Railway Marketplace for Data, Know-How and Services

Stefan H. Reiterer[1] Alexander Meierhofer[1] Ivan Vidovic[2] Marco Forberger[2] Benjamin Stuntner[3] Jochen Nowotny[3]

[1]Virtual Vehicle Research, Austria, {stefan.reiterer,alexander.meierhofer}@v2c2.at
[2]voestalpine Digital Track Management GmbH, Austria, ivan.vidovic@voestalpine.com, m.forberger@vadtm.com
[3]Track Machines Connected Gesellschaft m.b.H., Austria, {benjamin.stuntner,jochen.nowotny}@tmconnected.com

Abstract

Like any other industry the railway sector undergoes great transformations. To make transportation by rail more sustainable and affordable infrastructure, maintenance is key. However, this needs a managed and collective effort of many players in order to make this efficient and easily accessible to experts of many different fields: Measurement experts, Simulation experts and Engineers who perform the maintenance. In this work, we present a market place which leverages modern technologies and standards like GAIA-X, Functional Mock-up Interface (FMI), Distributed Co-Simulation Protocol (DCP) or System Structure and Parameterization (SSP) to provide accessible and easy to use services and retrieve crucial data from other providers, while protecting sensitive data and know-how. This market place is developed under the Umbrella of the The European Rail Joint Undertaking (ERJU) of the European Union. This work is an update of our publication in Eisenbahn Ingenieur Kompendiumm 2024 (**EIK2024**) and contains several additions regarding the implementation and the usage of Modelica standards. A working prototype will be available in 2026.

Keywords: Data Spaces, Railway, FMI, SSP

1 Introduction

The railway industry is currently experiencing a unprecedented boom, which causes countless challenges. While almost all parties involved are working on innovative solutions, promoting, driving and accelerating innovation to face these challenges, it has become clear that a new form of cooperation and exchange is necessary. The time between an idea, prototype and the approval of new products is still extremely long and hence lacks the necessary speed to solve the tasks at hand. One of the most important issues, as it turns out, is the regulated exchange of knowledge, information, data, mathematical models, simulations, etc. in a controlled manner with defined rules and conditions. It is not enough to simply provide data and but also sharing it anonymously between parties. It should be possible for a third party to obtain an existing model, in return for monetary compensation (for example), enrich it with new data in order to possibly develop a better model, while protecting the know-how. On the one hand, this saves a huge amount of resources of all kinds. On the other hand, and this is the much more important aspect, it makes it possible to build on existing knowledge, refine it and thus develop new models, acquire new findings, etc. much faster and bring them timely to market. It also would open up the possibility of collaborative co-simulations. There are many different scenarios like track-vehicle interaction, interaction between vehicles and infrastructure with the environment of many standardized building blocks which could benefit of leveraging co-simulations for cooperation of different parties from manufacturers to maintainers for vehicles and infrastructure.

However, the reality is very different. Not only is it challenging from a technical point of view. There is also no consensus on how such tasks should be handled, how knowledge can be exchanged or obtained and under what framework conditions this is even possible. The legal, financial and security aspects play a decisive role here.

There are also technical challenges. While nowadays data space solutions exist (see Section 2 below) they are hard to configure and setup. This is why they are not as widely used as they could. It is therefore crucial to provide easy to use solutions which need serious development effort.

1.1 The European Rail Joint Undertaking

The European Rail Joint Undertaking (ERJU) is the direct successor to Shift2Rail, the European initiative to promote innovation, technology and cooperation in the rail sector. ERJU is unique in terms of its scale, both financially and in terms of the number of partners involved. The project volume amounts to around €1 billion, the European Union is funding the initiatives with a further €500 million and several hundred companies, scientific partners and rail operators are involved in the program. (*ERJU* 2023). The latter include Deutsche Bahn, the Austrian Federal Railways (ÖBB), ADIF, SNCF and many more. voestalpine Railway Systems (vaRS) is not only the only system solution provider represented in the consortium, it is also one of the 26 founding members of the ERJU. This is partly due to the excellent results from Shift2Rail.

