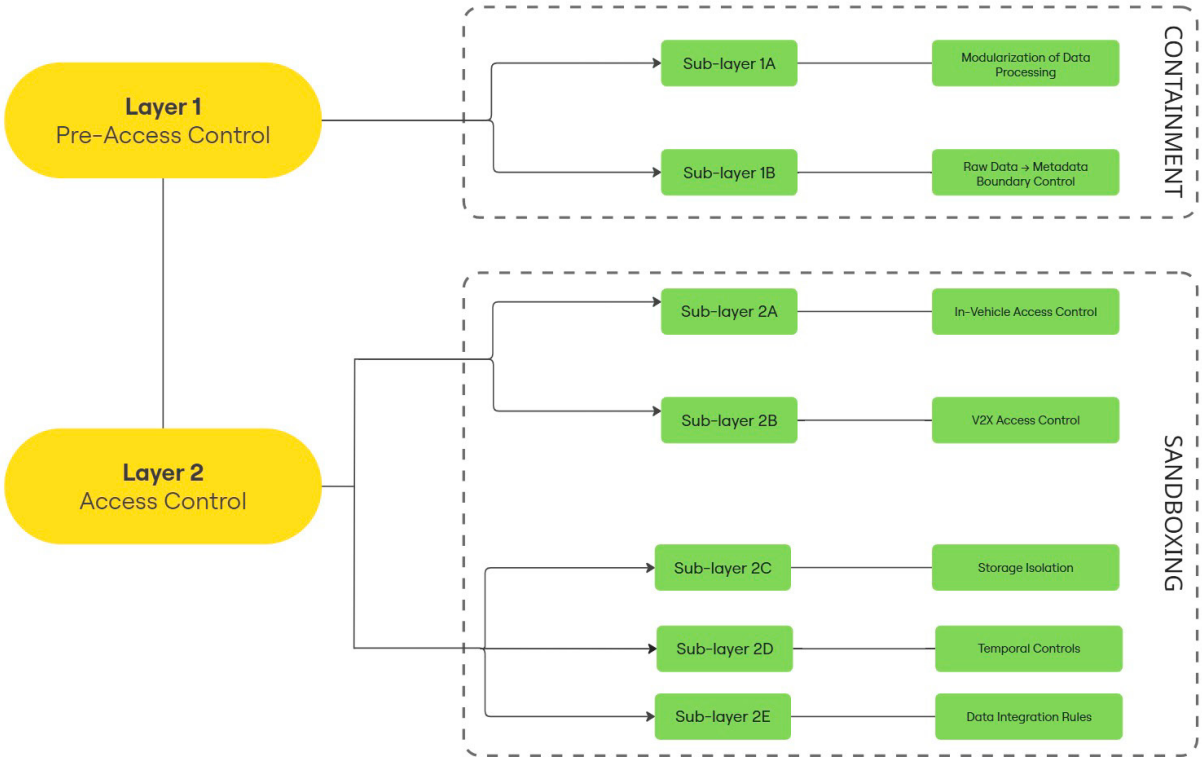

Fig. 3. Layered Access Control Framework

minimize the risk of leakage of this meaningful metadata. This sub-layer protects the raw data, which is very sensitive when it gets out of the vehicle, and also the metadata, which reflects meaningful patterns of the user.

In layer 2, which is called "*access control*", a sandboxing method for the historical data are proposed. With this layer, the aim is to protect both personal and non-personal data from revealing a specific person by separating these data types with sandboxing. This layer has two main groups; the first group of this layer is related to vehicle operation data, and the second group is related to storage and data integration. In the sub-layer 2A, which is called "*in-vehicle access control*": handles the sandboxing of the data generated by the vehicle sensors, its driver, passengers and apps as well as enforcing of access control according to the context of each data. In the sub-layer 2B, which is called "*V2X access control*": handles the data that comes from the external actors such as other vehicles, and infrastructure. These two sub-layers create the first group of the sub-layer of layer 2. In the sub-layer 2C, which is called "*storage isolation*": process the storage of the historical data into a separate sandbox to prevent the combination of the personal and non-personal data, as well as historical non-personal data that could lead to re-identification. In the sub-layer 2D, which is called "*temporal controls*": this sub-layer defines the rules for the time period of the data to be stored, and also enforces who will get access to what data and when based on the defined access control policies. In sub-layer 2E, which is called "*data integration rules*", defines and continuously updates the set of rules on which data can or cannot be stored together. The sub-layers 2C, 2D, and 2E are the second group of layer 2.

6. Conclusion

In this paper, we presented a conceptual context-based access control framework for AVs. The current practice of data collection, processing and communication in AVs fail to adhere to the "*privacy by design*" requirement of GDPR in the context of personal data. The proposed conceptual access control framework contributes towards achieving

"*privacy by design*" requirement of GDPR. The access control has two layers with sensitivity classification of real-time and historical data. It applies containment and sandboxing depending on the privacy sensitivity of the data and according to the context-based policies. It contributes to privacy by design in AVs by restricting the combination of different data types that could potentially risk the revelation of personally identifying data. The presented conceptual access control framework will be implemented and evaluated in the future.

References

[1] Aulino, L., Saager, M., Harre, M.C., Espindola, L., 2020. Consideration of Privacy Aspects in the Area of Highly Automated Driving. An Intention Recognition Use Case Section I: Articles. European Journal of Privacy Law & Technologies (EJPLT) 2020, 252–264. URL: https://heinonline.org/HOL/P?h=hein.journals/ejplt2020&i=628.

