

Rail to Digital automated up to autonomous train operation

D7.2 – IP assessment and Data Ownership Model

Part 1 Legal Abstract

Due date of deliverable: 31/05/2024

Actual submission date: 09/12/2024

Leader/Responsible of this Deliverable: Dr. Philipp Neumeier, DB InfraGO AG

Reviewed: Y/N

Document status		
Revision	Date	Description
01	31/05/2024	First issue (Executive Summary and Conclusions missing)
02		TMT Standards added
03		Draft reviewed by partners
04		Final version after partner reviews and law firm edits
05	05/11/2024	Deliverable divided into 2 Parts

Project funded from the European Union's Horizon Europe research and innovation programme		
Dissemination Level		
PU	Public	x
SEN	Sensitiv – limited under the conditions of the Grant Agreement	

Start date: 01/12/2022

Duration: 18 months

ACKNOWLEDGEMENTS



a. This project has received funding from the Europe's Rail Joint Undertaking (ERJU) under the Grant Agreement no. 101102001. The JU receives support from the European Union's Horizon Europe research and innovation programme and the Europe's Rail JU members other than the Union.

REPORT CONTRIBUTORS

Name	Company	Details of Contribution
Dr. Philipp Neumeier	DB InfraGO AG	Data Factory Project Circle Lead
Natalie Kelle	DB AG	In-house Lawyer (Legal opinion)
Christoph Jacobs	DB AG	In-house Lawyer (Legal opinion)
Sebastian Dubiel	DB Systel GmbH	R2DATO WP7 Project Manager
Prof. Dr. Christoph Wagner	Morrison & Foerster LLP	External Lawyer (Author of legal analysis)
Frederic Antoine	ATSA	Expert Review
Dominik Kevicky	AZD	Expert Review
Saro Thiyagarajan	FT/Wabtec	Expert Review
Lars-Kristian Vognild	NRD	Expert Review
Tom Jansen	NS	Expert Review
Sebastiaan Linssen	NS	Expert Review
Oliver Lehmann	SMO	Expert Review
Philippe David	SNCF	Expert Review
Hardik Jain	THD/Thales	Expert Review
Maik Baehr	Siemens	Expert Review
Bastian Simoni	Alstom	Expert Review
Lars Bergmann	Siemens	Expert Review
Michele Bruzzo	Hitachi	Expert Review
Hájek Jiří	AZD	Expert Review
Michal Novak	AZD	Expert Review
Rao Xiaolu	SBB	Expert Review

Disclaimer

The information in this document is provided “as is”, and no guarantee or warranty is given that the information is fit for any particular purpose. The content of this document reflects only the author’s view – the Joint Undertaking is not responsible for any use that may be made of the information it contains. The users use the information at their sole risk and liability.

EXECUTIVE SUMMARY

In light of the ever-growing demand for passenger and freight transportation, the DB InfraGO AG (hereinafter “**InfraGO**”), along with other European rail transport companies (hereinafter “**RTC**”), rail infrastructure companies (hereinafter “**RIC**”) and industry partners (hereinafter “**Partner Organizations**”), is participating in the “Rail to Digital and Automated Train Operations” (R2DATO) project, which aims to leverage digitalization and automation to develop the Next Generation Automatic Train Control (ATC) and scalable Digital and Automatic (up to Autonomous) Train Operation (DATO) capabilities to enhance the capacity of the existing rail networks. This requires the acquisition, storage, simulation, and processing of a significant amount of sensor data for machine learning training, AI use cases, and validation purposes. This challenge is addressed by Work Package 7 (hereinafter “**WP7**”) of the R2DATO project, which is setting the framework for a collaborative solution in the rail sector, the Pan-European Data Factory.

Across Europe, various Partner Organizations develop solutions to collect, store and annotate sensor data and build up all required infrastructure for processing the data in so-called “**Data Centers**”. The aim is for these Data Centers to interconnect with those of other Partner Organizations across Europe, creating a collective data repository known as the “**Pan-European Data Factory**.” The main objective of this integration is to facilitate access to the data stored within the individual Data Centers under predefined conditions, allowing not only the Partner Organizations to utilize this data but also third parties via Open Data Sets.

As part of WP7 of the R2DATO project, InfraGO has undertaken an initial legal evaluation focusing on non-technical aspects of the Pan-European Data Factory that addresses: (I.) A feasible setup for non-discriminatory access to data and computation infrastructure, (II.) data ownership and (III.) the publication of Open Data Sets. To this end, InfraGO has engaged the services of external legal counsel of Morrison & Foerster LLP and consulted with experts in the field, including ATSA, AZD, FT/Wabtec, NRD, NS, SMO, SNCF, THD/Thales, Siemens, Alstom, Hitachi and SBB.

- I. To ensure non-discriminatory access of stakeholders to the data and computation infrastructure of the Pan-European Data Factory, it is essential to establish a central coordinating body the so-called Pan-European Entity. This entity should be independent of the individual Partner Organizations and have a reliable, long-term organizational framework. It is recommended to structure the entity as a registered association dedicated to ensuring interoperability among the network of individual Data Centers in different European countries, thereby enhancing the overall effectiveness and efficiency of data management and access in the European rail sector. The decentralized structure of the Pan-European Data Factory should only require the establishment of a coordinating and mediating body with a narrowly defined mandate, which should be manageable without significant human and material resources. It will not produce its own data and will not provide its own server capacity or transmission capacity but will merely act as a coordination vehicle for the operators of the individual Data Centers.

The tasks of the Pan-European Entity will include (1) the establishment and operation of the backbone network that connects the individual Data Centers, (2) the definition of and (3) monitoring compliance with common standards, interfaces and conditions for the collection, processing, provision and use of data to ensure legal compliance and interoperability, as well as the (4) adoption of nondiscriminatory criteria for third-party access to the Pan-European Data Factory.

The establishment of the Pan-European Entity will not require merger control clearance under EU law by the EU Commission provided that it primarily serves auxiliary functions that relate to the individual Data Centers of the Partner Organizations as outlined above. In that case the Pan-European Entity is unlikely to perform all the functions of an independent economic entity and thus does not meet the so-called full-function criterion of the Merger Regulation.

However, the Pan-European Data Factory may be subject to notification obligations under national merger control regimes that do not provide for a full-function criterion, as is the case under German law. It is therefore recommended that a comprehensive merger control review be carried out at the pan-European level, based on the revenues of the undertakings cooperating within the framework of the Pan-European Entity as soon as they are known. This review should include all countries where these undertakings generate revenues. It was not possible to carry out such a comprehensive merger control review in all relevant jurisdictions at the time of this analysis since the participating undertakings in the Pan-European Entity are yet to be determined.

- II. The data ownership involved in the project may attract data access claims of third parties. Claims for data access may arise under antitrust law. Depending on how data offerings develop in the future, the individual Data Centers and Pan-European Data Factory will likely hold a relatively strong (if not dominant) market position regarding the data offerings or AI models that are suitable for enabling automatic rail operations. Third-party access claims against the providers of the Data Centers, i.e. the Partner Organization operating them, are likely to be justified if these third parties wish to use the data or AI models available in the individual Data Center or Pan-European Data Factory to train their own AI models or independently carry out the certification of autonomous trains. Such access will have to be granted under FRAND conditions, i.e. under fair, reasonable, and non-discriminatory terms, which are defined centrally by the Pan-European Entity. Thus, access to the individual Data Centers or Pan-European Data Factory must generally be granted for an appropriate fee.

Data access claims by third parties under other regulations are unlikely. The Data Act specifically only applies to claims by “users”, who in this case are the RTCs that own and operate the train and not the operators or the Data Centers. Access claims by public sector bodies are subject to a demonstration of an exceptional need to carry out their statutory duties in the public interest, which exists if the data is necessary to respond to a public emergency such as a natural disaster. Furthermore, even if a data access claim exists under the Data Act, it would only relate to the sensor data collected by the sensors on the train and the associated metadata, not the processed data such as the compressed sensor data, the selected sensor data with annotations or the artificial sensor data.

- III. Regarding the publication of the Open Data Set and accompanying software, it is recommended to adopt a dual licensing model. A dual licensing model would allow non-commercial use to be free of charge under an open-source license, while commercial use would require a commercial license subject to a fee. This approach balances the goal of democratization and the possibility for collaborative (further) development of the Open Data Sets with the need for control and potential for monetization by the Partner Organizations. Licensing of the Open Data Set and software will be managed by the Partner Organizations,

either jointly or individually, on the basis of centrally defined conditions by the Pan-European Entity.

It is advisable that any further developments by users, particularly those under a commercial license, should include a license back to the Partner Organizations, potentially with appropriate remuneration. The commercial license can either be offered as a purchase or rental license. In case of a rental license the preferred compensation model needs to be considered.

Abbreviations and Acronyms

AG	Stock corporation (Aktiengesellschaft)
AktG	German Stock Corporation Act (Aktiengesetz)
ATC	Next Generation Automatic Train Control
AO	German Tax Code (Abgabenordnung)
BGB	German Civil Code (Bürgerliches Gesetzbuch)
BKartA	German Federal Cartel Office (Bundeskartellamt)
CEF 2	EU funding program “Connecting Europe Facility 2”
CJEU	Court of Justice of the European Union
Connected Product	An item that obtains, generates or collects data concerning its use or environment and that is able to communicate product data via an electronic communications service, physical connection or on-device access, and whose primary function is not the storing, processing or transmission of data on behalf of any party other than the user (Art. 2 para. 5 DA).
DA	Data Act (Regulation (EU) 2023/2854)
Data Center	Unit/infrastructure where the storage, annotation and processing of sensor data is carried out. Core component of the “Data Factory”.
Data Factory	Construct with the common denominator that their core component is a Data Center.
Data Holder	A natural or legal person who is entitled or obliged under this Regulation, under applicable Union law or under national law implementing Union law to use and provide data - including product data or related service data, where contractually agreed - that it has accessed or generated during the provision of a related service” (Art. 2 para. 13 DA).
DATO	Scalable Digital and Automatic (up to Autonomous) Train Operation
e.V.	Registered association under German law (eingetragener Verein)
FRAND	Fair, reasonable, and non-discriminatory
InfraGo Data Factory	Sensors, data transmitters and the Data Center as core component.

