DISCUSSION

DATA PROTECTION AND PRIVACY ISSUES IN THE USE OF AUTONOMOUS VEHICLES

Eva Fialová*, Ján Matejka**[1]

Abstract: *Although autonomous vehicles will have many benefits, not only in the field of road safety but also in personal autonomy for those who cannot drive non-autonomous vehicles and are thus dependent on the help of others (the elderly and handicapped), the collection of data in real-time may hurt the privacy of persons whose personal data is processed.*

This paper deals with the data collected and processed in the operation of autonomous vehicles, their further use, and some of the problematic issues related to this process. This paper offers a solution consisting of the adoption of special legislation on processing personal and non-personal data because the distinction mentioned above will blur and be fluid in an enormous amount of data. This approach would be analogous to electronic communications data, irrespective of their personal and non-personal form. The paper does not seek to assess how to process personal data in the operation of autonomous vehicles under the GDPR but instead points out that the GDPR does not provide a suitable legal framework.

Keywords: *Autonomous vehicles, GDPR, non-personal data, personal data, privacy*

1. INTRODUCTION

Autonomous vehicles will need data to operate. They will receive data from the outside; from other vehicles (V2V), the infrastructure (V2I), or from third parties (data providers). Autonomous vehicles will also generate the data and send it back to other vehicles, the infrastructure, car manufacturers, public authorities, and other public or private entities.[2] In the operation of autonomous vehicles, an incredibly large amount of data will be processed.[3]

Autonomous vehicles will process:

1) Data generated by sensors, cameras, radars and lidars (devices for measuring distances by laser light), and thermal imagining devices. These components will capture information about the vehicle's operation and surroundings.

* JUDr. Eva Fialová, LL.M. Ph.D., The Institute of State and Law of the Czech Academy of Sciences, Prague, Czech Republic. ORCID: 0000-0001-7904-9250.

** JUDr. Ján Matejka, Ph.D., The Institute of State and Law of the Czech Academy of Sciences, Prague, Czech Republic. ORCID: 0000-0003-0051-2754.

[1] This paper was supported by the Technology Agency of the Czech Republic under grant No. TL02000085 Civil Liability for Damage Caused by Operation of Autonomous Vehicles (CiLiAV).

[2] National Automobile Dealers Association and the Future of Privacy Forum. Personal Data In Your Car. In: *Future of Privacy Forum* [online]. [2021-03-04]. Available at: <https://fpf.org/wp-content/uploads/2017/01/consumer-guide.pdf>; or Data And The Connected Car. In: *Future of Privacy Forum* [online]. [2021-03-04]. Available at: <https://fpf.org/wp-content/uploads/2017/06/2017_0627-FPF-Connected-Car-Infographic-Version-1.0.pdf>.

[3] It is estimated that sensors of one vehicle alone can produce between 380 TB to 5 100 TB of data in one year. WRIGHT, S. Autonomous cars generate more than 300 TB of data per year. In: *Tuxera* [online] 2.7.2021 [2022-05-22]. Available at: <https://www.tuxera.com/blog/autonomous-cars-300-tb-of-data-per-year/>

2) Location data using GPS and GIS (Geographic Information System) that will locate and navigate the vehicle.[4]
3) Data about an owner, passengers, and a driver of a semiautonomous vehicle, their settings and preferences, biometric data (facial or voice recognition or fingerprints to identify a user of the vehicle, to open and start the vehicle, or to perform specific tasks by a voice command[5]) and current driver's condition (driver alert system).

Most of the data will not be *a priori* personal data. It will be the data necessary for the actual operation of the vehicle, or rather for driving it. However, even this data in connection with data about individuals will qualify as personal data. They will need to be handled by the legal principles of personal data protection. For example, there will be the location data of the vehicle, which is not personal data itself, but, in connection with data about the vehicle user, or an address where the vehicle is normally parked overnight, the location data become personal data as well.[6] Other data will be personal by its very nature (biometric data, personal settings, etc.). It is likely that users will not be aware of the processing of information about themselves or will not be aware of the volume of the data and linkability with other personal information from smartphones or smart home apps.

2. AUTONOMOUS, AUTOMATED AND CONNECTED VEHICLES

There are different levels of autonomy or automatization. The most used scale of automatization/autonomy level is NHTSA/SAE scale of automation.[7] Fully autonomous vehicles are expected to be only five vehicles. In level 3, the driver present in the vehicle could take over the driving in some circumstances. Level 4 vehicles can operate in a particular area or on predefined roads. The article addresses vehicles mainly from levels 3 to 5. When we talk about autonomous vehicles, we will also understand an automated level 3 vehicle.