The focus of ERJU is less on basic research or innovations (which are of course welcome). Rather, the program strives to bring solutions, innovations and technologies developed in recent years to the next level (TRL 5 and higher) by demonstrating and ultimately approving the use of these new products under real conditions. The objectives are:

1) Test, approve and launch products more quickly or make them marketable, 2) tighten the cooperation between rail, industry and science and improve on joint solutions and approaches, 3) strengthen the resilience of rail to current and future challenges and thus ultimately achieve the goals of Shift2Rail and ERJU, 4) become the number one most sustainable form of transportation for people and goods.

ERJU runs until 2029 and will be advertised and processed in two calls, although only the first call has currently been published, which will now run until the end of 2026. The next one will follow in the near future.

2 Data Spaces

In this section, a proper definition of data spaces is provided and it is discussed why they are crucial to the future of the railway industry. Also, a short review of existing projects from automotive industries is given to show some examples of how a rail data space might potentially look like.

2.1 Data Space Market Places

In order to simplify cooperation between individual partners in the rail industry, it is important to make the exchange of data and services as easy as possible. Similar efforts and projects already exist in the automotive sector. Examples include Catena-X (catena-x.net) for the supplier sector and the ENVITED Marketplace (www.envited.market) which is used to exchange geodata and maps to aid the development of driver assistance systems. A marketplace for the exchange of geodata of infrastructure and other data for the various players in the European rail sector is currently developed to make essential data accessible to developers and service providers. This often poses a challenge, as IPs and know-how have to be protected accordingly. This problem should now be solved. One particular aspect that needs to be considered within this framework is the provision of simulations as a service, as implemented in the ÖBB Rail4Future project (`rail4future.com`), among others. It should be possible to purchase the necessary data sources to be able to carry out complex simulations by oneself or through partners. These services can then be purchased or offered on the planned marketplace. Further possible use cases of the marketplace will be discussed in detail in the course of the article. One of the most important considerations is that centralized approaches should be avoided: So, instead of a central location where all data and services are collected, a federated system should be established where all players can offer their data and services and potential customers can use the services locally. The advantage of this is that large amounts of data are not duplicated many times and data waste is avoided. In addition, data protection is guaranteed and the integrity and trustworthiness of the data offered is ensured. In order to meet and take account of these requirements, it is necessary to set up a "Federated Data Space". This term is discussed in more detail in the next section, which also looks at other important technical requirements.

2.2 Federated Data Spaces

A federated data space can essentially be characterized by four features:

1. Unlike databases, data spaces do not store data, but only provide links to the actual data.
2. They do not force the data provided into a standardized schema, but can be made available as they are.
3. Redundancy of data is permitted.
4. They can overlap and a participant can be part of several data spaces (see (Otto 2022), (Franklin, Halevy, and Maier 2005)).

GAIA-X (gaia-x.eu) serves as the basis for this project. It is a European initiative to create a secure and trustworthy digital ecosystem for data and services. This initiative aims to strengthen Europe's sovereignty in data processing and digital infrastructure, particularly in the areas of cloud computing and data management. Federated data spaces (in the sense of GAIA-X) also offer so-called "federation services" that function across the boundaries of several data spaces (see (Franklin, Halevy, and Maier 2005)). The architecture of GAIA-X is based on the IDS ("international data space") architecture (see (Siska, Karagiannis, Drobics, et al. 2023), (Otto et al. 2018)), which provides for the connection of the individual participants via so-called IDS connectors (see (Siska, Karagiannis, Drobics, et al. 2023) and Figure 1 for a graphical representation). These "connectors" help to convert the data correctly and enable any payment transactions between the individual parties.

