Rail to Digital automated up to autonomous train operation

D7.2 – IP assessment and Data Ownership Model

Part 2 Partner Insights and Strategic Feedback on the Pan-European Data Factory Framework

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
06	18/12/2024	Answers to questions raised by SMO added

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

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

This document consolidates the comments, questions, and suggestions provided by R2DATO project partners on the *Pan-European Data Factory Legal and Regulatory Framework*. This feedback does not originate from Morrison & Foerster LLP, DB In-fraGO AG, DB AG, or DB Systel GmbH but rather represents the perspectives and operational needs of key project partners. It serves as a valuable reference for future development and refinement of the legal opinion and associated regulatory assessments, particularly regarding data ownership, access rights, licensing, and compliance with European directives.

Partner contributions in this document underscore the importance of:

1. **Data Ownership and Access**: Addressing complexities in data ownership and defining access rights that balance public and private interests.

2. **Licensing Models**: Advocating for dual licensing structures, including open-source licenses, to foster innovation.

3. **Technical and Operational Feasibility**: Addressing concerns about technical implementation and cybersecurity associated with the Pan-European Data Factory's centralized model.

4. **Regional Considerations**: Ensuring non-EU countries such as Switzerland and Norway are explicitly acknowledged within the framework, along with other member state-specific legal contexts.

The aggregated feedback here forms a foundation for continued dialogue within the R2DATO consortium. It is intended to support decision-making for future updates and the formulation of policies, ensuring that all stakeholder concerns are appropriately considered in the development of the pan-European Data Factory.

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

LIST OF FIGURES AND TABLES

1   INTRODUCTION

The purpose of this document is to present the consolidated feedback from R2DATO project partners regarding the *Pan-European Data Factory Legal and Regulatory Framework*. This feedback is essential for aligning the project's legal, technical, and operational components with the diverse needs and perspectives of its partners across Europe.

The Pan-European Data Factory is envisioned as a foundational structure to centralize data collection, sharing, and processing for fully automated and data-driven rail operations. However, given the complex regulatory landscape and the varying data ownership and access requirements among stakeholders, it is crucial to address partner-specific concerns and recommendations as documented here. Each contribution reflects distinct operational needs, regional regulatory priorities, and strategic considerations.

The feedback collected here serves a dual purpose:

1. **To Inform Ongoing Development**: By integrating partner input, this document ensures that the Data Factory's regulatory framework remains relevant, compliant, and effective across jurisdictions.

2. **To Facilitate Structured Dialogue**: The document functions as a resource for future consultations within the consortium, enabling a structured approach to refining the Data Factory's legal and regulatory aspects.

This document is structured as follows:

- A comprehensive summary of feedback from each partner is provided, detailing specific comments on key areas such as data ownership, access regulations, licensing, and technical feasibility.

- Key recommendations for incorporating this feedback into the main deliverable are outlined in the concluding section, which emphasizes the next steps for collaborative refinement.

2   R2DATO PARTNER FEEDBACK

The feedback summarized in this part exclusively reflects the opinions and suggestions of the R2DATO project partners. It does not originate from Morrison & Foerster LLP, DB InfraGO AG, DB AG, or DB Systel GmbH. The comments below provide specific insights into partner perspectives on data ownership, licensing, regional inclusion, technical feasibility, and operational impact. These inputs contribute to a broader understanding of the collaborative needs and expectations within the R2DATO consortium for the Pan-European Data Factory framework.

The table below outlines partner-specific feedback, requests for clarification, and strategic recommendations across critical areas such as data ownership, data access regulations, technical requirements, and governance implications. The feedback provided will be systematically reviewed to refine the existing framework, enhance inclusivity for European and non-EU partners, and ensure that legal provisions and data access rights meet the operational and technical requirements of all stakeholders.