Besides autonomous and automated vehicles, we also recognise the connected vehicles. Connected vehicles *"can be defined as a vehicle equipped with many electronic control units (ECU) linked together via an in-vehicle network and connectivity facilities allowing it to share information with other devices both inside and outside the vehicle."*[8] Connected vehicles will receive data from the outside, from other vehicles (V2V), the infrastructure (V2I), or third parties. Connected vehicles will also generate the data and send

[4] Anatomy of Autonomous Vehicles: Is GIS Really Under the Hood of Self-Driving Cars? In: *GIS Geography* [online]. 22.1.2022 [2022-05-22]. Available at: <https://gisgeography.com/autonomous-vehicles-gis-self-driving-cars/>.

[5] Commission Nationale de l'Informatique et des Libertés (CNIL). Connected vehicles and personal data. In: *CNIL* [online]. [2021-02-22], p. 21. Available at: <https://www.cnil.fr/sites/default/files/atoms/files/cnil_pack_vehicules_connectes_gb.pdf>.

[6] FOWLER, G. A. What does your car know about you? We hacked a Chevy to find out. *The Washington Post.* [online] 17.12.2019 [2022-05-22]. Available at: <https://www.washingtonpost.com/technology/2019/12/17/what-does-your-car-know-about-you-we-hacked-chevy-find-out/>.

[7] Society of Automobile Engineers International (SAE International). J3016 Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles. p. 1. In: *SAE* [online] [2022-05-22]. p. 1. Available at: <https://www.sae.org/standards/content/j3016_201806>.

[8] Guidelines 01/2020 on processing personal data in the context of connected vehicles and mobility-related applications. In: *European Data Protection Board* [online]. [2022-05-22]. Available at: <https://edpb.europa.eu/system/files/2021-03/edpb_guidelines_202001_connected_vehicles_v2.0_adopted_en.pdf>, p. 8.

it back to other vehicles, the infrastructure, or third parties. They will receive data from the outside, from other vehicles (V2V), the infrastructure (V2I), or from third parties. Autonomous vehicles will generate the data and send it back to other vehicles, the infrastructure, or third parties. All autonomous vehicles will be connected vehicles, but not all connected vehicles will be autonomous.

Next to autonomous (and connected) vehicles are being developed that are autonomous in the true sense of the word. These vehicles use only the data they generate or record before their operation. The vehicles themselves do not communicate with the outside world.[9] Although these vehicles can be used for military purposes, for example, future civilian autonomous vehicles will most likely be based on interacting with the outside world.[10]

3. DATA PROCESSING AND PURPOSES OF PROCESSING

As mentioned above, autonomous vehicles will communicate with their surroundings. The French supervisory authority CNIL has identified three basic models for transmitting personal data, depending on their use.[11] The first model is personal data that should not be transmitted outside the vehicle (*IN-IN*). These data are, for example, the data necessary to authenticate the user or adjust the vehicle's interior according to the user's needs. The second model includes data transmitted outside the vehicle (*IN-OUT*), e.g., for the provision of commercial services by third parties, geolocation data in case of theft, or data sent via eCall. The last transmission model is a model where data is transmitted outside the vehicle but returned to it (*IN-OUT-IN*). These personal data include the data necessary for vehicle maintenance or remote installation of updates or data required for information from an early warning system.

Specific communication via short-range microwave technology will occur between autonomous vehicles and between them and the infrastructure. The interconnection between various entities in the so-called C-ITS (collaborative intelligence transportation system) enables autonomous traffic to operate. These systems are based on an open network, allowing the interconnection between the stations of this system. Personal data will be processed within it.[12]

It is already expected that (autonomous) vehicles will be equipped with a whole range of different devices processing data. One of these devices is an event data recorder (hereinafter: "EDR"). This equipment will be mandatory in all vehicles after 2024. Regulation 2019/2144 on type-approval requirements for motor vehicles and their trailers, and systems, components, and separate technical units intended for such vehicles, as regards

[9] GLANCY, D. J. Privacy in Autonomous Vehicles. *Santa Clara Law Review*. 2012, Vol. 52, No. 4, p. 1174.

[10] FRIEDRICH, B. Verkehrliche Wirkung autonomer Fahrzeuge. In: M. Maurer et al. *Autonomes Fahren. Technische, rechtliche und gesellschaftliche Aspekte*. Berlin, Hedelberg: Springer Verlag. 2015, p. 349.

[11] COMMISSION NATIONALE DE L'INFORMATIQUE ET DES LIBERTÉS (CNIL). *Connected vehicles and personal data*. p. 19.