GAIA-X also stipulates that data transfers should be secured by corresponding policies and certificates. The safeguarding and certification of simulation data and results poses an additional challenge, as the corresponding services that are supposed to perform this safeguarding automatically must be co-developed by the providers due to the fact that this assessment is anything but trivial. Such aspects were considered in the UPSIM project (`upsim-project.eu`), among others, and corresponding processes, standards and automated toolkits were proposed to determine the credibility of data and results in the context of simulations (see (Ahmann et al. 2022)). In order to be able to offer simulations as a distributed service, the aspect that these services can possibly be connected in series must also be taken into account. This aspect is already being considered in Rail4Future. For this purpose, so-called co-simulation process graphs are used (see (Reiterer, Schiffer, and Benedikt 2022)), which represent a description of the process and dependencies on data sources and can be managed within graph databases. As the name suggests, these special graphs represent processes and co-simulations. By definition, a co-simulation is a collection of individual simulations that exchange their results with each other (see (Gomes et al. 2018)). With the help of these "co-simulation process graphs", instructions for automation software (e.g. Jenkins) can be generated and executed using a service. The special graph structure also makes it possible to represent the occurring co-simulations between different simulation participants and to generate a configuration of the co-simulation that can be received by different co-simulation protocols in order to enable the orchestration of the co-simulation and the data exchange between the individual participants (see (Reiterer, Schiffer, and Benedikt 2022) for a simple example). This can also be used to optimize process flows and ultimately use valuable resources, such as computing capacity and licenses, more efficiently (see (Reiterer and Kalab 2021)). This fairly general representation not only facilitates automation, it also enables the standardization and exchange of processes, as these can be stored and passed on like recipes and individual process steps and data sources can be exchanged or expanded. This simplifies collaboration between individual participants enormously when it comes to offering and exchanging services. For example, it is now possible to offer not only individual simple simulations as a service, but also complex co-simulations that consist of many partial simulations and are fed from a wide variety of data sources, and not necessarily just in one central location, but distributed. As co-simulation process graphs only store links and instructions instead of entire data sets or simulations, they form a good basis for fully implementing the federated data space principle in the area of simulation.

It may be noted that data space technologies are in a constant development and changes to the core technology can happen any time. However, even with the advent of new versions, technologies and standards as well as the fundamental concepts stay the same as in GAIA-X.

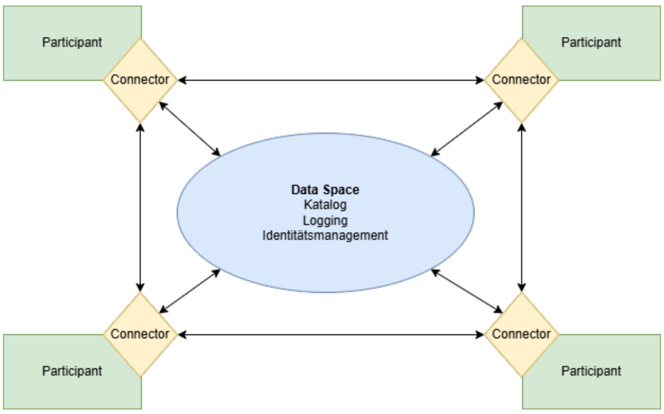

Figure 1. Structure of the Federated Data Space

3 The Railway Market Place

In the course of the ERJU cooperation, as described in Section 1.1, a demonstrator was defined along which the concept of a marketplace for railroad and railroad maintenance data is to be tested. The aim of this ERJU marketplace demonstrator is to provide the basic functionality for customer interaction and interaction with external platforms and federated data spaces. The ERJU Marketplace Demonstrator is used to gather customer experience during a user-friendly test phase. The concept is also the basis to present our innovation to a broad field of potential customers and partners - to get them on board by using our marketplace in the first version as their point of sale for the rail market on the one hand and as consumers on the other. Both customers and partners are spread across the entire rail industry. Railroad companies, manufacturers but also operators of machines, infrastructure and measurement technology. Every single company that wants to use data and services relating to rail transport and infrastructure should be able to benefit from the railroad marketplace.

The targeted customers do not have a broad overview of existing data and services in this area. Concepts for segmenting and structuring the data are therefore of high value, as most customers are specifically interested in finding the information that is valuable for the region or rail network for which they are responsible. Since there is currently no well-organized overview and data exchange facility across the rail industry, the marketplace helps to reduce the effort of creating the same data multiple times by different companies. With faster and wider availability of data, decisions can be made based on data and objectively, ultimately reducing the negative impact on infrastructure availability. In addition, any energy that can be saved is also of great value to our environment as it helps to reduce $CO2$ pollution and other negative side effects.

Potential partners are also spread across the industry and offer data and/or services relevant to the rail and infrastructure management business. They can use the marketplace as an additional sales channel to become more visible in this market area. This often results in new opportunities from previously unknown interested parties for the data for their own special use cases. The exchange of data should also be made easier for the partners through clearly defined format and quality criteria that should apply to both parties of possible transactions. A mockup of the marketplace view is shown in Figure 2.

