LEGAL PROTECTION FOR DATABASES IN THE REALM OF AUTONOMOUS MOBILITY: A COMPREHENSIVE APPROACH

Zdeněk Lokaj,* Martin Šrotýř,** Miroslav Vaniš,*** Ivo Janda,**** Tomáš Ščerba*****

Abstract: *In the dynamic realm of autonomous mobility, understanding legal protection for databases is crucial. The upcoming article focuses on key aspects in this area. It clarifies fundamental concepts, including both sui generis and copyright safeguards. Additionally, it outlines important characteristics for assessing potential legal protection. The article also introduces a unique set of coefficients, making complex legal concepts more tangible. This innovative approach offers makers concrete insights into their database's level of legal protection. By combining theoretical understanding with practical application, the article equips makers in the autonomous mobility sector to implement effective strategies for securing legal protection. Furthermore, the article will delve into simplified typologies of databases within autonomous systems. It will also provide a step-by-step illustration of how to determine the degree of legal protection for these databases. This process will include supplementary information aimed at refining the assessment of legal protection, offering makers the necessary guidance for accurate decisions. Through this comprehensive approach, the article aims to provide makers with practical tools to navigate the complex landscape of database legal protection efficiently.*

Keywords: *Autonomous mobility, autonomous vehicles, law protection, traffic-data databases, sui generis, copyright safeguard*

I. INTRODUCTION[1]

The advent of autonomous vehicles onto roads marks the dawn of a transportation revolution, promising not only enhanced efficiency and safety but also a transformation in the way we perceive movement. Yet, this new paradigm comes with legal and ethical questions that demand meticulous examination and resolution. Among these inquiries lies the complex issue of protecting databases that harbor crucial data for the functioning and advancement of autonomous mobility.

As data utilization surges within the domain of autonomous mobility, gaining a comprehensive understanding of database legal protection becomes pivotal. This comprehension enables database creators to strategically plan protective measures and seamlessly integrate data into their strategies. Moreover, the dynamic nature of autonomous mobility necessitates innovative and adaptable approaches to database legal protection, ones that can reflect the specific challenges and characteristics of this field.

* Associate Professor, Ing. Zdeněk Lokaj, Ph.D., LL.M. Associated Professor at the Department of Applied Informatics in Transportation, Faculty of Transportation Sciences, Czech Technical University in Prague, Prague, Czech Republic. ORCID: 0000-0002-0624-0430.

** Ing. Martin Šrotýř, Ph.D. Department of Applied Informatics in Transportation, Faculty of Transportation Sciences, Czech Technical University in Prague, Prague, Czech Republic. ORCID: 0000-0003-0049-4381.

*** Ing. Miroslav Vaniš, Ph.D. Department of Transport Telematics, Faculty of Transportation Sciences, Czech Technical University in Prague, Prague, Czech Republic. ORCID: 0000-0002-7589-6206.

**** JUDr. Ivo Janda, Ph.D. Partner at White & Case LLP, Prague, Czech Republic.

***** JUDr. Tomáš Ščerba, Ph.D. Former Local Partner White & Case LLP, Prague, Czech Republic. ORCID: 0009-0002-9898-6594.

[1] This article is part of the project "Law protection of traffic-data databases in autonomous driving systems", which is co-financed from the state budget by the Technology Agency of the Czech Republic (www.tacr.cz) under the ÉTA 5 Programme.

The objective of this article is to dissect the key facets of legal protection for databases in the context of autonomous mobility. It concentrates on elucidating fundamental concepts surrounding database protection, encompassing sui generis and copyright protection. Furthermore, it introduces a novel approach utilizing specialized coefficients to precisely assess the legal protection level of databases. This approach aims to furnish database makers with concrete insights into the extent of legal protection attainable for their databases. By amalgamating theoretical understanding with pragmatic application, the article equips stakeholders in the autonomous mobility sector to implement strategies for ensuring the legal protection of their databases.

II. LITERATURE REVIEW

The article by Vaniš[2] et al. navigates the intersection of databases and legal protection within the realm of autonomous mobility. By exploring the fundamental databases pivotal to this domain, the article highlights their role in urban connectivity. It presents a comprehensive framework for evaluating the characteristics that impact decisions regarding legal protection, offering practical insights by applying these traits to selected databases. In addition, the article delves into the legal framework surrounding database protection, providing a roadmap for understanding the symbiotic relationship between technology and legal protection in the context of autonomous mobility. This paper stems from the article, further elaborating on its concepts and contributing to a deeper understanding.

Andraško et al.[3] examine current legal frameworks in the EU regarding data management in connected vehicles. They define essential concepts for autonomous vehicles and their corresponding legal framework in relation to fundamental rights. The paper addresses the challenge of assigning responsibility and liability among various actors involved in processing personal data under the General Data Protection Regulation (GDPR). This analysis highlights the complexity of determining liable entities in certain cases of data processing within connected and automated vehicles, particularly considering the broad interpretation of joint controllership.

Mlada et al.[4] explore the evolving area of autonomous vehicle control, ranging from driving assistance to full automation. The focus lies on the relationship between data management and privacy concerns within autonomous systems. These systems rely on sensor data and information from various sources, raising the potential for the mishandling of personal data. Such mishandling could lead to breaches of digital privacy and violations of the GDPR in the European Union. The article delves into the importance of proper personal data handling not only for GDPR compliance but also for ensuring the smooth de-

[2] VANIŠ, M., et al. Possibilities of legal protection for autonomous mobility databases based on their characteristics. 2023 Smart City Symposium Prague (SCSP). In: *IEEE Xplore* [online]. 15. 6. 2023 [2023-06-12]. Available at: <https://ieeexplore.ieee.org/abstract/document/10146235>.