[12] ARTICLE 29 DATA PROTECTION WORKING PARTY. Opinion 03/2017 on Processing personal data in the context of Cooperative Intelligent Transport Systems (C-ITS) (WP 252). In: *Europa.eu* [online] [2021-11-04]. Available at: < https://ec.europa.eu/newsroom/article29/items/610171>, p. 4.

their general safety and the protection of vehicle occupants and vulnerable road users (General Safety Regulation, hereinafter: "GSR") already anticipates the autonomous mobility and foresees a uniform format for data exchange, as well as systems for recording data on the vehicle's position and its surroundings in real-time.

EDP won't be the only mandatory vehicle equipment under GSR. The autonomous vehicles will have to be also equipped with i.e. driver availability monitoring system, driver drowsiness, and attention warning, advanced driver distraction warning, or alcohol interlock installation facilitation. Those devices will necessarily process the personal data of the driver of the level 3 vehicle.

The data from or about the autonomous vehicle and its operation will be or may be processed for various purposes and, for those purposes, transferred to third parties. The purposes for data processing will encompass

1) vehicle('s) operation and safety
2) infrastructure condition and traffic
3) research and development
4) investigation of accidents
5) user's settings and preferences
6) provision of commercial services and infotainment to the passengers
7) other purposes defined by public authorities (tax collection, analysis of crimes, etc.)

The data about the occurrence of an accident will have to be handed over to insurance companies to process and resolve the damage event; in other words, to find out what was the cause of the damage and process claims, and eventually ascertain whose fault gave rise to liability. Currently, the member states would have to determine what data will be used to establish liability.[13] Insurance companies will also need the data to calculate risk and set up their business models.[14]

Already, with the increasing use of assistance systems and, therefore the volume of data, the European Automobile Manufacturers Association (ACEA) warns that numerous companies are increasingly interested in accessing vehicle data for their subsequent commercial use: providers of emergency services, insurance companies for determining the price of insurance premiums, providers of financial services, operators of road infrastructure and tolls, providers of entertainment and travel services, operators of social networks and search engines, marketers, etc.[15]

The massive potential of commercial use is represented by (personal) data related to the movement of persons, i.e., who, where to, where from, when, and how. This data can be used to estimate the user's income using the vehicle type she uses and where she usually works, live, or does the shopping. It could also be possible to determine which users travel together in the vehicle. This data could be used to create user profiles and show targeted

[13] EVAS, T. A common EU approach to liability rules and insurance for connected and autonomous vehicles. In: *European Parliament* [online] 2018 [2022-05-22]. p. 27. Available at: <https://www.europarl.europa.eu/RegData/etudes/STUD/2018/615635/EPRS_STU(2018)615635_EN.pdf>.

[14] KRAUSOVÁ, A., MATEJKA, J. Autonomous Vehicles And In-Vehicle Data In The Context Of Motor Insurance. *The Lawyer Quarterly*, 2/2020, p. 159.

[15] ACEA. Strategy Paper on Connectivity. In: *ACEA* [online]. [2021-11-04]. p. 3. Available at: <https://www.acea.be/uploads/publications/ACEA_Strategy_Paper_on_Connectivity.pdf>.

commercials directly in the vehicle. According to Glancy, offers for goods and services will be delivered to users in real-time, based on the location of the vehicle.[16] Currently, there are rules for employers on the conditions under which and to what extent they can monitor the location of their vehicles, and hence employees, via the GPS. At the time of autonomous mobility, anyone using an autonomous vehicle could be monitored this way.

4. LEGAL REGULATION OF DATA PROCESSING

Controllers of personal data are obliged to comply with data protection law. The processing of personal data in the European Union is generally regulated by Regulation No. 2016/679 on the protection of individuals concerning the processing of personal data and on the free movement of such data and repealing Directive 95/46/EC (General Data Protection Regulation, hereinafter: "GDPR").

However, GDPR is too general. The controllers will have to comply with the data minimisation principle, process the data for a period necessary for the purpose, adopt appropriate safety measures, etc. GDPR leaves to the data controllers how to fulfill their obligations in this legal norm. Besides that, GDPR lays down rules just for the processing of personal data.

The amount of data will be enormous. Together, they may compile a detailed picture of the car users. Autonomous driving will also be characterized by the plurality of users (passengers) of the autonomous vehicle during one drive. Detailed information about the persons involved in the operation of the autonomous vehicle will be available to public authorities and private entities for various purposes mentioned above. The quantity of data could serve to the authorities to possibly monitor the activities of these persons or to use the information in any other way. For the private entities, the data will represent a precious source of information as well.