It is also important to highlight that usage of data space technologies is complicated, hence, it is important to provide ways like simple web interfaces and services to hide the burden of configuration from the end users. This is one of the key aspects for the design of the market place we propose here and the main reason such data space solutions are not broadly adopted as there is a lot of configuration effort necessary to allow safe data transfers and provide services in a consistent manner.

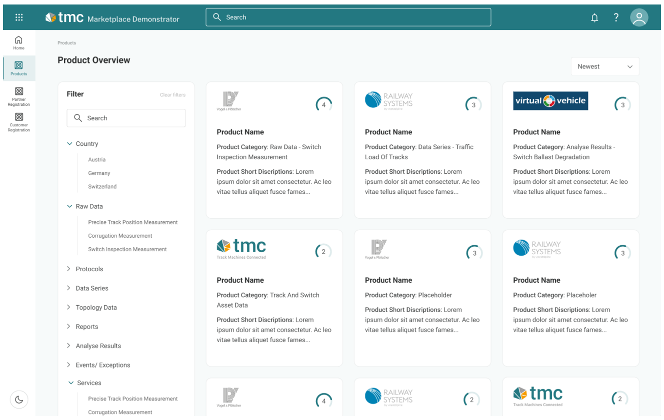

Figure 2. Marketplace View

3.1 Architecture & Services

An important milestone in the ERJU project and the implementation of the marketplace demonstrator is the definition of the architecture. The focus here is on state-of-the-art software technologies and microservice concepts, as these have proven to be extremely powerful in meeting the requirements for flexibility, scalability and maintainability in the fast-moving world of digitization. The basic idea behind microservice architecture is to divide a large and complex software application into smaller, independent services called microservices. Each microservice is responsible for a specific function or task and can be developed, deployed and scaled autonomously.

A detailed overview of the architecture can be seen in Figures 3 and 4. Figure 3 shows the architecture of the marketplace at its core, while Figure 4 shows how external services can be integrated.

The marketplace consists of the core components described in the following sections.

3.1.1 Data Catalog

For the marketplace, the data catalog plays a crucial role in the provision of information and services. This catalog, which is part of the service repositories and the data catalog repository, is at the heart of this state-of-the-art rail data marketplace. The service repository is a carefully managed collection of service self-descriptions. These self-descriptions contain a wealth of information that enables potential customers to understand exactly what type of service is being offered. The details of the data types are described in the data catalog repository. This tells customers what content is available and in what form it can be provided. This clear overview enables customers to search specifically for the data they need. Both data and services are presented as products on the marketplace interface. This creates a user-friendly environment in which customers can easily search for the information they need. Not only provides the data catalog overview pages, but also the product detail pages use these repositories to present the products and product information in an appealing way on the marketplace GUIs.

3.1.2 TrackDB: Railway Infrastructure Topology

To filter data and services as accurately and efficiently as possible, it is essential to not only locate the data country-specifically and geographically, but also to map it within the topology of a railroad infrastructure operator. Since this network model is not static but constantly changing, TrackDB proves to be a crucial tool for mastering the dynamics of the railroad topology network. TrackDB is an innovative service and part of the marketplace that makes it possible to create and manage the rail topology network based on various sources. This service goes far beyond traditional static network modeling and continuously adapts to the requirements and changes in the rail infrastructure by opening up freely accessible sources but also accessing partly (partially) closed data based on partnership with infrastructure operators.

The key to TrackDB's effectiveness lies in its ability to dynamically capture and update information about track sections, switches, signals and other elements of the rail network. This is done in real time, allowing users to access up-to-date data at all times. This means that the content provided by service partners and data providers can also be clearly assigned to the network, allowing this information to be filtered down to sections of specific tracks or even individual components such as switches.

3.1.3 Services

Another way to expand the product range is to host third-party services and ensure the smooth operation of these services. This aspect significantly expands the functionality of the marketplace and offers service providers the opportunity to integrate their specialized solutions directly into the platform. The "Service and Data Center" marketplace acts as a central hub for hosted services from third-party providers. Providers have the opportunity to host their services in this environment in order to make them accessible to a broader target group. This opens up new business opportunities and promotes innovation in the rail industry.

3.2 Simulation Services and Co-Simulation Standards

A special key feature of the market place is to provide simulation services for predictive maintenance.