[3] ANDRAŠKO, J., et al. Sustainable data governance for cooperative, connected and automated mobility in the European Union. *Sustainability*. 2021, Vol. 13, No. 19. In: *MDPI* [online]. 24. 9. 2021 [2023-06-10]. Available at: <https://www.mdpi.com/2071-1050/13/19/10610>.

[4] MLADA, M., et al. Protection of personal data in autonomous vehicles and its data categorization. 2022 Smart City Symposium Prague (SCSP). In: *IEEE Xplore* [online]. 13. 6. 2022 [2023-06-06]. Available at: <https://ieeexplore.ieee.org/document/9792557>.

velopment and user adoption of autonomous vehicles. It outlines the functioning of data within autonomous systems, examines data processing from autonomous vehicles, categorizes data, and assesses potential security risks. By addressing privacy issues during the development of autonomous systems, the article suggests measures and steps to protect the personal data of individuals involved in these systems. This article serves to provide insights into the challenges and solutions related to privacy rights in the context of autonomous vehicle technology.

Vaniš et al.[5] delve into the transformation of vehicles from mechanical devices to sophisticated electronic systems with assistance technologies, ultimately progressing toward fully autonomous vehicles. The focus is on data generated by these autonomous vehicles and their legal implications. The article outlines the existing legal framework surrounding data within autonomous vehicles, followed by a presentation of data types categorized based on their degree of personality. The article concludes by proposing software design to assist manufacturers and administrators in addressing data protection challenges. This comprehensive examination of data classification and legal considerations provides valuable insights into data management in the realm of autonomous vehicles.

Szigeti, Csiszár and Dávid Földes[6] focus on modeling an information system for autonomous vehicles and their operations. Their article deals with the architecture and functions of this complex system, encompassing both operators and users. Additionally, it addresses functions related to passenger and operator information management, along with efficient data structuring. The study's findings impact information process planning and development projects, paralleling this article, where data integration and protection are crucial for the effective operation of new systems.

Benyahya et al.[7] conduct comprehensive analysis of GDPR requirements pertinent to the automated city shuttles (ACSs) ecosystem, addressing data collection, storage, usage, and transmission. The study delves into data processing principles, data subjects' rights, and data controllers' obligations, emphasizing the multifaceted roles of ACSs stakeholders. Furthermore, the study explores the alignment of privacy laws with security technologies and examines the gap between legal definitions and practical implementation of privacy-preserving techniques. In light of GDPR challenges, the study advocates for bolstering data protection laws. This interdisciplinary approach aims to effectively navigate overlapping stakeholder roles and the complex implementation of data privacy-preserving methods within the ACSs ecosystem.

As technological advancements rapidly progress to enhance the human experience, the demands for cybersecurity to adhere to the European Union GDPR have increased.

---

[5] VANIŠ, M., et al. Classification of non-personal data in autonomous vehicles. 2022 Smart City Symposium Prague (SCSP). In: *IEEE Xplore* [online]. 13. 6. 2022 [2024-04-28]. Available at: <https://ieeexplore.ieee.org/abstract/document/9792536>.

[6] SZIGETI, S., CSISZÁR, C., FÖLDES, D. Information management of demand-responsive mobility service based on autonomous vehicles. *Procedia Engineering*. 2017, Vol. 187, pp. 483–491. In: *ScienceDirect* [online]. [2023-05-27]. Available at: <https://www.sciencedirect.com/science/article/pii/S1877705817319343>.

[7] BENYAHYA, M. et al. The Interface of Privacy and Data Security in Automated City Shuttles: The GDPR Analysis. *Applied Sciences*. 2022, Vol. 12, No. 9, pp. 4413. In: *MDPI* [online]. 27. 4. 2022 [2023-05-28]. Available at: <https://www.mdpi.com/2076-3417/12/9/4413>.

The article by Jackson[8] delves into the intricate intersection of technology, cybersecurity, and data privacy by examining the GDPR's implications within the realm of artificial intelligence (AI). It focuses on the legal challenges posed by the GDPR for autonomous cybersecurity systems, which are at the forefront of technological development. These systems are anticipated to play a crucial role in addressing expanding cyber threats and resource limitations. Part I of the article explores the practical aspects of AI-based network defense systems, highlighting their potential, limitations, and technical application in the cyber domain. Part II delves into the GDPR's privacy implications and legal complexities, particularly as they relate to the development of AI-driven network defense. The article concludes by discussing legislative considerations for potential future US data privacy laws, proposing exemptions for information security-related data processing, liability limitations, permissible use of anonymized data, and defined uses of repurposed data. This exploration intersects with our discussion on the legal protection of databases in the context of autonomous mobility, as both topics involve the intricate relationship between technological advancement, legal regulations, and the evolving landscape of data privacy and security.

Finck and Pallas[9] critically examine the differentiation between personal and nonpersonal data under the GDPR from legal and computer science perspectives. The challenges of categorizing de-personalized data and the uncertainties in defining anonymous data are highlighted. The article also underscores the divergence in definitions among GDPR, the Article 29 Working Party, and national supervisory authorities. It delves into the technical foundations of anonymization and applies these principles to practical case studies involving personal data on blockchains. The conclusion suggests that residual risks persist even with anonymization, emphasizing the relevance of risk assessment within the GDPR framework, a notion that aligns with our efforts to establish guidelines for assessing and safeguarding the legal protection of databases in the dynamic domain of autonomous mobility.

Gaeta[10] addresses several regulatory challenges related to road safety and data protection in the context of autonomous driving. The article explores topics such as driver roles, the homologation of self-driving vehicles, and the regulation of various types of data used and generated by autonomous driving systems, taking into account the levels of automation outlined in the SAE standard J3016. With the evolution of autonomous vehicles, the need for comprehensive regulations is evident to ensure public safety and data privacy. The article's insights contribute to the broader discourse on how legal considerations play a crucial role in shaping the future of autonomous driving technologies.