Commercial use of personal data could be, in principle covered by GDPR. Commercial use of personal data may only be possible, with rare exceptions, with the consent of data subjects. However, some services could be provided if all passengers in the vehicle have given their consent (e.g., in-vehicle advertising). The question is whether the controller can process personal data at all if not all the passengers give their consent. To process personal data based on consent, if there are several passengers, a form of collective approval would be practical, whereby one data subject, usually the one determining the driving parameters, would grant consent for other vehicle passengers' data to be processed. This type of consent is not possible *de lege lata* because today's personal data protection and privacy concept is based on the individual's control of the data.[17]

The question is not only from whom to request the consent but also how to manage to change data subjects as the vehicle occupancy changes, for example, if the fleet owner and the user operate the vehicle, or the users of the autonomous vehicle will log in before they can drive it or use it. Each passenger could give the application provider their own

[16] GLANCY, D. J. *Privacy in Autonomous Vehicles*. p. 1195.

[17] KAMMOURIEH, L. et al. Group Privacy in the Age of Big Data. In: L. Taylor – L. Floridi – B. Van Der Sloot. *Group Privacy. New Challenges of Data Technologies*. Cham: Springer, 2017, p. 55.

informed consent to process their personal data. Those intending to use the vehicle will be overwhelmed with requests for approval. In that case, it can be assumed that consent will be given without reading information about processing personal data, as is the case with services provided via the Internet. By the principle of data protection by design and valid consent, vehicle users should also be able to opt in and not opt out in real-time.

The performance of a contract as the ground for personal processing data would only apply to the personal data of the vehicle user who concluded the agreement, for example, with the fleet operator. This concept will also be more beneficial for the other party to the agreement, as the sole person responsible for the performance of the vehicle rental contract will be the data subject. However, the other contracting party could process the personal data of other vehicle users.

5. RECOMMENDATION FOR FUTURE REGULATION

The theoretical solution for some types of data is anonymization. However, anonymization is not the panacea. Due to current advanced technologies and several datasets, there will always be a residual risk of identification.[18] In fact, this is what the legislator also cautiously admits. For instance, according to art. 6 para 4 of GSR, the data collected by event data recorders in cars must be anonymized, but, at the same time, the data can be made available to national authorities on the basis of EU or national law and in compliance with GDPR (sic!). Anonymization would not even be desirable if the cause of an accident had to be investigated.

The operation of autonomous vehicles will necessarily be comprised of numerous personal data processes and multiple data controllers and processors. It will not always be clear who is the controller and the processor of what data and which data subjects should address for executing their rights.[19] The analysis for determining the responsible persons made by Vellinga and Mulder, but the role of the particular person in the data processing in their study depends on whether this person fulfills the definition of the controller, resp. The processor.[20]

Working Party 29 (the predecessor of the European Data Protection Board) anticipated the adoption of a particular law regulating the processing of data in the C-ITS system, in particular, because the processing of personal data in these systems is complex, with an unlimited number of entities that could participate in the system.[21]

Processing such an amount of data may interfere with the individual's privacy. GDPR represents the harmonization of data processing but cannot by itself fulfill the requirements of transparency and predictability towards the data subjects. Processing of personal data by public authorities, i.e., when it results in the threat of privacy infringement, should be governed by special legislation that would meet those requirements.[22]

[18] FINCK, M., PALLAS, F. They who must not be identified—distinguishing personal from non-personal data under the GDPR. *International Data Privacy Law*. 2020, Vol. 10, No. 1, p. 35.

[19] EVAS, T. *A common EU approach to liability rules and insurance for connected and autonomous vehicles*, p. 26.

[20] MULDER, T., VELLINGA, N. Exploring data protection challenges of automated driving. *Computer Law & Security Review*. 2021, Vol. 40, p. 5.

[21] Ibid., p. 9.

[22] Decision of the Slovak Constitutional Court from 10. November 2019, no PL. ÚS 25/2019-117.

Moreover, the line between personal and non-personal data blurs with the amount of processed data. Data about objects may become personal if the use, ownership, or other relation with an individual can indirectly identify her. Personal data may be easily derived from information about the vehicle.[23] The unclear distinction between the personal and non-personal data and fluidly passing of the particular data from one category to another cause legal uncertainty as well as the tendency of responsible persons to escape the legal obligations arising from the processing of the data by labeling all the data as non-personal, even though it may not be the case in the course of the data processing.[24] Anonymization, as well as the non-personal character of data, thus contexts and time-dependent.[25]

The best solution would be the special legislation regarding personal processing data and non-personal data relating to the operation of autonomous vehicles. The special legal norm or norms may overcome the main problematic points, that is, 1) the nebulous line between personal and non-personal data and their presumed frequent passage from non-personal to personal data, 2) the use of the vehicle by more users individually and at the same time. The law could apply to both commercial and non-commercial use of the data processed in the operation of autonomous vehicles.