To enable the seamless integration of these services, service providers must offer "black box" services that can communicate via standardized interfaces. These interfaces serve as templates and enable interaction between the hosted services and the data available on the platform. This ensures that the hosted services can work effectively with the required data inputs.

If a service is filled with data and executed, the next step is to output results and feed them back into the marketplace. Hosted services should be able to provide the results in a standardized file format. This facilitates integration and data exchange between the various services and users on the platform. The ability to integrate hosted

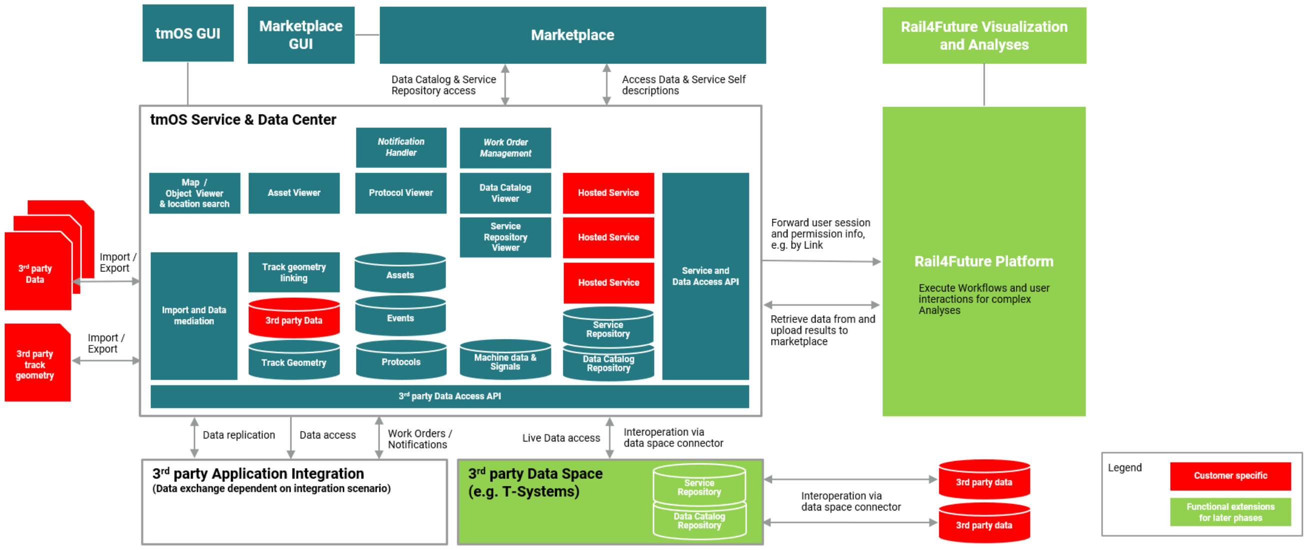

Figure 3. High level System Architecture

services into the platform opens up a world of potential. It allows customers to access a wide range of specialized services, from data analytics to predictive maintenance, without having to leave the platform and without having to worry about running their own platform.

3.2.1 The Role of Co-Simulation Standards in the Context of the Market Place

Co-Simulation standards like the Functional Mockup Interface (FMI) (FMI-Working-Group 2020) and the Distributed Co-Simulation Protocol (DCP) (Modelica Association Project DCP 2019) play a crucial role as they are leveraged to enable easy integration of new simulation services. First, mappings between input data and simulations can be created with ease and tools for reading out inputs and writing outputs can be provided. Second, they provide a standardized interface developers and engineers can rely on to provide a unified way to read and write data from and to simulations. This standardization is crucial as simulations are very diverse and rely on very different forms of input. The metadata file an FMU or DCP participant provides ways to auto generate forms for data provision as other participants know what they have to provide in order to make the simulation run and what to expect as output. We show an example in Section 3.2.2. Since the FMU defines the proper in and outputs this can be used to tell a potential user which data they need to provide (for example as CSV file) and which names the fields have. Another important benefit of co-simulation standards like FMI is that FMUs can be stored as standalone files that can be shared between different parties. This allows of easy exchange of models.

The only real drawback FMUs have is that it is not always easy to guarantee proper runtimes of the underlying models. So proper set-up sandbox environments are needed as well. A data space can help with that as the models can run with an enclosed environment (e.g. on site with the owning party) and just input and output data is exchanged, which only works for non real-time computations of course. However, solutions to this problem also could be provided in the future.