---

[8] JACKSON, B. W. Cybersecurity, privacy, and artificial intelligence: an examination of legal issues surrounding the European Union General Data Protection Regulation and autonomous network defense. *Minnesota Journal of Law, Science & Technology.* 2020, Vol. 21, No. 1, p. 169. In: *Minnesota Journal of Law, Science & Technology* [online]. 6. 6. 2020 [2023-05-30]. Available at: <https://scholarship.law.umn.edu/cgi/viewcontent.cgi?article=1476&context=mjlst>.

[9] 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, pp. 11–36. In: *Oxford Academic* [online]. 10. 3. 2020 [2023-05-22]. Available at: <https://academic.oup.com/idpl/article/10/1/11/5802594>.

[10] GAETA, M. C. The regulation of certain aspects of autonomous driving in the Italian legal system. *European Journal of Privacy Law & Technologies.* 2022, Vol. 1. In: *European Journal of Privacy Law & Technologies* [online]. [2023-06-12]. Available at: <https://universitypress.unisob.na.it/ojs/index.php/ejplt/article/view/1565/1077>.

The article by Míšek[11] investigates the interpretation of the amended Copyright Act's "Open Data Act" and proposes solutions to challenges encountered by database makers. It examines the official work exception and its legal framework, along with sui generis database rights and their exceptions. The article's focal point is the interplay between these two areas, specifically analyzing the feasibility of applying the official work exception to sui generis database rights. Moreover, it addresses the transitional provision introduced by the "Open Data Act." This discussion underscores the complexities faced by database makers. The article contributes to a nuanced understanding of how legal amendments affects the utilization of public sector databases and provides insights for resolving interpretation challenges for creators within this evolving legal framework.

Böhm[12] evaluates the forthcoming EU legislative changes regarding Event Data Recorders (EDR) and their suitability for elucidating accidents involving automated vehicles. It proposes a data model derived from real accident reconstructions involving Advanced Driver Assistance Systems (ADAS) equipped vehicles and tailored crash tests. Authors, contributors to the AHEAD working group, have formulated a data model aligned with automated driving requirements, with the model's structure outlined. The article underscores the advantages of preserving internal video/photo feeds from vehicle camera systems alongside object data.

Kuutti et al.[13] explore improving vehicle safety through driver assistance and autonomy. It emphasizes Deep Learning-based obstacle detection, radar-camera fusion, and perception enhancement in challenging conditions. It addresses a research gap in radar-based perception, providing insights into the current state of camera and radar-based systems for autonomous vehicles. The study contributes to the discussion on sensor fusion and perception, linking to the theme of data-driven safety in autonomous mobility.

The paper by Růžička et al.[14] delves into the realm of intelligent transport systems and the utilization of big data, focusing on specific sources and outcomes of this data within the context of the Czech Republic. While the content of the abstract does not explicitly address legal protection of databases, it connects to the broader landscape of data-related matters. As the article explores the use of data for proposing traffic solutions and draws comparisons between international practices and the local situation in Prague, it indirectly touches on considerations of data ownership, privacy, and potentially legal frameworks related to data utilization.

---

[11] MÍŠEK, J. Sui Generis Database Right and Official Work Exception. *Conference on Grey Literature And Repositories*. 2019. In: Národní technická knihovna – Institucionální digitální repozitář [online]. 24. 10. 2018 [2023-05-24]. Available at: <https://repozitar.techlib.cz/record/1294/files/Fulltext_Misek_EN.pdf>.

[12] BÖHM, K., et al. New developments on EDR (Event Data Recorder) for automated vehicles. *Open Engineering*. 2020, Vol. 10, No. 1, pp. 140–146. In: *degruyter.com* [online]. [2023-05-08]. Available at: <https://www.degruyter.com/document/doi/10.1515/eng-2020-0007/html?lang=de>.

[13] KUUTTI, S., et al. A survey of the state-of-the-art localization techniques and their potentials for autonomous vehicle applications. *IEEE Internet of Things Journal*. 2018, Vol. 5, No. 2, pp. 829–846. In: *Cornell University* [online]. [2023-06-11]. Available at: <https://arxiv.org/abs/2303.04302>.

[14] RŮŽIČKA, J., et al. Big Data Application for Urban Transport Solutions. Smart City Symposium Prague (SCSP) IEEE. 2022. In: *IEEE Xplore* [online]. 13. 6. 2022 [2023-05-03]. Available at: <https://ieeexplore.ieee.org/abstract/document/9792538>.

III. DATABASE CHARACTERISTICS

In this section, we will focus on various characteristics and key parameters of databases that play a significant role in determining the legal protection of databases. Each of these characteristics provides a new insight into databases and influences our approach to legal aspects related to their protection and effective utilization.

III.1 Method of Data Collection

This characteristic focuses on the method of collecting relevant data. The key consideration for this characteristic is the subsequent preprocessing complexity of this data. We base our analysis on the following data collection and data modification methods before storing them in the database:

- Continuous Data Collection (CDC): Data collection occurs continuously, and data can be directly stored in the database. This scenario arises when values are stored in the database that are either entered through a user interface or are simply copied from other sources without requiring conversion.
- Standardized Data Modification (SDM): Data needs to be modified in a standardized process after being captured, before being stored in the database. This case is common with data from detectors, where accompanying materials typically describe the type of data a given detector records. Based on this description, data preprocessing occurs, often through scripts that transform the data into the required format.
- Custom Data Modification (CDM): Data needs to be modified after being captured, but no standardized procedure is available. This situation arises when only measured data is available, but there is uncertainty about how to further process it. This is frequently encountered in proprietary solutions or areas lacking clear standardization procedures. An example is data from Roadside Units within C-ITS data, where standardization might be lacking.
- User-Generated Data (UGD): Data is generated by users based on other data. This is a typical scenario when creating a database, such as digitizing a paper-based record system into electronic format.