Partner	Comments	Reference to abstract
SBB (Swiss Federal Railways)	- General assessment: Welcomes the legal opinion but has several remarks and suggestions. - Data Ownership and Access: Requests clearer formulations regarding data ownership, specifically the rights of data users and providers, highlighting the need to balance public and private interests. - Licensing: Suggests clearer definitions of licensing terms and supports dual licensing, with an emphasis on open-source licenses to foster innovation. - Highlighting Switzerland: Requests explicit mention of Switzerland as a partner, given that the text predominantly refers to EU countries. Notes that this might also apply to Norway.	Data ownership and access (pp. 19-22), Licensing (pp. 28-29)
SNCF (French National Railways)	- General assessment: Sees the document as a starting point reflecting the German perspective (mainly Deutsche Bahn's). - Legal Bindingness: Clarifies that their review should not be interpreted as consent or a legal basis for the network. Requests the addition of a paragraph stating that the opinion is an open opportunity and should not be considered a collective position of the consortium for future work. - Future Discussions: Requests that the document be treated as a foundation for future	-

	discussions, rather than a legally binding document.	
Chalmers University of Technology	- General assessment: Provides constructive feedback focused on scientific validity and methodology. - Legal Aspects of Data Access: Stresses the importance of clear regulations on data access, especially concerning European directives and their implementation across member states. Suggests more detailed exploration of the Data Act's relevance to the project. - Future Research: Recommends addressing potential future research fields and evolving legal frameworks. - Additional feedback: Requests a clearer document structure and more practical examples of legal analysis in various EU member states.	-
NS (Dutch Railways)	- General assessment: Considers the document a good starting point but notes significant gaps. - Technical Feasibility: Criticizes the insufficient discussion on technical implementation and operational feasibility of the Pan-European Data Factory. - Focus on German Law and EU Data Act: Suggests that the legal opinion is too focused on German law and the EU Data Act, overlooking other relevant laws in other EU member states. Requests a more comprehensive analysis of these aspects. - Specific Dutch Law Requirements: Requires more specific references for Dutch law in relation to NS data. - Grant Agreement Articles: Misses references to Articles 8 and 9 of the Grant Agreement regarding data sharing and ownership among R2DATO/ERJU participants. - Licensing: Requests suggestions on how to organize licensing among consortium members.	Data Access Regulations (pp. 19-22)
SMO (Siemens Mobility)	• Did you also analyze the approach of a decentralized data ecosystem consisting of self-organizing data spaces connected via standardized interfaces?	

	- Many market participants in the rail industry have running data spaces in place. Some of them are already connected via APIs. Why do we need an additional Pan-European Data Factory as a central entity to manage access? - The central Pan-European Data Factory might adversely affect current digital business models of several market participants. What may be the motivation to support this approach? - The establishment of a central Pan-European Data Factory harbors the risk of creating a rail data monopolist selling data. Which measures prevent this? - The topic of a Pan-European Data Factory will impact all European market participants. How can we integrate the opinion/feedback of smaller market participants who are not involved in our working group? - Will there be an obligation for market players in the rail industry to participate in the Pan-European Data Factory framework? Is there a disadvantage for those who won't participate? - Which data is in the scope of the Pan-European Data Factory (e.g., raw data, metadata, processed data, personal data, operational data)? - If all European rail data, including security-critical data, is managed by one central Data Factory, this could be a target for hostile cyber-attacks. Which measures prevent this? - Data sovereignty: How will the ownership rights of the data generated by participating companies be adequately protected and safeguarded within the Pan-European Data Factory? - The processing and provision of data sets by the individual data factories	

	costs money. How will this be financed? • Does the Pan-European Data Factory earn money, or how is it financed? What detailed cost-benefit analyses have been conducted to evaluate the added value of the Pan-European Data Factory for participating companies? • According to the EU Data Act, data can only be used and shared by the data holder based on contractual agreements. Why should customers agree on providing their data to a Pan-European Data Factory which is centrally coordinating the access to the data? • Chapter 2.1 refers to the EU Data Act. However, there is no relation between the Pan-European Data Factory and the EU Data Act identifiable. Why are five pages referring to the EU Data Act? • ☐ How does this approach differentiate from Railway-X, Mobility Data Space, and European Rail Data Space?	

Table 1 Partner Feedback Summary on Legal Opinion and Data Ownership Model

2.1 ANSWERS TO QUESTIONS RAISED BY SMO

Importance of Differentiating Key Concepts

Differentiating the **Pan-European Entity** and the **Pan-European Data Factory** ensures clear separation of **organizational** and **technical responsibilities**. The Entity focuses on **coordination** and **standardization**, while the Data Factory handles **data processing** and **AI development**, enabling efficient collaboration and goal alignment.