[2] Daoudagh, S., Marchetti, E., Monreale, A., . The GDPR Compliance Through Access Control Systems. Doctoral Thesis. Universita di Pisa.

[3] Enev, M., Takakuwa, A., Koscher, K., Kohno, T., 2016. Automobile Driver Fingerprinting. Proceedings on Privacy Enhancing Technologies URL: https://petsymposium.org/popets/2016/popets-2015-0029.php.

[4] European Commission, . Communication from the commission to the european parliament and the council. URL: https://primarysources.brillonline.com/browse/human-rights-documents-online/communication-from-the-commission-to-the-european-parliament-and-the-council;hrdhrd46790058. series Title: Human Rights Documents online.

[5] European Data Protection Board, . Guidelines 01/2020 on processing personal data in the context of connected vehicles and mobility related applications | European Data Protection Board URL: https://www.edpb.europa.eu/our-work-tools/our-documents/guidelines/guidelines-012020-processing-personal-data-context_en.

[6] European Union, . Regulation (EU) 2016/ 679 of 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, and repealing Directive 95/ 46/ EC (General Data Protection Regulation) .

[7] Finck, M., Pallas, F., 2020. They who must not be identified—distinguishing personal from non-personal data under the gdpr. International Data Privacy Law 10, 11–36. URL: https://academic.oup.com/idpl/article/10/1/11/5802594.

[8] Gadotti, A., Rocher, L., Houssiau, F., Creţu, A.M., de Montjoye, Y.A., . Anonymization: The imperfect science of using data while preserving privacy. Science Advances 10, eadn7053. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC466941/.

[9] Hu, V.C., Ferraiolo, D., Kuhn, R., Schnitzer, A., Sandlin, K., Miller, R., Scarfone, K., 2014. Guide to Attribute Based Access Control (ABAC) Definition and Considerations. Technical Report. National Institute of Standards and Technology. Gaithersburg, MD. URL: https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-162.pdf.

[10] Kuner, C., Bygrave, L.A., Docksey, C., Drechsler, L., Tosoni, L., 2021. The EU General Data Protection Regulation: A Commentary/Update of Selected Articles. SSRN Electronic Journal URL: https://www.ssrn.com/abstract=3839645. publisher: Elsevier BV.

[11] Kwak, B.I., Woo, J., Kim, H.K., 2016. Know your master: Driver profiling-based anti-theft method, in: 2016 14th Annual Conference on Privacy, Security and Trust (PST), pp. 211–218. URL: https://ieeexplore.ieee.org/document/7906929.

[12] Lokaj, Z., Šrotýř, M., Vaniš, M., Janda, I., Ščerba, T., 2023. Detection of Non-Personal Data Leading to User Identification, Including Related Recommendations for the Field of Autonomous Mobility. The Lawyer Quarterly 13. URL: https://old.tlq.ilaw.cas.cz/index.php/tlq/article/view/571. number: 4.

[13] Mendes, B., Araújo, M., Goes, A., Corujo, D., Oliveira, A.S., 2025. Exploring V2X in 5G networks: A comprehensive survey of location-based services in hybrid scenarios. Vehicular Communications 52, 100878. URL: https://linkinghub.elsevier.com/retrieve/pii/S2214209625000051. publisher: Elsevier BV.

[14] Neupane, S.R., Sun, W., 2025. Advanced Data Classification Framework for Enhancing Cyber Security in Autonomous Vehicles. Automation 6, 5. URL: https://www.mdpi.com/2673-4052/6/1/5. number: 1 Publisher: Multidisciplinary Digital Publishing Institute.

[15] Next Move Strategy Consulting, . Global autonomous vehicle market size 2030. URL: https://www.statista.com/statistics/1224515/av-market-size-worldwide-forecast/.

[16] Purtova, N., 2022. From knowing by name to targeting: the meaning of identification under the GDPR. International Data Privacy Law 12, 163–183. URL: https://academic.oup.com/idpl/article/12/3/163/6612144. publisher: Oxford University Press (OUP).

[17] Purtova, N., Leenes, R., 2023. Code as personal data: implications for data protection law and regulation of algorithms. International Data Privacy Law 13, 245–266. URL: https://doi.org/10.1093/idpl/ipad019.

[18] Rossi, L., Walker, J., Musolesi, M., 2015. Spatio-temporal techniques for user identification by means of GPS mobility data. EPJ Data Science 4, 11. URL: https://doi.org/10.1140/epjds/s13688-015-0049-x.

[19] Salami, E., 2022. Balancing Competing Interests in the Reidentification of AI-Generated Data. European Data Protection Law Review (EDPL) 8, 362–376. URL: https://heinonline.org/HOL/P?h=hein.journals/edpl8&i=379.

[20] Zhang, Q., Zhong, H., Cui, J., Ren, L., Shi, W., 2021. AC4AV: A Flexible and Dynamic Access Control Framework for Connected and Autonomous Vehicles. IEEE Internet of Things Journal 8, 1946–1958. URL: https://ieeexplore.ieee.org/abstract/document/9169695.

[21] Znidi, F., Morsy, M., Rathore, H., 2024. Future on Wheels: Safeguarding Privacy in Tomorrow's Connected Vehicles-FUTURE-SP. IEEE Access 12, 179857–179878. URL: https://ieeexplore.ieee.org/document/10766571.

[22] Zuboff, S., . Surveillance Capitalism or Democracy? The Death Match of Institutional Orders and the Politics of Knowledge in Our Information Civilization 3, 20. URL: https://journals.sagepub.com/doi/epub/10.1177/26317877221129290.