III.2 Extent of Captured Data

A significant characteristic in certain cases can be the extent of captured data. Here, it is necessary to distinguish between two approaches:

- The extent of data captured and
- The extent of data that is further stored in the database.

In some cases, these two approaches might coincide. For instance, with an induction loop, it is possible to set the recording frequency, and data is measured and recorded simultaneously.

However, when dealing with data from autonomous vehicles, the situation becomes more complex. This is due to the possibility of facing database capacity issues when attempting to simultaneously store this data. Such problems could arise in a scenario where the vehicle is not specially equipped to handle the volume.

Based on the aforementioned information, we are only addressing the extent of data being stored in the database. Drawing from discussions and expert insights, we propose the following ranges:

- Small (S): The database accepts data at a rate of 0.1 MBit/s.
- Medium (M): The database accepts data at rates between 0.1 MBit/s and 10 MBit/s.
- Large (L): The database accepts data at rates exceeding 10 MBit/s.

III.3 Substantial Investment

This parameter determines the significance of the contributor's investment in acquiring or establishing the database, encompassing both the quantitative (and therefore financial) contribution and the intangible effort required to achieve the final database, data quality, and provided information.

From the perspective of this article, we will consider the following categorization of the contributor's investment:

- Significant Investment (SI): This category involves an investment that is objectively substantial and not entirely negligible. Such an investment demands significant human, technical, and financial resources and is not easily replicable by anyone else.
- Partially Significant Investment (PSI): This category represents a degree of importance and effort from the contributor's end. While it might not be as pivotal as a significant investment, it still requires a certain level of financial, material, and intellectual resources.
- Insignificant Investment (II): This category pertains to an investment that is trivial or inconsequential in terms of the contributor's effort and significance. This type of investment can be easily replicated by others and does not necessitate substantial financial or material resources.

III.4 Data Utilization

The primary utilization of data from databases always lies in their transformation into specific information that would otherwise be unattainable. However, this generally applies to all databases, and hence, it is not explicitly listed in the data utilization categories.

For the sake of simplicity, let us consider traffic databases. We have several fundamental options for data utilization:

Providing Real-Time Traffic Information (RTTI): In the case of a real-time updated database, it's possible to extract information and broadcast it directly to relevant areas to inform pertinent parties based on the nature of the information. Navigation devices can be a major recipient of these messages, allowing them to update drivers' current routes. Such a database could aid transportation operators in traffic management.

Scientific Research (SR): Many databases, not just in transportation, can also be used for scientific purposes. For instance, in the realm of transportation, this could involve various predictions and simulations. Another more specific use is database utilization in data mining, which involves uncovering connections not immediately evident through standard usage and examination.

Integration with Other Databases (IOD) for Additional Information: Many databases (for instance, those recording data from induction loops) have limited value if not inte-

grated with other databases. It is only in these cases that acquiring information becomes meaningful.

Long-Term Strategies (LTS): Evaluating data from these databases allows for proposing measures aimed at enhancing safety or traffic flow (e.g., constructing new roads or traffic lights, speed limitations in problematic areas, etc.) based on the data analysis.

III.5 Originality

This subsection focuses on evaluating the originality of a database's structure, which is categorized into three levels:

- Minimal Originality (MO): A database with minimal originality lacks complete independence and primarily constitutes the collection and arrangement of existing information. Due to its insufficient structural originality, this type of database may not be suitable for copyright protection.
- Moderate Originality (MdO): A database with moderate originality introduces new elements that bring a certain level of structural uniqueness. This type of database could be eligible for copyright protection if it meets the relevant requirements for structural originality.
- High Originality (HO): A database with high originality features a wholly independent and genuinely unique structure. This type of database is most suitable for copyright protection, especially if its structure constitutes the author's original intellectual creation.
- The assessment of a database's structural originality is crucial in determining the suitability of various forms of legal protection and can influence decisions regarding database protection. This incorporated criterion offers an additional perspective for evaluating a database's structure and its potential safeguarding.

III.6 Data Types in Terms of Legislation

Although this characteristic is not directly related to the legal protection of the database, it is important to consider legislative obligations that pertain to stored data. This characteristic focuses on the type of data stored in databases in terms of legislation. While the project primarily aims at the legal protection of databases, it is also essential to account for legislative responsibilities that apply to stored data and vary according to their type. Within this characteristic, we differentiate between the following two data types:

- Personal Data (PD): This category involves data that can be used to identify a specific individual, directly or indirectly. Ensuring the privacy and security of personal data is a crucial legal and ethical consideration.
- Non-Personal Data (NPD): This category includes data that does not directly identify individuals and is generally considered non-sensitive. However, it is still important to adhere to relevant regulations when handling such data.

In-depth exploration of the protection of personal data and non-personal data processed in the context of autonomous mobility is a focal point of other research projects.[15]

[15] e.g. CK02000188 – Protection of non-personal data and databases in autonomous systems or TL03000691 – Privacy and personal data proteciton in autonomous driving systems.

IV. DATABASE OVERVIEW

This section provides a comprehensive exploration of databases intricately linked to the field of autonomous mobility. It is understood that this overview is not exhaustive. The primary aim is to empower readers with the ability to categorize their databases within the scope of the databases mentioned below. Subsequent sections also delineate the characteristics of these databases, including the determination of the level of legal protection. For clarity, this section is divided into two subsections: "Traffic Databases" and "Operational Databases."