Term	Pan-European Entity	Pan-European Data Factory (PEDF)
Type	Organizational and administrative body	Technical and data-driven platform
Focus	Coordination, standardization, and interoperability	Data collection, processing, and AI development
Role	Acts as a mediator and standard-setting authority	Comprises distributed data centers and platforms
Data Generation	Does not generate or process data	Collects, processes, and provides data
Structure	Independent entity (e.g., association)	Network of interconnected data factories

Figure 1 Differentiating Key Concepts

2.1.1 Did you also analyze the approach of a decentralized data ecosystem consisting of self-organizing data spaces connected via standardized interfaces?

Yes, the approach of a decentralized data ecosystem was analysed. The **Pan-European Data Factory** is based on the technical idea of a **decentralized ecosystem**, but also enables the partial centralization of **IT assets**, such as a **backbone network**, which serves as the technical basis for connecting the various **data spaces**. Even in case of a fully decentralized approach, a **central coordinating body (the Pan-European Entity)** remains crucial to ensure **interoperability**, **security standards** and **standardized interfaces**. Without this central coordination, there is a risk that different **data rooms** will form fragmented systems that cannot communicate efficiently with each other.

The **Pan-European Entity** ensures that **standards** are enforced, and **data flows** are smoothly integrated. This combines the advantages of **decentralization** with the necessary central mechanisms for **consistency** and **efficiency**.

2.1.2 Many market participants in the rail industry have running data spaces in place. Some of them are already connected via APIs. Why do we need an additional Pan-European Data Factory as a central entity to manage access?

Although existing **data spaces** are partially connected via Application Programming Interfaces (**APIs**), the main motivation of the **Pan-European Data Factory** is **large-scale standardization** and coordination across all **stakeholders**. Existing systems, if not properly integrated, could lead to **data islands** that are not compatible with each other. The **Pan-European Data Factory** provides an overarching **infrastructure** that

ensures **data rooms** work together seamlessly and use **standardized interfaces**. This prevents **fragmentation** and ensures that all market participants, regardless of their size or technical capacity, have equal access to **high-quality**, **interoperable data**. The **Pan-European Data Factory** could also offer **central components**, such as the **backbone network**, which ensures efficient connection of all stakeholders. In addition, by involving **external users**, **data** and **tools** can also be made available beyond the circle of members, which creates the advantage of broader data availability.

2.1.3 The central Pan-European Data Factory might adversely affect current digital business models of several market participants. What may the motivation to support this approach?

The market participants' existing **infrastructures** and **business models** are not replaced, but integrated. This integration enables companies to continue using their own systems while at the same time accessing a much larger volume of **interoperable data**. **Uniform data formats** and **high-quality standards** are ensured. This creates added value by making **data** available on a **larger scale**, which is particularly beneficial for **AI developments** and **automation** in the rail sector. Companies benefit from **economies of scale** and have access to data that they could not access on their own without having to fundamentally change their own **business models**.

2.1.4 The establishment of a central Pan-European Data Factory harbors the risk of creating a rail data monopolist selling data. Which measures prevent this?

No, the risk of a **data monopoly** is avoided by the design of the **Pan-European Data Factory and the Pan-European Entity**. The Pan-European Entity acts as a **coordinator**, not as a **data owner**. It ensures that data is accessible on the basis of **non-discriminatory rules**. To avoid monopoly risks, **competition control mechanisms** are obeyed at **EU and national levels.**Access to central **IT assets**, such as the **backbone network**, remains open and standardized without creating a dominant position. There should be standardized model contracts that regulate the contribution and use of data. This will ensure that the Pan-European Data Factory is available to all interested companies on equal and fair terms. To maximize transparency, we could consider making the standardized model contracts publicly available.

2.1.5 The topic of a Pan-European Data Factory will impact all European market participants. How can we integrate the opinion/feedback of smaller market participants who are not involved in our working group?

The framework of the **Pan-European Data Factory** guarantees **non-discriminatory access** for all **market participants**, regardless of their size. Market participants can use and contribute **data** and **services** both as members and as external users. This is achieved through **open data initiatives** and the creation of **low-threshold technical**

access. **Cost-effective options** are being developed for smaller players to enable them to participate without having extensive **IT resources**. The use of **standard interfaces** reduces the effort required for integration, which also makes access practicable for smaller market participants.