GbR	Partnership under the Civil Code (Gesellschaft bürgerlichen Rechts)
gGmbH	Non-profit GmbH (gemeinnützige Gesellschaft mit beschränkter Haftung)
GmbH	Limited liability company (Gesellschaft mit beschränkter Haftung)
GoA4 rail operations	Highly automated rail operation up to fully automated, unaccompanied rail operations
GWB	German Competition Act (Gesetz gegen Wettbewerbsbeschränkungen)
InfraGO	DB InfraGO AG
JV	Joint Venture
NS	Dutch Railways
Open Data Set	Data set that is accessible to as many companies, research institutions and other interested parties as possible and allows all those interested parties to develop innovative solutions.
OSS licenses	Open-source licenses
Pan-European Data Factory	Pan-European pool of data from all Data Centers
Pan-European Entity	Central coordinating body of the Pan-European Data Factory that is independent of the partners, with a reliable, long-term organizational structure that serves to fulfill the jointly defined purpose.
Partner Organizations	Rail transport companies, rail infrastructure companies and industry partners
SaaS	Software-as-a-Service
SEP	Standard-essential patent
SCE mbH	European cooperative Society
SNCF	National Company of the French Railways
StGB	German Criminal Code (Strafgesetzbuch)
Related Service	A digital service, other than an electronic communications service, including software, which is connected with the product at the time of the purchase, rent or lease in such a way that its absence would prevent the connected product from performing one or more of its functions, or which is subsequently connected to the product by the manufacturer or a third party to add to,

	update or adapt the functions of the connected product (Art. 2 para. 6 DA).
RIC	Rail infrastructure companies
RTC	Rail transport companies
R2DATO	Rail to Digital and Automated Train Operations
TFEU	Treaty on the Functioning of the European Union
User	A natural or legal person that owns a connected product or to whom temporary rights to use that Connected product have been contractually transferred, or that receives related services (Art. 2 para. 12 DA).
WP7	Work Package 7

TABLE OF CONTENTS

Acknowledgements.....	2
Report Contributors	2
Executive Summary	4
Abbreviations and Acronyms.....	7
TABLE OF CONTENTS.....	10
List of Figures and Tables	11
1 Introduction	12
2 Legal Analysis	14
2.1 Setup for non-discriminatory access of stakeholders to data and computation infrastructure	14
2.1.1 Necessity of a central coordinating body	14
2.1.2 Merger control considerations	17
2.2 Essential Aspect of Data Ownership.....	19
2.2.1 Data ownership in the narrower sense.....	19
2.2.2 Data access claims.....	20
2.3 Publication of an open data set.....	28
2.3.1 Possible licensing types	28
2.3.2 Considerations to be observed in case of (partial) commercial licensing	29
4 Conclusions.....	30
References.....	32

LIST OF FIGURES AND TABLES

Figure 1 Ownership and access to the Pan-European Data Factory.....	13
---	----

1 INTRODUCTION

DB InfraGO AG (hereinafter "**InfraGO**") is participating in the "Rail to Digital and Automated Train Operations" (hereinafter "**R2DATO**") project funded by Europe's Rail Joint Undertaking (EU-RAIL or ERJU) together with various European rail transport companies (hereinafter "**RTC**"), rail infrastructure companies (hereinafter "**RIC**") and industry partners. To meet the increasing demand for transportation of both passengers and freight, R2DATO will take the advantages of digitalization and automation to develop the Next Generation Automatic Train Control (ATC) and deliver scalable Digital and Automatic (up to Autonomous) Train Operation (DATO) capabilities in order to enhance the capacity of the existing rail networks.

The development of highly automated rail operation up to fully automated, unaccompanied rail operations (hereinafter "**GoA4 rail operations**") including automated perception and incidence management requires the acquisition, storage, simulation and processing of a large amount of sensor data for ML training for AI use cases and validation purposes. This challenge is addressed by Work Package 7 (hereinafter "**WP7**") of the R2DATO project. It is setting the framework for a collaborative solution in the rail sector, the Pan-European Data Factory, to jointly collect, store and annotate sensor data and build up all required infrastructure for processing the data as a prototype.

In the context of the project R2DATO and for the purpose of this legal analysis, the term "**Pan-European Data Factory**" is to be understood as follows:

Across Europe, RTCs, RICs and industry partners may develop individual solutions to collect, store and annotate sensor data and build up all required infrastructure for processing the data. The storage, annotation and processing in each of those individual solutions may be carried out in one or more units which are referred to as "**Data Center**".

The Data Center, in turn, is the core component of the "**Data Factory**". InfraGO, as an example, generates data through a set of technical components including sensors and data transmitters. The data generated will be transmitted to and stored in a Data Center. The entire construct, i.e. sensors, data transmitters and the Data Center, constitutes the "InfraGo Data Factory". Other use cases for Data Factories are also conceivable. The common denominator, however, is that their core component is a Data Center.

InfraGO's Data Center is not intended to remain stand alone; rather, it may be connected with Data Centers of other European partner organizations such as RICs, RTCs or industry partners (hereinafter "**Partner Organizations**") to form a pan-European pool of data from all Data Centers (hereinafter "**Pan-European Data Factory**"). The primary goal is to enable access to the Data Center of all participants under predefined conditions. Moreover, it is intended that third parties will have access to these networked Data Centers in accordance with specific regulations and to create and make available one or more Open Data Sets with the help of the Pan-European Data Factory. As many companies, research institutions and other interested parties as possible should be able to access these Open Data Sets across borders. To this end, among other things, the Partner Organizations participate in the research projects "RailDataFactory" in the framework of the EU funding program "Connecting Europe Facility 2" (hereinafter "**CEF 2**") and R2DATO. As part of the "RailDataFactory" project, among other things, InfraGO, in collaboration with the National Company of the French Railways (hereinafter "**SNCF**") and Dutch Railways (hereinafter "**NS**"), conducted a "Study on a pan-European Data Factory for automated rail operation". The study focused, inter alia, on the scalable "backbone network" required for a pan-European Data Factory (i.e. the telecommunications infrastructure with very high data transmission rates that shall connect the individual Data Centers) and the data platform based on it.

To this end, InfraGO undertook to provide for an initial legal evaluation of the following non-technical aspects of the Pan-European Data Factory:

- (i) Develop a feasible setup for non-discriminatory access of stakeholders to data and computation infrastructure,
- (ii) address non-technical aspects that are related to the essential aspect of data ownership, and
- (iii) address non-technical aspects that are related to the publication of an “**Open Data Set**” (i.e. a data set that is accessible to as many companies, research institutions and other interested parties as possible and allows all those interested parties to develop innovative solutions), in particular analyze and complete a sector-wide data ownership model towards the release of the Open Data Set.

This paper summarizes a recommendation for the set-up of data connectivity, ownership and access to the Pan-European Data Factory.

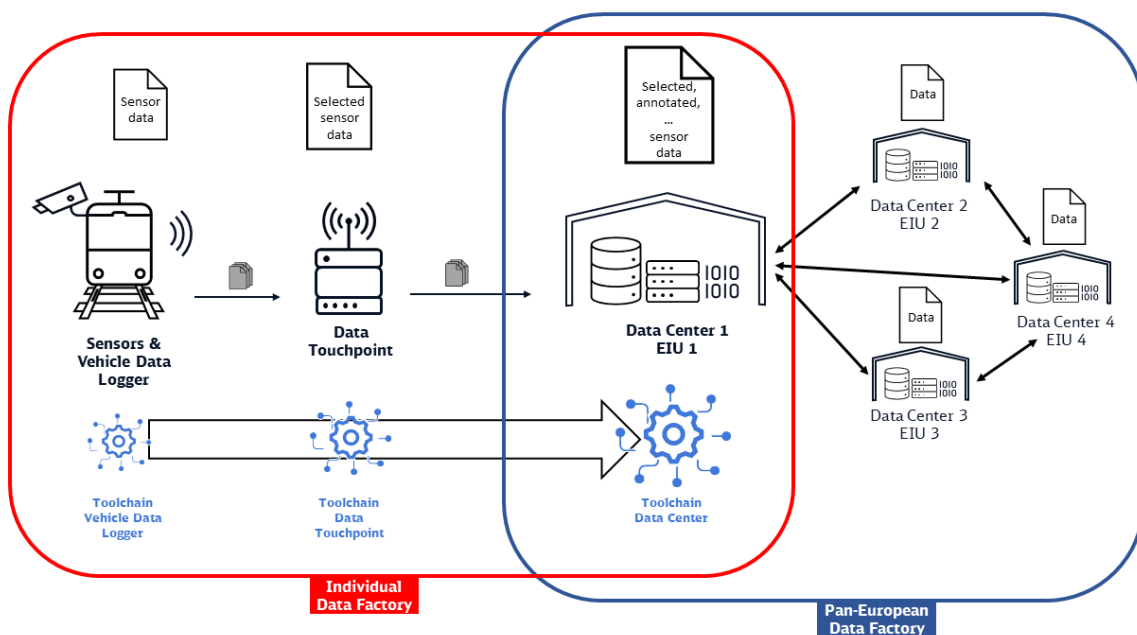


Figure 1 Ownership and access to the Pan-European Data Factory

2 LEGAL ANALYSIS

2.1 SETUP FOR NON-DISCRIMINATORY ACCESS OF STAKEHOLDERS TO DATA AND COMPUTATION INFRASTRUCTURE

2.1.1 Necessity of a central coordinating body

While InfraGO is able to control and take responsibility for the development, operation and use of the InfraGO Data Factory, the envisaged pan-European networking and coordination of several Data Centers inevitably involves Partner Organizations developing and operating Data Centers in different European countries. Their individual projects must be coordinated in a manner that ensures that the respective Data Centers are interoperable, i.e. have compatible data sets and interfaces. In addition, the networking of individual Data Centers must be organized via a backbone that must be jointly controlled and financed. It is difficult to imagine that these tasks can be reliably organized in a decentralized manner via bi-lateral contracts between more than two partners. Rather, it is to be assumed that the organization of a pan-European Data Factory requires a central coordinating body that is independent of the partners, with a reliable, long-term organizational structure that serves to fulfill the jointly defined purpose ("**Pan-European Entity**"). In the following analysis, we assume that the Pan-European Entity should only take on tasks that the Partner Organizations cannot handle on a decentralized basis. The organizational type to be chosen for this entity should be functionally oriented to the tasks at hand and require as little organizational effort as possible.