This legislation, whether in one single norm or more norms, would have to cover the categories of processed data and the purposes of the processing, the retention period (general and in case of an accident), security measures, and the types of recipients. The subsequent data processing of some recipients (public authorities, insurance companies) and their entitlement should be reflected in special legal regulations.

The legislator could get inspired by the processing of data in electronic communication, mainly that its rules apply to the processing irrespective of whether or not the data are personal or non-personal. The autonomous vehicle could be considered a terminal device analogous to the terminal equipment under the ePrivacy Directive.[26]

As the proposal or ePrivacy Regulation (Recital 2) says: "*The content of electronic communications may reveal highly sensitive information about the natural persons involved in the communication, from personal experiences and emotions to medical conditions, sexual preferences and political views, the disclosure of which could result in personal and social harm, economic loss or embarrassment. Similarly, metadata derived from electronic communications may also reveal very sensitive and personal information. These metadata includes the numbers called, the websites visited, geographical location, the time, date, and duration when an individual made a call, etc., allowing precise conclusions to be drawn regarding the private lives of the persons involved in the electronic communication, such*

[23] ARTICLE 29 DATA PROTECTION WORKING PARTY. Opinion 4/2007 on the concept of personal data (WP 136), In: *Europa.eu* [online]. [2021-11-04]. Available at: <https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2007/wp136_en.pdf>, p. 9.

[24] CHERCIU, N., CHIRVASE, T. Non-Personal Data Processing Why Should We Také It Personally? *European Journal of Privacy Law & Technologies*. 2020, No. 2, p. 188.

[25] STALLA-BOURDILLON, S., KNIGHT, A. Anonymous Data v. Personal Data False Debate: An EU Perspective on Anonymization, Pseudonymization and Personal Data. *Wisconsin International Law Journal*. 2016, Vol. 34, No. 2, p. 318.

[26] The vehicles connected are terminal equipment regarded by the European Data Protection Board. See *Guidelines 01/2020 on processing personal data in the context of connected vehicles and mobility-related applications*. p. 7.

as their social relationships, their habits, and activities of everyday life, their interests, tastes, etc." The data concerning electronic communication need not be per se personal data, but in connection with other data, this data may acquire the character of personal data. This statement applies perfectly to the data (personal and non-personal) processed in the operation of autonomous vehicles.

6. CONCLUSION

GDPR enshrines the basic principles but itself does not represent clear and definite rules for processing. To ensure the protection of privacy and personal data, but simultaneously to allow and unify the processing of personal and non-personal, the authors argue that the processing should have its ground in a particular legal regulation covering the processing of personal as well as non-personal data. The reason for this solution is an immense amount of data and the blur and the fluid distinction between the two categories of data. The authors argue for the regulation analogous to electronic communications data.

Overall, the authors of this text believe that processing personal and non-personal data in the operation of autonomous vehicles must consist of strict adherence to principles implemented in four key areas. Awareness of these principles represents the *conditio sine qua non* of legal and safe implementation of the operation of autonomous vehicles.

The first such area is undoubtedly the area of the quality of national or EU legislation. The legislatures should pass special laws to protect autonomous vehicle data in this area. Generally, under these laws, data should be downloaded only with the vehicle owner's consent, subject to certain exceptions (e.g., court orders, vehicle safety research, or to service or repair the vehicle). The second (and related area) is anonymizing personal data obtained by autonomous vehicles, which should be anonymized. Still, steps should be taken to ensure that data reasonably cannot be re-identified, considering technological developments and regulatory guidance. For example, a unique mobile device identifier or IP address is "individually-identifiable information," a unique identifier associated with a vehicle could also be considered personally identifiable, even if an owner or passenger could not always be positively identified. Under a "privacy by design" approach, companies should be encouraged to "build security into their devices at the outset, rather than as an afterthought and to consider:

- conducting a comprehensive privacy or security risk assessment;
- minimizing the data they collect and retain, and
- reasonably testing all security measures before related products or services are marketed.

In sectoral directives or industry guidance, the stakeholders involved in developing autonomous vehicles may also take advantage of other industry resources. For example, the alliance of automobile manufacturers and related associations should have developed voluntary consumer privacy protection principles that provide an approach to customer privacy in developing automobiles and features utilizing innovative technologies.

The last significant area is transparency and adherence to the data protection principles, including the notice and consent principles, which are vital under the fair information practice principles that serve as the foundation of many privacy laws and frameworks.