2.1.6 Will there be an obligation for market players in the rail industry to participate in the Pan-European Data Factory framework? Is there a disadvantage for those who won't participate?

Participation in the **Pan-European Data Factory** is not mandatory. Even if there is no legal obligation, the platform should be designed in such an attractive way that the added value of participation motivates companies to join, as a member or external user.

2.1.7 Which data is in scope of the Pan-European Data Factory (e.g. raw data, meta data, processed data, personal data, operational data)?

The scope of the **Pan-European Data Factory** covers a variety of **data types**, including **multimodal sensor data**, **metadata**, **processed data**, and **AI/ML models**. When integrating **AI models**, clear **guidelines** and **standards** will be implemented to ensure that these models are **interoperable** and can be used in different environments without being tied to specific technical platforms. In addition, **legal** and **technical measures** are introduced to ensure that **sensitive data** such as **personal data** is handled in accordance with the applicable **regulations**.

2.1.8 If all European rail data, including security-critical data, is managed by one central Data Factory, this could be a target for hostile cyber-attacks. Which measures prevent this?

The **Pan-European Entity** is primarily designed as a **central coordination point** and sets and monitors security standards, The technical **IT components** of the **Pan-European Data Factory** remain decentralized at the participating companies. Each organization retains responsibility for the **cybersecurity** of its own systems and data. However, the Pan-European Data Factory will also enable **hybrid approaches** in which certain central **IT assets** can be brought in. An example of this would be a **centralized backbone network** that serves as the backbone for connecting the various **decentralized data rooms**. All components would also be subject **cybersecurity guidelines** and **protective measures**. To ensure the security of the entire infrastructure, **common security standards** are developed and ensured through regular **audits**, **certification procedures**, and **compliance monitoring**. The combination of **decentralized** and **centralized components** ensures that the benefits of decentralization are maintained, while centralized assets are deployed in a targeted manner to ensure efficient data flows and **robustness**. This reduces vulnerability to **cyber threats** and enables a **scalable** and **secure data infrastructure**.

2.1.9 Data sovereignty: How will the ownership rights of the data generated by participating companies be adequately protected and safeguarded within the Pan-European Data Factory?

Data sovereignty is guaranteed by a clear **legal framework. There should be standardized model contracts that regulate the contribution and use of data.** These contracts should of course be in line with the Data Act, in particular Chapter IV.

2.1.10 The processing and provision of data sets by the individual data factories costs money. How will this be financed?

The **funding mechanisms** of the **Pan-European Data Factory** are still being developed. Namely, a precise distribution of revenue and a full **cost-benefit analysis** are currently carried out. Several funding options are conceivable.

First of all, companies that are members of the Pan-European Data Factory could make mutual **contributions on a cooperative basis** and use the **resources** as part of their **membership**.

One possible source of revenue is that **data** (raw data, metadata, models, datasets, etc.), **tools** (specific data pipelines, software frameworks, etc.), and **computing capacity** (storage, computing power) will be offered through the Pan-European Data Factory. Companies that are members of the Factory could make use of additional services for a **fee**.

The payment of membership fees is also conceivable, whereby this income could be reinvested in the **maintenance**, further development, and expansion of the **IT infrastructure** and the coordination office of the Pan-European Data Factory.

In addition, the Pan-European Data Factory could also include **external users** who are not members. External users would have access to certain **datasets** and **tools** for a **fee** but would not be able to fully benefit from the advantages available to members, such as **co-determination rights** or privileged access to certain internal data sources. This option would allow non-members to contribute to the funding of the infrastructure while expanding access to data, which could be particularly attractive for **research institutions**, **start-ups**, or companies with limited resources.