2.1.1.1 Tasks of the Pan-European Entity

The Pan-European Entity is responsible for the following tasks:

- **Establishment and operation of the "backbone"**: The Pan-European Entity is responsible for the establishment and operation of the network between the individual Data Centers. The network shall operate like a backbone which links the individual Data Centers to form a (virtual) Pan-European Data Factory and enables access to the data sets of all connected Data Centers for all Partner Organizations from any direction. This requires a transmission path network that enables data transfer with a very high data throughput rate. It is assumed that no separate network infrastructure needs to be set up or operated for this, but that suitable transmission capacities can be rented from specialized network operators. This commissioning would have to be prepared and coordinated by the Pan-European Entity, also with regard to the sharing of related costs and further contractual modalities. The contract with the service provider can, but does not necessarily have to, be awarded via the Pan-European Entity. It is also conceivable that the Partner Organizations award the contract jointly if this contributes to a more streamlined organization of the Pan-European Entity.
- **Definition of common standards and interfaces**: The Pan-European Entity will need to define common standards and interfaces and, where appropriate, conditions for the collection, processing (e.g. anonymization), provision and use of data to ensure that individual Data Centers are legally compliant in all relevant jurisdictions and that the Data Centers are connected and fully interoperable.
- **Monitoring compliance with standards and other agreements**: To ensure compatibility and compliance with legal requirements, such as data protection and data security, the Pan-European Entity must monitor compliance with the defined standards and agreements – if necessary, via a service provider – and, if necessary, check and certify the conformity of certain procedures before they are introduced and implemented.

- **Conclusion of contracts for access to the Pan-European Data Factory:** The Pan-European Entity will also have to adopt nondiscriminatory criteria according to which third parties are granted access to the Pan-European Data Factory. The Partner Organizations must then base their contracts with third parties on these criteria. In order to keep the Pan-European Entity as lean as possible, it should not act independently vis-à-vis third parties, i.e. it should not conclude any access contracts, etc.

As a result, the decentralized structure of the Pan-European Data Factory only requires the establishment of a coordinating and mediating body with a narrowly defined mandate, which should be manageable without significant human and material resources. It will not produce its own data and will not provide its own server capacity or transmission capacity but will merely act as a coordination vehicle for the operators of the individual Data Centers. Within the framework of the Pan-European Entity, those responsible for the individual Data Centers will meet in regular intervals to discuss the aforementioned issues and decide on a course of action. Implementation and action towards third parties will then be carried out by the operators of the individual Data Centers, i.e. the Partner Organizations.

2.1.1.2 Preferred legal form: registered association

A possible legal form for the Pan-European Entity is an organization under association law. The advantages of a registered association under German law (eingetragener Verein, hereinafter “**e.V.**”) or an equivalent under the law of another EU member state are, inter alia, their simple formation and accession of new members. The members decide on the election of the board or other committees, which can be set up flexibly. Given the manageable number of members envisaged here (Partner Organizations of other European countries), there is no danger of majorization by a large number of new members. Each member could be granted equal membership rights. Important decisions could require unanimity, all other decisions could be taken by majority vote. The association has its own legal personality, which excludes any recourse to the members under liability law. All other details could be based on the model articles of association known for existing pan-European entities such as Gaia-X.

Projects such as Gaia-X and Catena X, which, on the initiative of the EU Commission, form an organizational framework for secure and interoperable data traffic between European industrial and software companies, are also organized as associations. For Gaia-X the legal form of an association under Belgian law is suggested. The necessary documents are available online.¹ In discussions with the Partner Organizations, it makes sense to propose organizational structures that have proven themselves across borders and are used for similar projects.

According to German association law, the requirement of non-commerciality applicable to associations under Section 21 of the German Civil Code (Bürgerliches Gesetzbuch, hereinafter “**BGB**”) must be observed, which imposes a strict mandate on the activities they can undertake. A relevant case group that might exclude the formation as registered association under German law is an “association with entrepreneurial activity in a domestic market”. This involves systematic, remunerative activities offered in an internal market to members, where the association acts as a provider of services typically offered by others outside the membership relationship. If, however, in the case of the Pan-European Data Factory, the activities of the Pan-European Entity are limited to coordination and certification services, an internal market relationship is likely to be denied. This is because these services could exclusively be provided by the Pan-European Entity and there are logically no market alternatives. Additionally, remuneration (between the association and its members) seems avoidable if membership fees independent of services are sufficient to cover the association’s costs.

2.1.1.3 Advantages of a registered association compared to other legal forms

Other legal forms are not suitable for the intended purpose or are less appropriate for other reasons:

- Compared to the partnership under the Civil Code (Gesellschaft bürgerlichen Rechts, hereinafter “**GbR**”) or the general partnership (offene Handelsgesellschaft, hereinafter “**oHG**”), the e.V. with its own legal personality has the advantage of excluding direct liability. In contrast, partners in a GbR or oHG are liable without limitation and personally.
- A limited liability company (Gesellschaft mit beschränkter Haftung, hereinafter “**GmbH**”) offers the desired limitation of liability for partners but incurs a significantly higher administrative and financial burden compared to the e.V. because the admission of new shareholders, changes to the shareholder structure and articles of association always require the involvement of a notary. Given the currently envisaged lean mandate for the Pan-European Entity (coordination and certification), the effort involved is likely to be considered disproportionate. The variant of a non-profit GmbH (gemeinnützige Gesellschaft mit beschränkter Haftung, hereinafter “**gGmbH**”) is only available if the purpose and management of the company’s business meet the requirements set out in Section 52 German Tax Code (Abgabenordnung, hereinafter “**AO**”) (non-profit, charitable or ecclesiastical purposes), which is likely not the case here.
- The stock corporation (Aktiengesellschaft, hereinafter “**AG**”), which is similar to the e.V. in structure, also offers the desired limitation of liability and would also allow the admission of additional members without notarization. However, this legal form causes an even greater administrative effort in operation compared to the GmbH. In addition, the German Stock Corporation Act (Aktiengesetz, hereinafter “**AktG**”), with its principle of strict adherence to the articles of association, offers little flexibility to implement tailored governance solutions for the project.
- The European cooperative Society (“**SCE mbH**”) offers legal certainty throughout the EU thanks to its anchoring in EU law, and its use avoids being tied to the law of just one EU member state. Under tax law, the SCE and its German counterpart benefit from a favorable arm’s length advantage calculation, as more favorable prices may be charged to members in this context. However, it can be assumed that the requirements for the establishment of a SCE mbH are not met. A SCE mbH engages in an exchange relationship with its members and provides them with marketable services, which is unlikely to be the case for the Pan-European Entity planned here.
- If it is independent of the founders, a foundation (Stiftung) could be advantageous under merger control law and underline the claim of a neutral institution vis-à-vis third parties. However, setting up a foundation is more complex than establishing an e.V. In Germany, foundations are established by an act of state and require sufficient basic assets (at least EUR 100,000) and proof of a suitable foundation purpose. Under German law foundations are also subject to the principle of capital preservation, i.e. ongoing costs may not be financed from endowed capital. If, on the other hand, the e.V. is structured as a cooperative non-full-function company, merger control could also be avoidable. Thus, the foundation would not be more advantageous in this respect.

2.1.1.4 Organizational Design

All essential organizational aspects of the Pan-European Entity would be defined in an association statute when it is founded. These include the name, registered office, financial year and rules on the acquisition and termination of membership. The rights and obligations of the

members and the composition and responsibilities of the association's bodies must also be regulated. This must include a board of directors and, optionally, advisory boards and working groups. Governance and the passing of binding resolutions must also be regulated. The articles of association of Catena-X Automotive Network e.V. provide an illustrative example of how the articles of association could be structured.² Membership of the e.V. should primarily include the national owners of the Data Centers, namely the Partner Organizations. Additionally, other industry partners, particularly those with international operations and relevant expertise, may also be considered for membership under specific circumstances. Co-control or veto rights should only be held by the members themselves.

2.1.2 Merger control considerations

In the event of institutionalized cooperation between the parties involved at a pan-European level (for example within the framework of the proposed e.V.), merger control clearance may be required. This applies in particular if the Partner Organizations involved each have a turnover that exceeds the relevant thresholds. Until clearance is granted, a strict prohibition on implementation would then apply, which could possibly also delay certain preparatory acts.