IV.1 Traffic Databases

In the context of autonomous mobility, the first subsection focuses on databases, which play a key role in the acquisition of traffic information. These databases, ranging from detectors to C-ITS, provide the essential information needed for autonomous vehicles to navigate and make informed decisions in dynamic traffic environments.

• Database from Detectors

This database contains information from various detectors that monitor traffic on roads. Single detectors track basic parameters such as speed and travel time, while paired detectors enable the measurement of vehicle operational characteristics based on paired points. This data is crucial for traffic analysis and optimization.

• Floating Car Database, Crowdsourcing Database

This database aggregates information from mobile devices and vehicles participating in traffic. Floating Car Data includes details about vehicle location, speed, and route, while crowdsourcing data is collected voluntarily from users sharing transportation information. These data are valuable for monitoring real-time traffic and tracking transportation trends.

• C-ITS Database

The database for Cooperative Intelligent Transportation Systems includes information from interactions between vehicles and infrastructure. This data encompasses hazard warnings, operational condition updates, and other communication elements between vehicles and traffic infrastructure. C-ITS data are essential for enhancing safety and operational efficiency.

• Mobile Operator Database

This database contains anonymized data about the movement of mobile devices, obtained from mobile operators. Such data enables the analysis of people and vehicle movement, impacting infrastructure planning and transportation services.

• Transportation Database from JSDI/NDIC:

The Unified Traffic Information System (JSDI) and the National Data Information Center (NDIC) provide transportation data, including real-time traffic, road conditions, and traffic event information. These details are pivotal for traffic management and providing current information to drivers.

• Accessibility-Related Database

Databases related to the availability of parking spaces and charging stations play a significant role in the context of autonomous mobility. These databases provide information about the availability of parking spots and charging stations for autonomous vehicles, which is crucial for proper route planning and efficient utilization of available resources.

• Database of traffic accidents

This database contains information about traffic incidents, including location, time, involved vehicles, and injuries. These data are fundamental for safety analysis in road traffic and identifying high-risk areas.

IV.2 Operational Databases

This subsection delves into databases that are operational in nature. In this context, the term pertains directly to autonomous vehicles.

• EDR/DSSAD Databases:

Event Data Recorder (EDR) and Data Storage System for Automated Driving (DSSAD) databases contain information from vehicle black boxes after accidents. These data are crucial for accident reconstruction and enhancing vehicle safety.

• Vehicle Control Unit Databases:

This database includes data from vehicle control units that record various operational parameters of vehicles. These data are important for traffic monitoring, vehicle maintenance, and performance optimization.

• Databases Related to Higher Levels of Automation:

This database focuses on collecting data from vehicles with advanced automation levels, like autonomous vehicles, encompassing vehicle behavior, sensor inputs, decision-making processes, and interactions with the environment. It is crucial for training machine learning algorithms, refining autonomous systems, and advancing the safety and capabilities of self-driving vehicles in real-world scenarios.

• Infotainment Databases:

Infotainment databases hold data integrating multimedia, navigation, communication, and vehicle functions, enhancing in-car experiences. These databases provide access to audio, video, navigation maps, real-time traffic updates, and connectivity features through interactive interfaces. They offer passengers and drivers entertainment, communication, and navigation options, contributing to driving safety and convenience.

V. TYPES OF LEGAL PROTECTION

From the legal point of view, considering the laws of the European Union and the Czech Republic, a database is defined with regard to the protection of and access to data. According to Directive 96/9/EC of the European Parliament and of the Council of 11 March

1996 on the legal protection of databases (the "Database Directive"), a database is to be understood as a "collection of independent works, data or other materials arranged in a systematic or methodical way and individually accessible by electronic or other means". The quoted passage provides the basis for the definition of database across all EU countries, including the Czech Republic and its Act No 121/2000 Coll., on Copyright (the "Copyright Act"), which transposes the Database Directive nearly verbatim.
Both the Database Directive and the Copyright Act distinguish two possible legal protection regimes in relation to databases: the copyright protection regime and a *sui generis* protection regime.

V.1 Copyright protection

A database that represents an author's own intellectual creation in the selection or arrangement of its content is protected under copyright law as a collective work. Such a database enjoys copyright protection automatically by virtue of the law, and no further action or registration are required – as would be the case, for example, for industrial rights.

Copyright protection of databases does not extend to their content and does not interfere with the rights to that content (which itself may be the subject of copyright protection, provided it is a unique creative product of the author). It applies equally to the completed work and to its individual parts and stages of development. A database that satisfies the condition of being the author's own intellectual production is therefore protected by copyright from the moment in which it becomes in any way objectively perceptible, whether or not it has been disclosed.

In addition to the right to claim authorship, the author of a copyrighted database (or its licensed user, where necessary for access to and normal use of the database), generally has the exclusive right to perform or approve the following:

(i) translation, processing, adjustments, or any other alteration of the dababase;
(ii) temporary or permanent reproduction of the database or any part thereof;
(iii) any distribution of the database or copies thereof (including the results of the operations under i. above) to the public; and
(iv) any communication, display or demonstration of the database (including the results of the operations under i. above) to the public.

A specific case would be the so-called employee copyright, where the employer of the author who created the database to fulfill his/her obligations under an employment contract or equivalent contract is entitled to exercise the copyright during his/her lifetime or existence.

Unauthorized interference with a copyrighted database may result in a number of claims, including, for example, a claim to desist from further tampering, to remove a modification previously made, to seek an apology or damages, or to recover unjust enrichment.

As the nature of copyright law suggests, the key to ensuring, with some degree of likelihood, that a database will be granted this legal protection regime is the originality criterion.