Public funding could also be considered for financing, especially for **research and development (R&D)** activities.
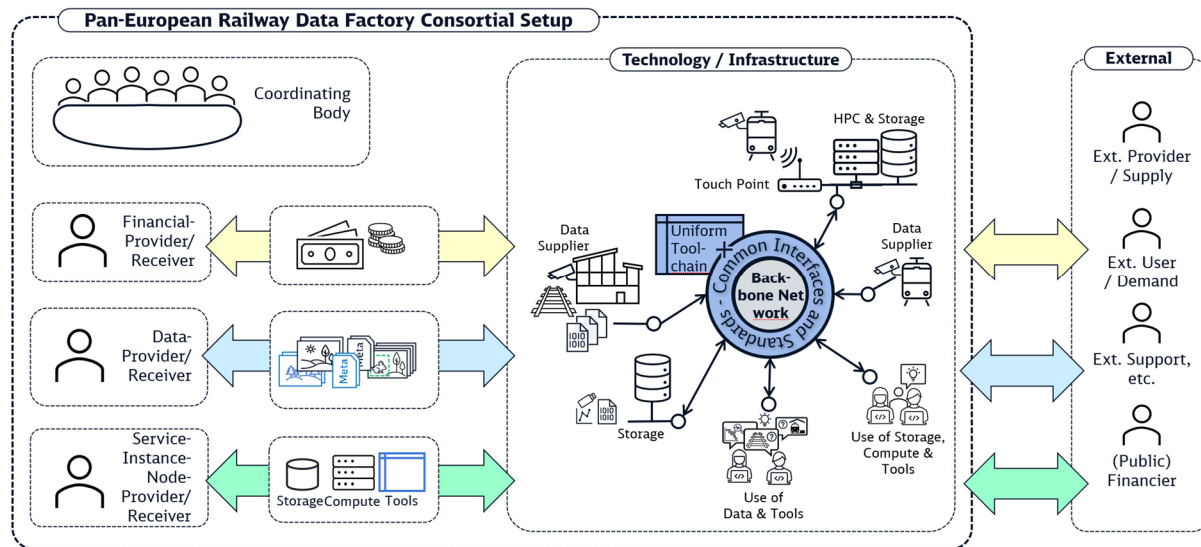
Figure 2: Possible setup

2.1.11 Does the Pan-European Data Factory earn money or how is it financed? What detailed cost-benefit analyses have been conducted to evaluate the added value of the Pan-European Data Factory for participating companies?

Various **financing models** are conceivable. The **Pan-European Data Factory** could generate revenue through **membership fees**, **access costs** for external users, or (**dual-)licensed models** in which commercial users pay for access to certain data. At the same time, **public funding** and **R&D activities** could support the financing. The **cost-benefit analysis** aims to provide added value for both **members** and **external users** through access to shared data resources.

2.1.12 According to the EU Data Act, data can only be used and shared by the data holder based on contractual agreements. Why should customers agree on providing their data to a Pan-European Data Factory which is centrally coordinating the access to the data?

Contractual agreements are a component of the **Data Act**. However, it stipulates claims beyond **contractual agreements**. At the same time, **contractual agreements** can be made beyond the scope of the **Data Act**. **Data** could be provided by members of the **Pan-European Data Factory** or **external users**. The recipient would not be the **Pan-European Data Factory** or the **Pan-European Entity** but members or **external users**. **Access** would be based on **contractual agreements** directly between the **provider** and the **recipient**. **External users** would grant **access** to their **data** in exchange for **access** to the **Pan-European Data Factory**. If they do not wish to share their **data**, they could be asked to pay a **fee** instead.

2.1.13 Chapter 2.1 refers to the EU Data Act. However, there is no relation between the Pan-European Data Factory and the EU Data Act identifiable, so why are five pages referring to the EU Data Act?

Chapter 2.1 addresses the aspect of **data ownership** and possible **access claims** arising therefrom. The **Data Act** enables **data access** and **sharing** across the **EU**. It defines clear rules for **data usage**, **access**, and **transfer**, in particular claims of so-called **users** against so-called **data holders**. It was essential to evaluate whether any **third party** might have mandatory **statutory claims** against the **Pan-European Entity** or any **Partner Organizations** participating in the **Pan-European Data Factory**, in addition to the claims resulting from the **contractual framework** set out above. The **Data Act** does not grant any such claims to **data access** for any **third parties**. Rather, the **Data Act** only applies to claims of **"users"** (e.g., **RTCs**), whose **data access** is to be **contractually agreed** anyway. While such claims can very likely not be derived from the **Data Act**, they might result from **antitrust law.**