2.1.2.1 Merger control at EU level

Under EU law only the acquisition of control is subject to notification. This includes both the acquisition of sole control and the acquisition of joint control (cf. Art. 3 Merger Regulation). In the event of cooperation between several companies in a newly established entity (Joint Venture, hereinafter “JV”) – as is the case with the Pan-European Data Factory – a notification is only triggered if the JV permanently fulfills all the functions of an independent economic entity (so-called full function criterion, cf Art. 3 para. 4 Merger Regulation). These two test criteria are likely not met:

- The establishment of the association constitutes an acquisition of joint control by the initially participating undertakings, if the governance of the association is structured in such a way that all shareholders (the parent companies) or, as the case may be, at least two must reach agreement on major decisions concerning the controlled undertaking (the joint venture).³ The clearest form of joint control exists where there are only two parent companies which share equally the voting rights in the joint venture.⁴ However, joint control may exist even where there is no equality between the two parent companies in votes or in representation in decision-making bodies or where there are more than two parent companies. This is the case where minority shareholders have additional rights which allow them to veto decisions which are essential for the strategic commercial behavior of the joint venture⁵ (such as appointments of senior management, business plan, market-specific rights, investments, etc.)⁶. If unanimity were required for such (strategic) decisions, this would constitute joint control. If, however, a simple majority (e.g. 2 out of 3 votes) were sufficient, changing majorities would be possible, meaning that there would be no joint control in principle. Thus, the governance of the association is decisive for the question of whether there is joint control or not.
- Further, the association may not meet the full function criterion.
 - To fulfill this characteristic, the JV must be active in a market and must perform the functions that are also performed by the other companies in this market. Therefore, the JV must have a management dedicated to the day-to-day business and sufficient resources such as financial resources, personnel, and tangible and intangible assets. If a JV only performs a certain (auxiliary) function in the business activities of the parent companies and does not have its own

market access or market presence, the full function criterion is not met. This is the case, for example, with JVs that are limited to research and development or production activities, or if a JV is essentially limited to the distribution or sale of the parent companies' products and therefore operates primarily as a sales agency.⁷

- Here the e.V. would not itself operate the Data Centers at a pan-European level. Rather, the e.V. would only be responsible for certain auxiliary functions that relate to the individual Data Centers of the Partner Organizations (such as certifying compliance with cybersecurity standards or the conditions for the collection, processing (e.g. anonymization), provision and use of data). In this respect, the establishment and organization of the backbone for networking the individual Data Centers would also represent only an auxiliary function for the Partner Organizations. The association will also – presumably with the exception of a small secretariat to organize coordination meetings – not have its own personnel capacities, but employees of the Partner Organizations will act jointly through it (for example, when it comes to organizing the “backbone”).

2.1.2.2 Merger control according to German Law

Under German law, in addition to the acquisition of control, the acquisition of shares or voting rights⁸ is also subject to merger control. The relevant threshold is the acquisition of a shareholding or voting right of at least 25%. Unlike EU law, German law does not have a full-function criterion (cf. Section 37 para. 1 no. 3 German Competition Act (Gesetz gegen Wettbewerbsbeschränkungen, hereinafter “**GWB**”)). It is relevant to note that the acquisition of shares or voting rights is relevant even in the case of an e.V.⁹

According to these standards, the formation of the association by up to four initially involved undertakings with equal shares will likely be regarded as relevant merger under German merger control law. It would then be decisive whether the turnover thresholds of German merger control are exceeded. In this respect, the full turnover of the RICs at group level would have to be considered (cf. Section 37 para. 1 no. 3 sentence 3 **GWB**).¹⁰ The provisions on merger control apply if, in the last financial year prior to the merger, the undertakings involved generated combined worldwide sales revenues of more than EUR 500 million and at least one participating undertaking generated sales revenues of more than EUR 50 million in Germany and another participating undertaking generated sales revenues of more than EUR 17.5 million in Germany (cf. Section 35 para. 1 **GWB**). If these turnover threshold are exceeded, there is likely to be an obligation to notify the merger in Germany.¹¹ An exception would only apply if – with an equal distribution of shares – at least five companies were initially involved or if the shares and voting rights of the initially involved companies were distributed in such a way that only one of these companies exceeded 25%.

2.2 ESSENTIAL ASPECT OF DATA OWNERSHIP

Data ownership can be approached from two perspectives: Firstly, in a narrower sense, applying the standards of property law, and secondly, and more relevantly, in terms of the obligated party of data access claims of third parties.

2.2.1 Data ownership in the narrower sense

There is no ownership within the meaning of German property law, specifically Section 903 of the German Civil Code (“**Bürgerliches Gesetzbuch**”, abbreviated “**BGB**”) of individual digital, intangible data, even if this data, such as sensor data collected by the trains, has an economic value. Data is not object within the meaning of Section 90 BGB, as it is not a tangible item.¹² Furthermore, data also lacks the rivalry condition characteristic for property, as it can be simultaneously used by multiple individuals and replicated with minimal effort without loss.¹³

Some literature suggests, with reference to Section 202a of the German Criminal Code (Strafgesetzbuch, hereinafter “**StGB**”) (spying on data) and Section 303a StGB (alteration of data), an analogous application of Section 903 BGB to construct data ownership.¹⁴ However, the position paper of the Max Planck Institute for Innovation and Competition currently sees neither a justification nor a necessity for creating exclusive rights to data.¹⁵ There is no principle according to which rights to data should be assigned to a specific legal entity from the outset.¹⁶ Furthermore, there are no apparent economic reasons for assigning exclusive rights to data. The associated encroachment on professional freedom and freedom of competition would rather entail the risk of hindering other market participants that rely on data, as well as negative influences on the development of downstream data markets.¹⁷ Moreover, due to the protection afforded by various legal domains, such as virtual domiciliary rights¹⁸, protection of database works and databases under copyright law¹⁹, competition law²⁰ and tort law via Section 823 para. 2 BGB in conjunction with protective laws (Section 202a StGB, Section 202b StGB, Section 202c StGB or Section 303a StGB)²¹, it cannot be assumed that there is an unintended regulatory gap that needs to be filled by creating a new category of ownership specifically for data.²²

However, it is important to distinguish the individual digital, intangible data from the embodiment of information in or on a data carrier. To the extent that electronic information is stored on a data carrier such as hard drives or disks by magnetizing their surface in a certain way, protection of the information can be mediated by the data carrier if it meets the criteria of a tangible item. Therefore, if the data carrier is damaged, destroyed, or altered in its nature, there is an infringement on tangible property. Consequently, modifying the magnetization of storage media by altering or deleting the information stored on these data carriers constitutes a violation of property rights.²³ However, this protection fails if the information is not stored on a data carrier belonging to the rightful owner but on foreign servers (i.e. in case of outsourcing or the use of cloud providers).²⁴

As the data is initially stored on the train itself, the owner or operator of the hardware and software used to collect and store the data would have de facto control over this data and would be the “quasi-owner” of the data without corresponding contractual provisions. Subsequently, data processing takes place autonomously in the Data Centers (as server and cloud infrastructure) and thus independently of the train. Consequently, with regard to processed sensor data, it would be crucial to determine ownership of the servers being used.

2.2.2 Data access claims

2.2.2.1 Data Act (Regulation (EU) 2023/2854)

Access rights may result from the Data Act if and when applicable (cf. Art. 50 Data Act (hereinafter “**DA**”)).

2.2.2.1.1 Data access for “users”

According to Art. 4 para. 1 of the DA, a user who cannot access the data directly from the connected product or related service has a claim against the data holder to be provided with the readily available data. In addition, the user or a party acting on behalf of a user may request that readily available data are made available to a third party (Art. 5 para. 1 DA). The DA therefore does not introduce the right to data access for random third parties.

- **Requirements of the claim**

The data access claim requires that data of a connected product or related service is involved. Pursuant to Art. 2 para. 5 DA a “connected product” means *“an item that obtains, generates or collects data concerning its use or environment and that is able to communicate product data via an electronic communications service, physical connection or on-device access, and whose primary function is not the storing, processing or transmission of data on behalf of any party other than the user”* (hereinafter **Connected Product**”).

Pursuant to Art. 2 para. 6 DA a “related service” means *“a digital service, other than an electronic communications service, including software, which is connected with the product at the time of the purchase, rent or lease in such a way that its absence would prevent the connected product from performing one or more of its functions, or which is subsequently connected to the product by the manufacturer or a third party to add to, update or adapt the functions of the connected product”* (hereinafter **Related Service**”).

The train that collects data with its sensors is a Connected Product within the meaning of Art. 2 para. 3 DA that generates relevant product data. Vehicles are explicitly mentioned in the Recitals as a use case and *“data automatically generated by sensors”* is explicitly mentioned as being within the scope of the Regulation.²⁵ The sensor data collected by the sensors on the train during the journey are therefore likely to be covered by the data access claim. The same applies to the associated metadata.

On the other hand, there are strong arguments suggesting that the processed data (such as the compressed sensor data, the selected sensor data (with annotations) or the artificial sensor data) is not covered by the data access claim.