From the perspective of database owners, a certain safeguard of originality which does not require any significant investment, might be to consistently treat their databases, or

the contents of which they are made, as confidential information or trade secrets and to address this accordingly in a contractual manner.

The originality criterion can be further enhanced, albeit with increased investment, by leveraging the expertise and know-how of skilled staff involved in the creation of the database, but also, for example, by acquiring AI and advanced technologies that can generate unique results capable of bringing innovation to the entire industry.

From this angle, the use of data, especially in the field of science and innovation, can be an indicative criterion in determining the legal status of a database. After all, unoriginal and easily replicable databases would presumably not make any significant contribution to the field in the first place.

Finally, the method of data collection will be a similarly telling criterion. As a rule of thumb, the greater the involvement of human and non-standard factors (as opposed to automated processes), the higher the potential originality of the resulting database.

V.2 *Sui* generis protection

In addition to copyright protection, or in cases where the copyright regime does not apply (usually because the structure of the database does not achieve a sufficient degree of originality), databases may also be protected under what is known as the *sui generis* regime. *Sui generis* protection applies to a database which was obtained, verified or presented by making a qualitatively and/or quantitatively substantial investment.

The concept of investment includes, in particular, any means intended for the search of existing elements and their collection in the database, which shall not overlap with the means used for the creation of these elements as such. As confirmed by the CJEU,[16] said interpretation is backed up by the 39th recital of the preamble to the Database Directive, according to which the aim of the *sui generis* right is to safeguard the results of the financial and professional investment made in obtaining and collecting the contents of a database.

As further clarified by the CJEU in the above-mentioned case, investment in the creation of a database may consist in the deployment of human, financial or technical resources but must be substantial in quantitative or qualitative terms. The quantitative assessment refers to quantifiable resources and the qualitative assessment to efforts which cannot be quantified, such as intellectual effort or energy.

Regarding the abovementioned notion of investment, attention should be paid to the so-called "spin-off" doctrine, according to which *sui generis* rights shall be granted exclusively to protect investments directly attributable to producing a database, if they are to respect the spirit of the Database Directive. In other words, no *sui generis* protection should be afforded to databases which are spin-offs or by-products of another (or main) activity. In accordance with this theory, the CJEU has e.g. held in the above-mentioned case that where the relevant database is created practically simultaneously with the formation of its individual elements and no special investment is required for the actual arrangement of these elements in the database, no claim to *sui generis* legal protection of the database arises.

[16] Judgment of the Court (Grand Chamber), Case C-444/02 and C-46/02, Fixtures Marketing, 9 November 2004.

However, evaluating said doctrine should not result in excluding every spin-off effect solely in reliance on a theory.[17] In the view of the Advocate General in the cited opinion, protection should also be granted to databases where the obtaining of elements was initially for the purpose of an activity other than the creation of a database, meaning that "an external database, which is derived from an internal database, should also be covered by protection".

It is not yet clear how high a quantitative or qualitative value an investment must have to become substantial, and this value is likely to vary over time along with the real costs of acquiring the database. However, the current legal literature is inclined toward the view that a substantial investment is likely to be a contribution between the *de minimis* and substantial levels.[18] However, a factor which qualifies the investment of the database maker (i.e. the person who takes the initiative and the risk of investing) as substantial may be, for example, that the elements of the database were obtained from sources not available to the general public.[19]

Provided that the database meets the statutory conditions for granting *sui generis* protection (for which no further steps or registration are required, just as in the case of copyright protection), the maker of the database (or the person to whom such a right has been transferred) has the right to prevent the exploitation and/or use of all or a qualitatively substantial part of the database for a period of 15 years from the creation of the database, or, on the contrary, to grant consent to such use of the database. In this respect, the *sui generis* legal protection of databases differs substantially from copyright protection since it also protects the content of the database itself.

he use of the entire content of the database or of a qualitatively or quantitatively substantial part of it includes its (permanent or temporary, direct or indirect) reproduction, or making it available to the general public. The use of the database made available for one's own needs is not affected by this restriction, provided that the database is used commonly and reasonably, not systematically or repeatedly, and without prejudice to the legitimate interests of the database maker, and provided further that the user does not cause harm to the holder of rights protected by copyright and contained in the database.

To maximize the chances of their database being granted *sui generis* protection, it is particularly important for the database maker, or a person in a similar position, to consistently separate the process of the creation of the individual database elements from the creation of the database (i.e., the particular arrangement of these existing elements) as such. Thus, if the direct data collection method can be replaced by alternatives that require a separate investment (either of resources or energy), such a procedure should be the preferred option.

Furthermore, it is advisable that key players in the field of autonomous mobility consistently identify and treat the essential elements that often make up databases in this in-

[17] Opinion of Advocate General, Case C-46/02, Fixtures Marketing Ltd v. Oy Veikkaus Ab, 8 June 2004, paragraphs 50-51, 67.

[18] TELEC, I., TŮMA, P. § 88a [Podstatný vklad]. In: Ivo Telec – Pavel Tůma *Autorský zákon [The Copyright Act]. 2nd ed.* Praha: C. H. Beck, 2019, p. 846, marg. No. 3. Recital 7 of the Database Directive also provides some interpretative guidance in this matter.

[19] See the above-cited Fixtures Marketing case.

dustry as confidential information or, better yet, trade secrets. This will provide increased protection for these elements individually and will at the same time increase the chances of the database that is formed by them being afforded *sui generis* protection.

Furthermore, the scope of the data processed may be indicative for determining the *sui generis* regime for the database, as it may imply increased complexity. Where technically feasible, it is advisable to create databases with a higher volume of elements.