2.1.14 How does this approach differentiate from Railway-X, Mobility Data Space, and European Rail Data Space?

The **Railway-X** and **RDS** approaches overlap with the **Pan-European Data Factory** concept, while the latter specializes in the rail sector, targeting **autonomous train operations** and **large-scale data integration**. Its key goals are to **standardize** and share **multimodal sensor data** with rail context to train **AI models** and support **fully automated train control** across European rail networks.
The concept is a hybrid of **decentralized IT assets** and potentially centralized assets like a **Backbone Network**. The concept provides a common tool chain as well as co-ordinated **data standards**, **formats**, and **interfaces**. The concept of **decentralized data ecosystems** and **federated data management** is supported and ensures **interoperability**, **data standardization**, and **fair access** for all rail **participants**.
By comparison:

1. **Railway-X**: Railway-X is a more **decentralized initiative** focusing on facilitating **API-based data exchanges** between existing data spaces. In contrast, the **Pan-European Data Factory** involves a **central coordinating body** to manage access and ensure interoperability across different systems.

2. **Mobility Data Spaces**: While both the **Pan-European Data Factory** and **Mobility Data Spaces** promote **data sharing** within the transport sector, Mobility Data Spaces are part of a **broader initiative** that includes various **transportation modes**, whereas the **Pan-European Data Factory** is specifically focused on **railway data**.

3. **European Rail Data Spaces**: This initiative is **similar** to the **Pan-European Data Factory** but operates as a **distributed system** without a central coordinating entity. The **Pan-European Data Factory** introduces **central**

coordination to prevent fragmentation and standardize access across the **European rail industry**, offering a more structured approach to **managing data** and **integrating systems**.

Railway-X and Mobility Data Space focus on **decentralized data ecosystems**, the **European Rail Data Space** includes the **context of multimodal data**. In summary, the **Pan-European Data Factory** shares some goals but also combines some of the strategies of the aforementioned projects, adding a **central coordination** approach to enable **AI-based solutions** for **autonomous train operations**.

3 CONCLUSIONS

In conclusion, this document provides a comprehensive synthesis of partner feedback on the R2DATO *Pan-European Data Factory Legal and Regulatory Framework*. The feedback highlights critical themes that require attention to ensure the framework's viability, technical feasibility, and alignment with stakeholder needs across various legal, technical, and operational domains.

Key takeaways include:

1. **Data Ownership and Access Rights**:
   Partners emphasize the need for clear and balanced definitions of data ownership and access rights that reflect both public and private interests. Standardized contractual agreements, aligned with the EU Data Act, were identified as essential to support cross-border data sharing while safeguarding data sovereignty.

2. **Technical Feasibility and Implementation**:
   Concerns were raised about the operational feasibility of a centralized data factory, particularly regarding its potential overlap with existing decentralized data spaces. In response, the framework adopts a **hybrid approach**: combining centralized coordination (to ensure interoperability and security standards) with decentralized IT assets managed at the participant level. This mitigates risks of fragmentation while leveraging existing infrastructure.

3. **Cybersecurity and Data Sovereignty**:
   The centralization of certain data components raised concerns about vulnerability to cyber-attacks. The proposed measures include robust security standards, regular audits, and a distributed infrastructure to ensure resilience. Data sovereignty is safeguarded through clear legal frameworks and standardized model contracts to protect ownership rights.

4. **Licensing Models and Regional Inclusion**:
   Dual licensing strategies, including open-source options, are supported to foster innovation and equitable access. Feedback also underscores the importance of explicitly including non-EU partners, such as Switzerland and Norway, in the legal and operational framework.

5. **Financing and Added Value**:
   Financing mechanisms for the Pan-European Data Factory are under development, with multiple options explored, including membership models, access fees for external users, and public funding. A detailed cost-benefit analysis is ongoing to demonstrate the added value of the platform for all participants, ensuring financial sustainability.

6. **Distinction from Other Data Initiatives**:
   Specific differentiation from **Railway-X**, **Mobility Data Spaces**, and **European Rail Data Spaces** was clarified. While sharing some goals, the Pan-European Data Factory uniquely introduces a central coordinating entity to ensure

interoperability, AI model training, and the integration of multimodal data specific to autonomous rail operations.

In conclusion, the insights and responses provided in this document reflect a collaborative effort to address key legal, technical, and operational challenges associated with the Pan-European Data Factory. By building a framework that balances innovation, inclusivity, and compliance, the R2DATO consortium aims to establish a robust foundation for the future of automated rail operations in Europe. This effort underscores the commitment to fostering interoperability, safeguarding data sovereignty, and enabling sustainable growth across the rail sector.