- The right to access only applies to data generated by the Connected Product (here: the train) or Related Service linked to the Connected Product. According to the Recitals, this is *“data in raw form [...] that are automatically generated without any further form of processing, as well as data which have been pre-processed for the purpose of making them understandable and useable prior to subsequent processing and analysis [...]”*²⁶ Furthermore, according to the Recitals, the distinction between “pre-processed” data (which is subject to the access rights) and “inferred and derived” information from the data (which is not subject to the access rights) is crucial.²⁷ The “inferred or derived” information is *“the outcome of additional investments into assigning values or insights from the data, in particular by means*

*of proprietary, complex algorithms, including those that are a part of proprietary software [...].*²⁸

- Applying these principles, the processed and analyzed data (such as the selected sensor data (with annotations) or the artificial sensor data) is likely not subject to access rights under the DA because such data is precisely the result of additional investments, as significant expenditure is required for the review, annotation and other processing of the sensor data.
- The same is true for the compressed sensor data, as the analysis and compression of the sensor data is likely to result in an “analysis” of the sensor data and therefore relevant processing, which renders the compressed sensor data into “inferred and derived” information. Furthermore, there is no longer a relevant link to a Connected Product or Related Service, as the data processing takes place independently in the Data Center (as a server and cloud infrastructure) and therefore independently of the train as a Connected Product.²⁹

- **Claimant**

The “user” entitled to claim data access pursuant to Art. 4 para. 1 DA means “a *natural or legal person that owns a connected product or to whom temporary rights to use that Connected product have been contractually transferred, or that receives related services*” (cf. Art. 2 para. 12 DA) (hereinafter “**User**”). User is therefore first and foremost the owner of a Connected Product, but also anyone who possesses a Connected Product, for example as a tenant or lessee.³⁰

Users in this case are likely to be the RTCs that own and operate the train as a Connected Product. If there are other authorized parties (e.g. another owner), then they are also considered Users.³¹ Other parties involved in the individual Data Factories (especially industry partners and RICs), on the other hand, are typically not likely to be Users. The same is true for random third parties.

The User’s right to data output does not apply if the Connected Product is designed and manufactured in such a way that the User can access the data directly from the Connected Product (so-called “access by design”, cf. Art. 3 para. 1 DA).

- **Obligated party**

Under the DA, the obligated party of the data access claim, i.e. the “data holder”³², means “a *natural or legal person who is entitled or obliged under this Regulation, under applicable Union law or under national law implementing Union law to use and provide data - including product data or related service data, where contractually agreed - that it has accessed or generated during the provision of a related service*” (Art. 2 para. 13 DA) (hereinafter “**Data Holder**”).

In the context of the individual Data Centers, the industrial partners, i.e. the vehicle and equipment manufacturers, who connect digital services to the train are likely to be considered Data Holders with regard to the data generated by their connected products or related services, respectively. In particular, these will be the manufacturers or operators of software (including Software-as-a-Service (“**SaaS**”) services) that are used in the train to operate the sensors, provided that the industry partner has de facto access to this data (e.g. via remote access) and can therefore retrieve or generate this data within the meaning of the legal definition.

The Pan-European Entity will likely not be considered a Data Holder since it does not provide a digital (Related) Service. The networking and coordination of individual Data

Centers has no significance for the functionality of the sensors on the individual trains and does not change their functionality. Rather, it is a downstream step for the effective use of the collected and evaluated data.

- **Exclusion of the claim**

The right to data access is not granted without restrictions. Rather, the legislator has created a certain balance where business secrets are affected by the disclosure of data. The User is required to implement security measures to safeguard these business secrets prior to obtaining the data from the Data Holder (cf. Art. 4 para. 6 DA and Art. 5 para. 9 DA). In “exceptional circumstance”, however, the data access claim can be rejected if the disclosure of the data is highly likely to cause serious economic damage despite the security measures implemented by the User (cf. Art. 4 para. 8 DA or Art. 5 para. 11 DA). This is subject to a strict standard and must be sufficiently justified to the User based on objective elements, in particular the enforceability of trade secrets protection in third countries, the nature and level of confidentiality of the data requested, and the uniqueness and novelty of the Connected Product (cf. Art. 4 para. 8 DA or Art. 5 para. 11 DA).

Furthermore, the User may not use the product or service data obtained to develop a competing product to the Connected Product or to enable third parties to do so (Art. 4 para. 10 DA and Art. 6 para. 2 lit. I) DA). However, the development of competing products on downstream markets is permitted.³³ Additionally, the use of data to economically spy on the Data Holder, e.g. with regard to production methods, is prohibited (Art. 4 para. 10 DA and Art. 6 para. 2 lit. (e) DA). It is likely that in these scenarios the Data Holder is justified in refusing to make the data available to prevent misuse that contravenes legal prohibitions.³⁴

If users request that data is made available to a third party in accordance with Art. 5 para. 1 DA, the claim is excluded if the third party is a so-called gatekeeper within the meaning of the Digital Markets Act (Regulation (EU) 2022/192”, “DMA”) (cf. Art. 5 para. 3 DA).

Prototypes are excluded from the scope of the DA (see Recital 14).

- **Scope of the claim**

If a User asserts a data access claim against a Data Holder, the “readily available data” and the relevant “metadata necessary to interpret and use those data” must be provided (cf. Art. 4 para. 1 DA and Art 5 para. 1 DA). The readily available data is “*product data and related service data that a data holder lawfully obtains or can lawfully obtain from the connected product or related service, without disproportionate effort going beyond a simple operation*” (Art. 2 para. 17 DA). This does not include “*data generated by the use of a connected product where the design of the connected product does not provide for such data being stored or transmitted outside the component in which they are generated or the connected product as a whole*”.³⁵

- Product data is “data generated by the use of a connected product that the manufacturer designed to be retrievable, via an electronic communications service, physical connection or on-device access, by a user, data holder or a third party, including, where relevant, the manufacturer” (Art. 2 para. 15 DA).
- Related service data is “data representing the digitization of user actions or of events related to the connected product, recorded intentionally by the user or generated as a by-product of the user’s action during the provision of a related

service by the provider” (Art. 2 para. 16 DA).

- As described above only the sensor data and corresponding metadata will fall under the scope of the data access claim pursuant to Art. 4 para. 1 and Art. 5 para. 1 DA.

2.2.2.1.2 Data access for a public sector body

Art. 14 DA outlines the conditions under which public sector bodies, including the Commission, the European Central Bank or a Union body may have a right of access to data against the Data Holder. This access is contingent upon demonstrating an exceptional need, as set out in Article 15, to use the *“data, including the relevant metadata necessary to interpret and use those data, to carry out its statutory duties in the public interest”*.

- **Requirements for the claim**

An exceptional need exists

- if the “data requested is necessary to respond to a public emergency and the public sector body, the Commission, the European Central Bank or the Union body is unable to obtain such data by alternative means in a timely and effective manner under equivalent conditions” (Art. 15 para. 1 lit. (a) DA).
 - “Public emergency” is defined as *“an exceptional situation, limited in time, such as a public health emergency, an emergency resulting from natural disasters, a human-induced major disaster, including a major cybersecurity incident, negatively affecting the population of the Union or the whole or part of a Member State, with a risk of serious and lasting repercussions for living conditions or economic stability, financial stability, or the substantial and immediate degradation of economic assets in the Union or the relevant Member State and which is determined or officially declared in accordance with the relevant procedures under Union or national law”* (Art. 2 para. 29 DA).
 - and, insofar as non-personal data is concerned, if *“a public sector body, the Commission, the European Central Bank or a Union body is acting on the basis of Union or national law and has identified specific data, the lack of which prevents it from fulfilling a specific task carried out in the public interest, that has been explicitly provided for by law, such as the production of official statistics or the mitigation of or recovery from a public emergency; and [...] has exhausted all other means at its disposal to obtain such data, including purchase of non-personal data on the market by offering market rates, or by relying on existing obligations to make data available or the adoption of new legislative measures which could guarantee the timely availability of the data.”* (Art. 15 para. 1 lit. (b) DA).

Public sector bodies must adhere to certain formal requirements when requesting data access, as stipulated in Art. 17 DA. These include substantiating the existence of “exceptional need” and detailing any intentions to share the data with other public authorities (Art. 17 para. 4 and Art. 21 DA).

- **Obligated parties and scope of the claim**

The obligation to provide data applies to every Data Holder from whom the public sector body requests data, provided there is sufficient justification. For the Data Holders affected in the context of the Data Factories, see section 3, 3.2, 3.2.2.1.1.

Parallel to the scope of the data access claim of Users, the data, subject to release, includes not only the readily available data within the meaning of Art. 2 para. 17 DA but also the relevant metadata necessary to interpret those data (cf. Art. 14 DA).

Data Holders do not receive financial compensation for requests based on Art. 15 para. 1 lit. (a) DA, except for microenterprises and small enterprises (Art. 20 para. 1 DA). For requests under Art. 15 para. 1 lit. (b) DA, Data Holders generally receive financial compensation (cf. Art. 20 para. 2 DA), subject to exceptions outlined in Art. 20 para. 4 DA.

2.2.2.2 Claims arising from data access under antitrust law

Antitrust law claims for data access can arise from the antitrust abuse provisions (Art. 102 Treaty on the Functioning of the European Union (hereinafter “TFEU”) or Sections 19, 20 GWB). This would require a dominant or at least relatively strong market position and the refusal would have to constitute unfair or objectively unjustifiable discrimination. Such a right of access under antitrust law is likely to exist for third parties in return for a reasonable fee, especially if the data is used to train AI models or to carry out the approval of self-driving trains.

2.2.2.2.1 Dominant or relatively strong market position

In order to determine the market position, the relevant market must first be defined in product and geographic terms.

- **Market definition**

The purpose of defining the relevant product market is to determine the competitive forces to which the undertakings concerned are exposed.³⁶ This is based on the product or service offered and then, according to the so-called demand market concept, it must be examined whether the products or services offered by other suppliers are interchangeable from the point of view of the customers in terms of their characteristics, intended use or price level to cover a specific need.³⁷

- The relevant data from the Data Center includes the directly recorded sensor data and the processed data (such as the selected sensor data (with annotations) or the artificial sensor data). AI models are also to be developed from the data. The data and AI models ultimately serve to enable the technological and operational prerequisites for GoA4 rail operations.
- Looking at it from the perspective of the demanders, it is very likely that the relevant product market can only include other data or AI models that can also be used to enable the technological and operational requirements for GoA4 rail operations. This is because other data or AI models are not suitable from a demand perspective to fulfill the relevant intended use.