In addition to the above, the use of AI or technological means in some cases also speaks in favor of the challenging nature of database acquisition, verification or demonstration. Since the originality criterion is not decisive for *sui generis* protection, an investment in sophisticated automation of data processing is recommended wherever the minimal originality of the database means that copyright protection is most likely out of the question.

Last but not least, it is desirable that the various databases of a single maker be interconnected as part of a more holistic strategy or process, as there is a fair chance that the effort or material investment in their acquisition will be regarded as substantial in such context.

The position of the CJEU, from which the interpretation of all the essential criteria for granting *sui generis* protection derives, is in some respects still not clear; hence, it will be necessary to constantly monitor developments in this area so as to be able to adapt the above-suggested strategies accordingly.

VI. PROCEDURE FOR DETERMINING LEGAL PROTECTION OF DATABASES

For the purpose of determining appropriate legal protection for databases, this article establishes levels of legal protection for databases, which will be further discussed in the following Subsection 6.3. The resulting level of legal protection for a database will be separately calculated for copyright protection and sui generis protection based on the characteristics of databases described in Section 3. The resulting numerical value assigned to each characteristic will guide the determination of the corresponding level of legal protection for the database according to the scale provided in Table 6.

VI.1 Assignment of Ratings to Database Characteristics

The assignment of evaluations to individual characteristics followed a procedure based on multicriteria decision theory.

Upon determining the states of database characteristics (as per Section 3), the corresponding evaluations must now be located. For clarity, a separate table has been devised for each characteristic:

- Method of Data Collection – Table 1,
- Extent of Captured Data – Table 2,
- Substantial Investment – Table 3,
- Data Utilization – Table 4,
- Originality – Table 5.

Table 1 – Evaluation of Method of Data Collection

Method of Data Collection	Evaluation
Continuous Data Collection (CDC)	4
Standardized Data Modification (SDM)	3
Custom Data Modification (CDM)	2
User-Generated Data (UGD)	4

Table 2 – Evaluation of Extent of Captured Data

Extent of Captured Data	Evaluation
Small	4
Medium	3
Large	2

Table 3 – Evaluation of Substantial Investment

Substantial Investment	Evaluation
Significant Investment (SI)	4
Partially Significant Investment (PSI)	3
Insignificant Investment (II)	2

Table 4 – Evaluation of Data Utilization

Data Utilization	Evaluation
Real-Time Traffic Information (RTTI)	2
Scientific Research (SR)	4
Integration with Other Databases (IOD)	3
Long-Term Strategies (LTS)	2

Table 5 – Evaluation of Originality

Originalita	Evaluation
Minimal Originality (MO)	1
Moderate Originality (MdO)	2
High Originality (HO)	4

VI.2 Computation of Legal Protection Coefficients

In this section, based on the evaluation of individual database characteristics, coefficient calculations will be conducted to determine appropriate legal protection. Two coefficients will be computed – one for copyright protection (k_A) and the other for sui generis protection (k_S). The process for their calculation is as follows:

k_A = 2 * Method + Extent + Investment + Utilization + 3 * Originality
k_S = 2 * Method + Extent + 3 * Investment + Utilization + Originality

The coefficients in these criteria calculations vary based on their significance for the respective legal protection. Using these coefficients, the level of legal protection is determined, as described in the following subsection.

VI.3 Levels of Legal Protection for Databases

When evaluating and determining the legal protection of databases, we have devised a proposal for four levels (Yes, Rather Yes, Rather No, No) that serve for better understanding and assessment of the applicability of the respective legal protection regime. It is important to note that these levels are specific and are intended solely to provide guidance for evaluating the analyzed criteria.

The level of legal protection is assigned based on the calculation of the coefficient k_A for copyright protection and k_S for sui generis protection. This is succinctly summarized as follows:

- Yes: The database exhibits characteristics sufficient for copyright/sui generis legal protection. There is a high likelihood that this database meets the legal requirements and criteria for the given type of protection. Consideration should be given to adopting appropriate measures to ensure legal protection for this database.
- Rather Yes: The database likely exhibits characteristics sufficient for copyright/sui generis legal protection, but there are certain doubts or factors that could affect the possibility of achieving protection. It is recommended to conduct further analysis and review specific information and circumstances related to this database.
- Rather No: The database probably does not exhibit characteristics sufficient for copyright/sui generis legal protection, but in certain cases, it could achieve some form of protection if additional information with significant character were available or if there were substantial changes in the structure or content of the database. It is advisable to conduct further analysis and consider options for achieving protection.
- No: The database likely does not meet the criteria for copyright/sui generis legal protection. In such a case, it is unlikely that this database would fulfill the requirements and criteria for the respective protection regime. It is recommended to explore alternative forms of protection or strategies for ensuring its protection.

Table 6 – Assignment of legal protection levels

The value of coefficient k_A or k_S	Level of legal protection
12 – 17,5	No
17,5 – 23,5	Rather No
23,5 – 28,5	Rather Yes
28,5 – 34	Yes

VII. DECISION ON DETERMINING LEGAL PROTECTION

This section builds upon previous research focused on identifying databases associated with autonomous mobility and analyzing the characteristics of these databases in terms of potential legal protection regimes. Its main aim is to demonstrate the process of calculating degrees of legal protection for databases based on their characteristics.

The characteristics for each database from Section 4 are listed in Table 7. Each row corresponds to a specific database, and each column corresponds to the value of the characteristic. For clarity, abbreviations are utilized within the table. The meanings of these abbreviations can be found in the relevant characteristic subsection that is a part of Section 3.

Based on Table 7, the coefficients for copyright and sui generis protection are calculated. The results of these coefficients, as well as the resulting level of legal protection, are included in Table 8. This table also includes an additional column containing specific information for further refining the legal protection assessment. This information can provide more precise insights into the legal protection status of each specific database.

Table 7 – Database List and Their Characteristics

Characteristic / Database	Collection method	Extent	Substantial Investment	Data utilization	Originality
Detectors	SDM/CDM	Medium / Large	II	All types	MO/ MdO
Floating Car	SDM	Medium	II	All types	MdO
C-ITS	SDM/CDM/UGD	Medium	II	All types	MdO
Mobile Operator	SDM/CDM	Large	PSI	SR, IOD, LTS	MO
JSDI/NDIC	SDM/UGD	Medium / Large	II/PSI	All types	MO/ MdO
Accessibility	SDM	Medium	II	RTTI, SR	MO/ MdO
Traffic Accidents	UGD	Small	PSI	SR, IOD, LTS	MO
EDR/DSSAD	CDC/SDM	Small / Medium	II	SR, IOD	HO
Vehicle Control Unit	SDM	Medium / Large	II	SR, IOD	MO
Higher Levels of Automation	SDM	Medium / Large	PSI	SR, IOD, LTS	HO
Infotainment	CDC/SDM/CDM	Medium	II	SR, IOD	MO

Table 8 – Database list with calculated coefficients and information for refining legal protection

Protection / Database	Copyright	Sui generis	Further specific information for refining legal protection
Detectors	Rather No $k_A = 23$	Rather No $k_A = 21$	Technical specifications of detectors Context and data accuracy
Floating Car	Rather No $k_A = 21$	Rather No $k_A = 19$	Geographical coverage and data availability Specifics and diversity of data items
C-ITS	Rather No $k_A = 23$	Rather No $k_A = 21$	Technical specifications and standards Network infrastructure and interoperability
Mobile Operator	Rather No $k_A = 21$	Rather No $k_A = 23$	Database structure and organization Investments in the database
JSDI/NDIC	Rather Yes $k_A = 26$	Rather Yes $k_A = 26$	Unique attributes and data quality Innovations and technological progress
Accessibility	Rather No $k_A = 21$	Rather No $k_A = 19$	Data quality and updates Integration with other databases
Traffic Accidents	Rather No $k_A = 19$	Rather No $k_A = 22$	Detailed event description Quality and reliability of accident data
EDR/DSSAD	Yes $k_A = 29$	Rather No $k_A = 23$	Specific data collection methodology Complexity of data relationships
Vehicle Control Unit	Rather No $k_A = 20$	Rather No $k_A = 20$	–
Higher Levels of Automation	Yes $k_A = 30$	Rather Yes $k_A = 26$	Technical specifications and innovative technologies Unique data compilation and organization
Infotainment	Rather No $k_A = 20$	Rather No $k_A = 20$	Data sources

VIII. DISCUSSION

This article addresses the issue of legal protection for databases in the context of autonomous mobility. The main objective of the article is to present a process that database creators can use to preliminarily determine the legal protection of databases in this innovative field. The article examines various characteristics of databases and proposes coefficients for copyright protection and sui generis protection based on these characteristics. The result is different levels of legal protection for different databases in autonomous mobility.

In the future, a more detailed analysis of the individual characteristics of databases that influence legal protection could be conducted. This may include more specific examples and scenarios illustrating how individual characteristics affect the calculation of coefficients for copyright and sui generis protection.

Furthermore, there are plans to conduct a survey of specific databases related to autonomous mobility, including the precise determination of their characteristics. Based on this, minor adjustments to the coefficient calculation system and the characteristics themselves may be made.

Additionally, it is necessary to continue monitoring the development of autonomous mobility on an international level. This development may involve new technological and legislative documents directly related to the issue of legal protection for databases.

As a long-term goal stemming from this article, an alternative form of legal protection for databases specific to the field of autonomous mobility could also be considered.

IX. CONCLUSION

In conclusion, it is imperative not to overlook the significance of legal protection for databases within the realm of autonomous mobility. The continuously changing area of technology and data-driven innovations demands a comprehensive understanding of how databases can be safeguarded. As this article has highlighted, the determination of legal protection levels is a complex yet essential process, offering clarity to database makers and stakeholders. Recognizing that databases are not uniform in their characteristics and that their legal protection should be tailored accordingly underscores the importance of this endeavor.

This article significantly contributes to the field by proposing a systematic approach for determining the level of legal protection, as outlined in Section 6, based on the unique characteristics of databases within the context of autonomous mobility. By introducing a methodology that combines theoretical understanding with practical application, it empowers database makers to make informed decisions regarding legal protection. The introduction of coefficients for copyright and sui generis protection provides a structured framework for evaluating databases, allowing for a nuanced assessment of their legal status.

Furthermore, this article aims to empower database makers in the field of autonomous mobility by providing them with a foundational understanding of the legal protection aspects of their databases. It serves as a compass, guiding them through the intricate process of assessing their databases' characteristics and determining the appropriate level of legal protection. By offering a structured methodology and coefficients for evaluation, it equips database makers with the tools they need to navigate the multifaceted legal landscape. Ultimately, this article strives to enable them to make informed decisions, safeguard their intellectual property, and foster innovation in the rapidly evolving world of autonomous mobility.

In closing, this article underscores the paramount importance of addressing the legal protection of databases in the sphere of autonomous mobility. It presents a systematic approach, offering clarity to database makers and stakeholders, while also paving the way for future advancements in safeguarding intellectual property within this innovative industry. As autonomous mobility continues its trajectory of transformation, the need for robust legal frameworks becomes increasingly evident. This article, by bridging the gap between technology and law, aspires to be a cornerstone for database makers and the legal community, fostering a harmonious environment where innovation and protection coexist. It is our hope that this work will inspire further exploration, research, and collaboration in the ever-evolving landscape of autonomous mobility databases.