In geographical terms, there is much to suggest that the market covers the entire EU or EEA. This is because the data is likely to be tradable throughout Europe, without the national characteristics of the data being important. Rather, the aim is to generate a large amount of data in order to enable the conditions for GoA4 rail operations.

- **Market dominance or at least a relatively strong market position**

The individual Data Center is likely to have a dominant or at least relatively strong market position.

- A company is dominant if it is a supplier or buyer of a certain type of goods or commercial services on the relevant product and geographic market without competitors, is not

exposed to significant competition or has a superior market position in relation to its competitors (Section 18 para. 1 GWB).

- Relative market power exists if other companies, as suppliers or consumers of a certain type of goods or commercial services, are dependent in such a way that there are no sufficient and reasonable opportunities to switch to third-party companies and there is a clear imbalance to the countervailing power of the other companies (Section 20 para. 1 GWB). Such a dependency can also result from the fact that a company is dependent on access to data controlled by another company for its own activities (Section 20 para. 1a GWB).

By these standards - depending on how data offerings develop in the future - the providers of the Data Centers are likely to have at least a relatively strong (if not dominant) market position. This is because it seems unlikely that numerous providers will be active in the near future with corresponding data offerings or AI models that are suitable for enabling the prerequisites for GoA4 rail operations. This conclusion is supported by the fact that a large amount of data must first be collected from various sources in order to generate corresponding data sets and train AI models. Consequently, it can be assumed that the providers of the Data Centers, i.e. the Partner Organization operating them, are likely to be the standard addressees of the prohibition of abuse under antitrust law.

The dominant market position of the Pan-European Data Factory necessarily follows from the fact that the individual Data Centers themselves have a dominant market position, as the Pan-European Factory is a pooling of these. Therefore, third parties may also have a claim against the Partner Organizations operating the Data Centers for access to the Pan-European Data Factory. The Central Coordinating Body in the form of the registered association can, in principle, also be obligated to provide access. However, in view of its narrowly defined mandate, which is primarily aimed at coordinating the various Data Centers, efforts should be made to keep it as free as possible from such access claims.

2.2.2.2.2 *Unfair hindrance or objectively unjustifiable discrimination*

Market-dominant or relatively strong companies may not unfairly hinder other companies or treat them differently without objective reason.

A - hypothetically assumed - refusal of access to data is likely to constitute both a hindrance and unequal treatment.

- The term "impediment" is to be interpreted broadly and includes any impairment of the ability to operate in competition.³⁸
- There is also a high probability of unequal treatment. This is because it is planned that in principle the data from the Data Center will be accessible to any third party (e.g. in return for money, computer services or data), even if they do not operate their own Data Center. If every third party is therefore to be granted access to the data in principle, then the refusal of data access to certain third parties is very likely to constitute unequal treatment. In this respect, the similarity of the companies is to be assumed, especially since this characteristic is broadly interpreted³⁹ and no restrictions on the right of access for third parties are planned.

Whether an obstruction is unreasonable or whether there is an objective reason for discrimination is determined on the basis of a comprehensive weighing of the interests of the parties involved in the Data Center and Pan-European Data Factory on the one hand and third parties on the other), taking into account the objective of the German Federal Cartel Office (Bundeskartellamt, hereinafter "**BKartA**"), which is aimed at ensuring free competition and, in particular, the openness of market access.⁴⁰ Accordingly, there is much to be said in favor of the

Partner Organizations operating the Data Centers and collaborating within the framework of the Pan-European Data Factory making the data and AI models available to any third party for a reasonable fee:

- The starting point for the balancing of interests is, in principle, that the market dominator can also design its own distribution system and does not have to promote competitors.⁴¹
- However, this commendable interest reaches its limit where it runs counter to the objective of the BKartA, which is aimed at ensuring freedom of competition, and is directed in particular against the openness of market access or where there is a risk of a downstream market with its own share of value creation being dominated. Furthermore, a refusal to deal is considered unfair if the refused service is essential in order to operate on the derived market and the refusal creates barriers to market entry.⁴²
- As far as data access is concerned, the evaluations resulting from the creation of Section 19 para. 2 No. 4 GWB and Section 20 para. 1a GWB must also be taken into account when weighing up interests. According to the explanatory memorandum to Section 20 para. 1a GWB, an unreasonable impediment is conceivable if the data is to form the basis of significant own value creation by the access petitioner or if there is a threat of downstream markets being overtaken without access.⁴³
- Applying these principles, an overriding interest of third parties is likely to exist in any case if third parties want to use the data or AI models available in the Data Center or the Pan-European Data Factory to train AI models themselves or to carry out the approval of self-driving trains themselves. If the data is used in this way, significant added value is likely to be created. The data would also be important for market access for self-driving trains. Particularly in cases involving the further development of AI models and related applications, a claim to access is regularly desirable from a competitive perspective, especially as these can only be trained with specific and regularly very large data sets.⁴⁴ It should also be taken into account that AI is considered to be of outstanding importance for the future for numerous applications and developments.⁴⁵

Thus, in cases where third parties lack a mandatory legal entitlement to the data, it is probable that the Partner Organizations operating the individual Data Centers will still be obliged to grant access to the individual Data Centers or Pan-European Data Factory respectively under **FRAND** conditions, i.e. under fair, reasonable, and non-discriminatory terms, which are defined centrally by the Pan-European Entity.

2.2.2.2.3 Non-discriminatory access criteria

The concept of FRAND conditions originally stemmed from patent law. The patent holder, whose patent forms a standard-essential patent (hereinafter “**SEP**”) for anyone intending to use the standard, is typically required to offer licenses on FRAND terms.⁴⁶ In the EU, the application of FRAND conditions has been solidified by Court of Justice of the European Union (hereinafter “**CJEU**”) ruling in the case of Huawei v. ZTE^{47, 48}. In essence, the CJEU ruled that the refusal by the holder of a standard-essential patent to grant a license according to FRAND terms may constitute an abuse within the meaning of Article 102 TFEU.⁴⁹ In the Microsoft decision, the CJEU endorsed the Commission’s requirement for Microsoft to provide interoperability information on fair and non-discriminatory terms, enabling commercial users to develop and distribute operating systems for work group servers.⁵⁰

The purpose of the FRAND requirement is to establish licensing conditions which, on the one hand, result in adequate remuneration for the SEP holder for the use of his SEPs (thus rewarding the inventor), but which, on the other hand, also ensure that all market participants

have access to the components of standard-essential technologies protected by SEPs. This ensures broad utilization of the standard and prevents distortions of competition resulting from challenges in accessing technology.⁵¹

However, the terms fair, reasonable, and non-discriminatory are relatively vague and cannot be clearly defined. Guidance is provided by the new Art. 6 para. 12 sentence 1 of the Digital Markets Act, which mandates gatekeepers to apply fair, reasonable and non-discriminatory general conditions for commercial users' access to certain services. The framework defined in the DMA could serve as a compass for determining FRAND compliant conditions of access to the Data Centers or Pan-European Data Factory.

Regarding Art. 6 para. 12, Recital 62 specifies three manifestations of a lack of fairness:

- Imbalance between the rights and obligations of commercial users,
- Disproportionate advantage in favor of the gatekeeper in relation to the service provided by the gatekeeper and
- Discrimination against commercial users who offer similar goods or services as the gatekeeper (proximity to self-preferential treatment, Art. 6 para. 5).

The requirement that the license conditions must be "reasonable" establishes an outer limit. Reasonableness is defined by what can be demanded of the other party without disrupting the balance of interests. Unreasonable conditions are those that overburden the company concerned economically, constrict it, push it into legal gray areas or violate its sound entrepreneurial basis. In some cases, the lack of transparency of the terms and conditions may also constitute unreasonableness if they are so extensive, complicated or scattered that the package of rights and obligations is no longer comprehensible for commercial users.⁵²

Conditions are applied in a non-discriminatory manner if they apply equally to all similar users. Any unequal treatment must have a comprehensible objective justification, avoiding arbitrary decisions by the gatekeeper.⁵³

Consequently, in principle, access to the Data Center or Pan-European Data Factory must be granted for an appropriate fee. In this respect, significant costs incurred in connection with data generation, maintenance, storage and transmission,⁵⁴ such as setting up an interface, are eligible for consideration. The market value of the data, on the other hand, is generally not to be remunerated.⁵⁵ It should also be noted that no prohibitive access costs may be demanded, whereby it may have to be assumed that the market dominator is already sufficiently remunerated by its own data collection and use.⁵⁶ In this respect, there may also be cases in which access may be granted free of charge.

2.3 PUBLICATION OF AN OPEN DATA SET

One of the aims of the cooperation between the Partner Organizations with the help of the Pan-European Data Factory is to publish anonymized sensor data, annotations and further meta data as so-called Open Data Sets and thus make them available for the entire rail sector for various applications and users (possibly along with accompanying software). In view of the high costs and immense effort involved in generating this data, as many companies, research institutions and other interested parties as possible shall be able to access the data and develop innovative solutions.

To avoid the need for permanent access operation of the Data Centers or Pan-European Data Factory, it is currently assumed that the Open Data Set will not be offered as a SaaS solution, but will be provided to users as an “on-site” software solution (e.g. via download). The Open Data Set (possibly along with accompanying software) is then operated on the systems of the respective user. This requires the Partner Organizations operating the individual Data Centers to license the Open Data Set (and possibly accompanying software) to the users either jointly and severally or individually for the entire data and software pool. The conditions for licensing by the individual Partner Organizations are defined centrally by the Pan-European Entity.

2.3.1 Possible licensing types

Various commercialization options exist for licensing the Open Data Sets and accompanying software to users. Considering the respective licensing considerations outlined below a compromise in form of dual licensing is advisable.

2.3.1.1 Pure open-source license

In the case of a pure open-source license, commercialization of the product is generally not possible since usability of the product in line with the principle of open source presupposes free use. While this approach allows for collaborative (further) development of the Open Data Sets and accompanying software by the various licensees (which can reduce software maintenance costs), it also means that the Partner Organizations would not generate revenue through license fees.

However, it is not uncommon to offer additional services for the software, particularly support under a service level agreement, for a fee in addition to the basic product license. Offering such additional support services could present a commercialization opportunity for the Partner Organizations. However, as these support services would probably be provided by a technology partner of the Partner Organizations, the associated costs must be taken into account accordingly when calculating the remuneration.

2.3.1.2 Purely commercial license

On the other end of the spectrum, the highest form of control and monetization option is a commercial license to the user. In this scenario, imposing an additional confidentiality obligation on licensees regarding the Open Data Set and accompanying software can enhance its value as a trade secret of the Partner Organizations. However, this approach would contradict the idea of democratization.

2.3.1.3 Dual licensing

A compromise solution could be a so-called dual-licensing model, in which use for non-commercial purposes is possible free of charge under an open-source license (e.g. by non-profit organizations or university research institutions), while commercial use of the data and

software must be made under a commercial license subject to a fee.

2.3.2 Considerations to be observed in case of (partial) commercial licensing

If, in addition to the free provision of the Open Data Set and accompanying software through an open-source license, a commercialization via a dual licensing approach is sought, the rights of non-commercial users under the open-source license must nevertheless be restricted in a manner that enables parallel commercialization by the Partner Organizations. In particular, the users of the open-source version may not license or otherwise distribute the Open Data Set or accompanying software in any way that is not also restricted to non-commercial use.

The users of the Open Data Set and accompanying software should have the opportunity to further develop the provided data and accompanying software. However, such further developments should at least be made available for use through a license back to the Partner Organizations. Certain Open-source licenses (hereinafter “**OSS licenses**”) (so-called copyleft licenses) mandate that adaptations of the software code licensed under them must in turn be licensed under the same license and that the source code of the adaptations must be disclosed to third parties. However, these obligations only apply if the licensee distributes their developments, i.e. makes them accessible to third parties, not in the case of purely internal use by the licensee. The Partner Organizations can therefore not rely on this. In the case of a reverse license for proprietary further developments, i.e. which are not subject to an OSS license, an appropriate, customary remuneration for the reverse license must also be provided for reasons of copyright and tax law. This can be factored into the calculation of the license fee paid by the user to the Partner Organizations.

In the case of a commercial offer, it should also be considered whether the license is offered as a one-time license (purchase license) or as a time-limited right of use with ongoing license payments (rental license). Both are possible in principle, with a purchase license potentially being resold on the used software market. In case of the former, a one-off lump sum is paid. With the latter, different remuneration models are conceivable:

- A license fee could be charged per authorized user. This might entail either a number of concurrent users (regardless of the identity of individual employees) or named users.
- If several software copies are utilized (e.g. on servers at several company locations), a combination of a server license fee (for each software copy) and a license fee per user would also be conceivable. Monthly flat-rate packages could also be created, each comprising a certain number of users and servers, avoiding the need to register every new user as long as the package volume is not exceeded. Ensuring accurate billing of license fees would necessitate appropriate measures, such as the allocation of license keys for each software copy, the registration of authorized users and audit rights.
- Transaction-based, metered or pay-per-use licenses could also be envisaged. The license fee is then calculated according to how often a usage action (e.g. prompt input, generation of output, etc.) is performed or how much usage time is incurred. However, remuneration based on the number of accesses makes less sense for software used “on-site”, as it is more difficult to check the access figures than with a SaaS solution.
- Finally, a revenue-based license is also possible. In this case, the license fee is based on the user company reaching certain sales thresholds.

4 CONCLUSIONS

As part of WP7 of the R2DATO project, InfraGO has undertaken to provide a first legal assessment of the Pan-European Data Factory and to formulate recommendations for the establishment of data connectivity, ownership and access to the Pan-European Data Factory.

- I. In order to ensure non-discriminatory access of stakeholders to the data and computation infrastructure of the Pan-European Data Factory, it is essential to establish a central coordinating body that is independent of the individual Partner Organizations, with a reliable, long-term organizational structure that serves to fulfill the jointly defined purpose, the so-called Pan-European Entity. This entity will play a pivotal role in ensuring interoperability across the network of individual Data Centers across different European countries, thereby enhancing the overall effectiveness and efficiency of data management and access in the European rail sector. The tasks of the Pan-European Entity will include (1) the establishment and operation of the backbone network that connects the individual Data Centers, (2) the definition of and (3) monitoring compliance with common standards, interfaces and conditions for the collection, processing, provision and use of data to ensure legal compliance and interoperability, as well as the (4) adoption of nondiscriminatory criteria for third-party access to the Pan-European Data Factory. The decentralized structure of the Pan-European Data Factory should only require the establishment of a coordinating and mediating body with a narrowly defined mandate, which should be manageable without significant human and material resources. It will not produce its own data and will not provide its own server capacity or transmission capacity but will merely act as a coordination vehicle for the operators of the individual Data Centers.

The recommended legal form for the Pan-European Entity is a registered association. Such legal form provides a legal personality that protects the members from liability, allows for a simple incorporation and accession of new members, has low administrative burden and a flexible membership structure.

The establishment of the Pan-European Entity will not require merger control clearance under EU law by the EU Commission provided that it primarily serves auxiliary functions that relate to the individual Data Centers of the Partner Organizations as outlined above. In that case the Pan-European Entity is unlikely to perform all the functions of an independent economic entity and thus does not meet the so-called full-function criterion of the Merger Regulation.

However, the Pan-European Data Factory may be subject to notification obligations under national merger control regimes that do not provide for a full-function criterion, as is the case under German law. It is therefore recommended that a comprehensive merger control review be carried out at the Pan-European level, based on the revenues of the undertakings cooperating within the framework of the Pan-European entity as soon as they are known. This review should include all countries where these undertakings generate revenues. It was not possible to carry out such a comprehensive merger control review in all relevant jurisdictions at the time of this analysis since the participating undertakings in the Pan-European Entity are yet to be determined.

- II. The data ownership involved in the project may attract data access claims of third parties. Claims for data access may exist under antitrust laws. Depending on how data offerings develop in the future, the individual Data Centers will likely hold a relatively strong market

position (if not a dominant one) regarding the data offerings or AI models that are suitable for enabling the pre-requisites for GoA4 rail operations. The dominant market position of the Pan-European Data Factory necessarily follows from the fact that the individual Data Centers themselves have a dominant market position, as the Pan-European Factory is a pooling of these. Third-party access claims against the providers of the Data Centers, i.e. the Partner Organization operating them, are likely to be justified if these third parties wish to use the data or AI models available in the individual Data Center or Pan-European Data Factory to train their own AI models or independently carry out the certification of autonomous trains. Such access will have to be granted under FRAND conditions, i.e. under fair, reasonable, and non-discriminatory terms, which are defined centrally by the Pan-European Entity. Consequently, in principle, access to the individual Data Centers or Pan-European Data Factory must be granted for an appropriate fee. In this respect, significant costs incurred in connection with data generation, maintenance, storage and transmission, such as setting up an interface, are eligible for consideration.

Claims for data access by third parties under other regulations are unlikely. The Data Act specifically does not grant such data access claims for arbitrary third parties. Instead, the Data Act only applies to claims by “users”, who in this case are the rail transport companies that own and operate the train and whose data access is intended to be contractually regulated in any case. Additionally, the Data Act provides for a right of access by public sector bodies where they can demonstrate an exceptional need to use the data in order to carry out their statutory duties in the public interest. Such an exceptional need exists if the data requested is necessary to respond to a public emergency such as natural disasters. Furthermore, even if such a data access claim were to exist under the Data Act, it would only relate to the sensor data collected by the sensors on the train during the journey and the associated metadata, the processed data on the other hand, such as the compressed sensor data, the selected sensor data with annotations or the artificial sensor data, is likely not covered by the data access claim.

- III. Regarding the publication of the Open Data Set and accompanying software, it is recommended to adopt a dual licensing model. A dual licensing model would allow non-commercial use to be free of charge under an open-source license, while commercial use would require a commercial license subject to a fee. This approach balances the goal of democratization and the possibility for collaborative (further) development of the Open Data Sets with the need for control and potential for monetization by the Partner Organizations. Licensing of the Open Data Set and software will be managed by the Partner Organizations, either jointly or individually, on the basis of centrally defined conditions by the Pan-European Entity.

Furthermore, it is advisable that any further developments by users, particularly those under a commercial license, should include a license back to the Partner Organizations, potentially with appropriate remuneration. The commercial license can either be offered as a purchase or rental license. Assuming that the software is used “on-site”, instead of a SaaS solution to avoid permanent access to the individual Data Centers or the Pan-European Data Factory, the following compensation models can be considered in the case of a rental license: a fee per authorized user, a combination of a server license fee per software copy and separate fees per user, or monthly flat-rate packages covering a defined number of users and servers as well as a revenue-based license fee.

REFERENCES

- ¹ See <https://gaia-x.eu/what-is-gaia-x/about-gaia-x/>.
- ² *Catena_X_Articles_of_Association.pdf* (catena-x.net).
- ³ Commission Consolidated Jurisdictional Notice under Council Regulation (EC) No 139/2004 on the control of concentrations between undertakings, para. 63.
- ⁴ Commission Consolidated Jurisdictional Notice under Council Regulation (EC) No 139/2004 on the control of concentrations between undertakings, para. 64.
- ⁵ Commission Consolidated Jurisdictional Notice under Council Regulation (EC) No 139/2004 on the control of concentrations between undertakings, para. 65.
- ⁶ Commission Consolidated Jurisdictional Notice under Council Regulation (EC) No 139/2004 on the control of concentrations between undertakings, paras. 69-73.
- ⁷ Commission Consolidated Jurisdictional Notice under Council Regulation (EC) No 139/2004 on the control of concentrations between undertakings, paras. 94-95.
- ⁸ In addition, there is also the acquisition of significant competitive influence as a relevant concentration (cf. Section 37 para. 1 no. 4 GWB). However, a dominant influence of several companies does not lead to a partial merger being deemed to exist between them, as Section 37 para. 1 no. 3 sentence 3 GWB does not apply to this constellation. This likely rules out a notification requirement due to the non-fulfillment of the turnover thresholds pursuant to Section 35 para. 1 GWB.
- ⁹ *Bach*, Münchener Kommentar zum Wettbewerbsrecht, 4th ed. 2022, Section 37 para. 55; *Thomas*, Immenga/Mestmäcker Competition Law, 6th ed. 2020, Section 37 para. 188.
- ¹⁰ *Thomas*, Immenga/Mestmäcker Competition Law, 6th ed. 2020, Section 37 para. 223.
- ¹¹ Domestic effects are also likely to exist, especially since the parent companies are likely to be competitors and the market share of 20% is likely to be exceeded, see *BKartA*, Merkblatt Inlandsauswirkungen in der Fusionskontrolle, paras. 16, 21.
- ¹² Higher Regional Court Dresden in NJW-RR 2013, 27, 28; Regional Court Konstanz in NJW 1996, 2662; *Werner*, Arnold/Günther ArbR 4.0, § 5: Geistiges Eigentum und Wettbewerbsrecht im Arbeitsverhältnis para. 11; *Wagner*, MüKoBGB, 9th ed. 2024, BGB § 823 para. 285; *Grützmacher*, CR 2016 485, 489.
- ¹³ *Wagner*, in MüKoBGB, 9th ed. 2024, BGB § 823 para. 285; *Czychowski/Siesmayer*, in Tager/Pohle ComputerR-HdB, 20.5 Rechte an Daten paras. 19, 22.
- ¹⁴ *Zech*, CR 2015, 137, 143f.; *Hoeren*, MMR 2013, 486, 487ff.; *Wagner*, MüKoBGB, 9th ed. 2024, § 823 para. 381ff.
- ¹⁵ *Drex/Hilty/Desaunettes/Greiner/Kim/Richter/Surblyte/Wiedemann*, Ausschließlichkeits- und Zugangsrechte an Daten, Positionspapier des Max-Planck-Instituts für Innovation und Wettbewerb vom 16. 8. 2016 zur aktuellen europäischen Debatte, GRUR-Int. 2016, 914; *Werner* in Arnold/Günther ArbR 4.0, § 5: Geistiges Eigentum und Wettbewerbsrecht im Arbeitsverhältnis para. 12.
- ¹⁶ *Drex/Hilty/Desaunettes/Greiner/Kim/Richter/Surblyte/Wiedemann*, Ausschließlichkeits- und Zugangsrechte an Daten, Positionspapier des Max-Planck-Instituts für Innovation und Wettbewerb vom 16. 8. 2016 zur aktuellen europäischen Debatte, GRUR-Int. 2016, 914; *Werner* in Arnold/Günther ArbR 4.0, § 5: Geistiges Eigentum und Wettbewerbsrecht im Arbeitsverhältnis para. 12.
- ¹⁷ *Drex/Hilty/Desaunettes/Greiner/Kim/Richter/Surblyte/Wiedemann*, Ausschließlichkeits- und Zugangsrechte an Daten, Positionspapier des Max-Planck-Instituts für Innovation und Wettbewerb vom 16. 8. 2016 zur aktuellen europäischen Debatte, GRUR-Int. 2016, 914, 915.
- ¹⁸ *Werner* in Arnold/Günther ArbR 4.0, § 5: Geistiges Eigentum und Wettbewerbsrecht im Arbeitsverhältnis para. 14-19.
- ¹⁹ *Werner* in Arnold/Günther ArbR 4.0, § 5: Geistiges Eigentum und Wettbewerbsrecht im Arbeitsverhältnis para. 20-28.
- ²⁰ *Werner* in Arnold/Günther ArbR 4.0, § 5: Geistiges Eigentum und Wettbewerbsrecht im Arbeitsverhältnis para. 30-41.
- ²¹ *Werner* in Arnold/Günther ArbR 4.0, § 5: Geistiges Eigentum und Wettbewerbsrecht im Arbeitsverhältnis para. 42-47.
- ²² *Werner* in Arnold/Günther ArbR 4.0, § 5: Geistiges Eigentum und Wettbewerbsrecht im Arbeitsverhältnis para. 11; *Dorner*, CR 2014, 617, 621ff.

- ²³ Higher Regional Court Oldenburg ZD 2012, 177; Higher Regional Court Karlsruhe, NJW 1996, 200, 201; *Wagner*, MüKoBGB, 9th ed. 2024, BGB § 823 para. 286; *Grützmacher*, CR 2016, 485, 489.
- ²⁴ *Wagner*, MüKoBGB/Wagner, 9th ed. 2024, BGB § 823 para. 378; *Grützmacher* in CR 2016, 485, 489.
- ²⁵ cf. Recital 15 to the DA. The sensor could also be considered a Connected Product, although it is only an aid for recording the data. This question could become particularly relevant if, for example, the sensors are retrofitted to the train by the RTC.
- ²⁶ Cf. Recital 15 of the DA.
- ²⁷ Cf. Recital 15 of the DA.
- ²⁸ Cf. Recital 15 of the DA.
- ²⁹ See also Recital 16, according to which the DA should not apply to data “[...] *obtained, generated or accessed from the connected product, or which was transmitted to it, for the purpose of storage or other processing operations on behalf of other parties, who are not the user, such as may be the case with regard to servers or cloud infrastructure operated by their owners entirely on behalf of third parties, inter alia for use by an online service.*” See also *Assion/Willecke*, MMR 2023, 805, 806.
- ³⁰ The examples of tenants or lessees are mentioned in Recital 18.
- ³¹ See Recitals 18 and 21.
- ³² The obligation does not apply to “micro-enterprises” or a “small enterprise” or a “medium-sized enterprise” that has been classified for less than one year (Art. 7 para. 1 DA).
- ³³ See recital 30 of the DA.
- ³⁴ See also *Assion/Willecke*, MMR 2023, 805, 807.
- ³⁵ See Recital 20 of the DA.
- ³⁶ Federal Court of Justice, decision of 23 June 2020, KVR 69/19, para. 31.
- ³⁷ Federal Court of Justice, decision of 23 June 2020, KVR 69/19, para. 23.
- ³⁸ *Fuchs*, Immenga/Mestmäcker, Competition law, 7th ed., Section 19 para. 84.
- ³⁹ *Fuchs*, Immenga/Mestmäcker, Competition law, 7th ed., Section 19 para. 90.
- ⁴⁰ Federal Court of Justice, judgment of 24 October 2011, KZR 7/10, para. 37.
- ⁴¹ *Bechtold/Bosch*, Bechtold/Bosch, 10th ed. 2021, Competition law, Section 19 paras. 21, 24.; *Stancke*, Bunte/Stancke KartellR, Section 9 Missbrauchskontrolle, Sections 19–21 GWB, para. 88.
- ⁴² *Stancke*, Bunte/Stancke KartellR, Section 9 Missbrauchskontrolle, Sections 19–21 GWB, para. 91.
- ⁴³ RegBegr, BT-Drs. 19/23492, p. 81.
- ⁴⁴ See *Schweitzer / Haucap / Kerber / Welker*, Modernisierung der Missbrauchsaufsicht für marktmächtige Unternehmen, 2018, p. 136.
- ⁴⁵ *Podszun*, Immenga/Mestmäcker, Competition Law, 7th ed. 2024, Section 20a para. 139.
- ⁴⁶ *Schwab*, HK-DMA, 1st ed. 2023, DMA Art. 6 para. 12 para. 295.
- ⁴⁷ Judgement of 16 July 2015, *Huawei/ZTE*, C-170/13, ECLI:EU:C:2015:477.
- ⁴⁸ *Schwab*, HK-DMA, 1st ed. 2023, DMA Art. 6 para. 12 para. 295.
- ⁴⁹ Judgement of 16 July 2015, *Huawei/ZTE*, C-170/13, ECLI:EU:C:2015:477, para. 53.
- ⁵⁰ Judgement of 17 September 2007, *Microsoft*, T-201/04, ECLI:EU:T:2007:289, para. 193, 810.
- ⁵¹ Ann PatR, § 43 Standard Essential Patents (SEP) and their FRAND licensing para. 22.
- ⁵² *Schwab*, in *Podszun*, Digital Markets Act, 1st ed. 2023, DMA Art. 6 para. 12 para. 306.
- ⁵³ *Schwab*, in *Podszun*, Digital Markets Act, 1st ed. 2023, DMA Art. 6 para. 12 para. 307.
- ⁵⁴ BT-Drs. 19/25868, p. 118.
- ⁵⁵ *Podszun*, Immenga/Mestmäcker, Wettbewerbsrecht, 7th ed. 2024, Section 20a para. 147. However, the question of whether the market value is to be replaced in exceptional cases is disputed if it is a case of “forced commercialization”, i.e. if the data has not yet been made available to third parties. However, this is not relevant here because the data is generally to be made available to third parties.
- ⁵⁶ RegBegr, BT-Drs. 19/23492, p. 81.