Furthermore, the System Structure and Parameterization (SSP) standard (see e.g. (Hällqvist et al. 2021)) can be used to link different simulation services together. This enables cooperation between different simulation domains and helps with automatic service execution as many tools support these open Modelica standards. As already mentioned, this can be also supported by graph based process descriptions (Stefan H. Reiterer and Clemens Schiffer 2021)(M. Stefan H. Reiterer and Clemens Schiffer 2023) which allow to make reproducible computation processes by auto generating and orchestrating execution pipelines. The viability of this approach in the context of the rail industry was already demonstrated (Ozan Kugu 2023).

Additionally, the data space architecture enables us to use data space connectors to safely upload sensitive data which is necessary for a multitude of simulations. This helps to protect know how of third parties and adds an additional layer of security.

3.2.2 An Example Use Case: Ballast Degradation

As an example to illustrate the diverse interaction possibilities on the marketplace, a use case was chosen called "Monitoring of Track Settlement and Identification of Single-Faults at Switches". One of the goals there involves the monitoring and prediction of ballast degradation in the switch-area, especially below the frog. In order to successfully execute this use case, both the service and all required relevant information sources must be defined in the service repository and the data catalog repository. It should be noted that we want to outline the process to give a better understanding of how this works, rather than explaining the model in detail, because if a simulation is offered as a service the inner workings of a simulation

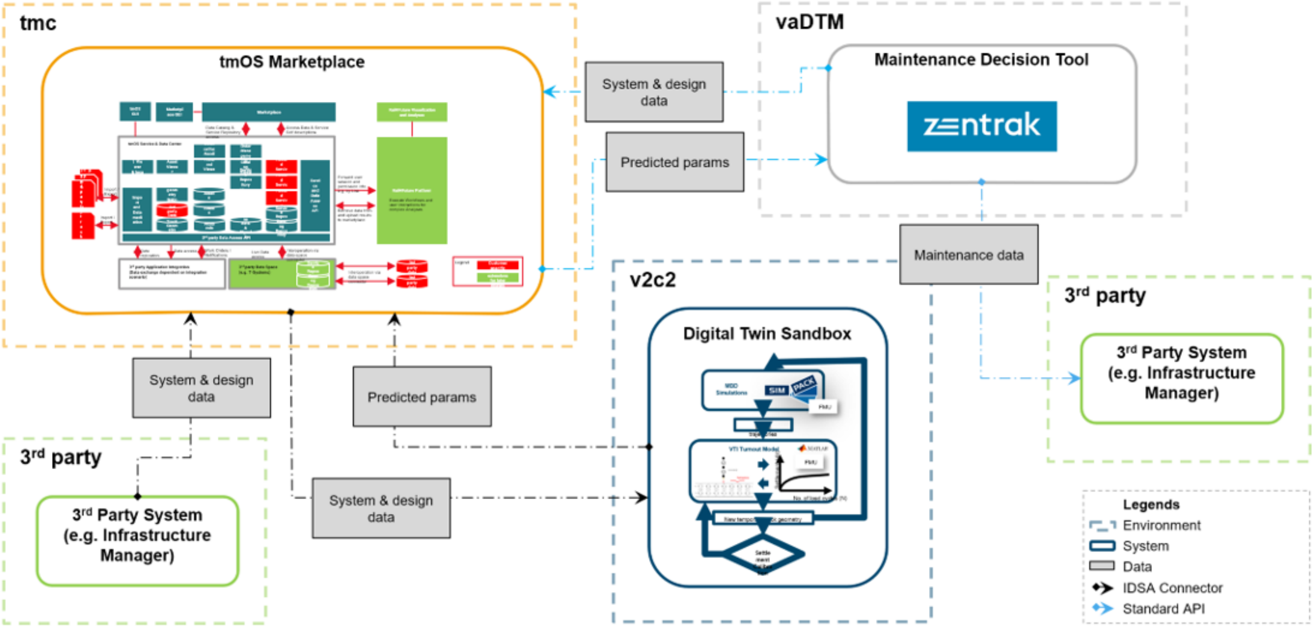

Figure 4. High level System Architecture with Integrated Services

have to be abstracted away within an FMU beforehand to make it usable. However, this heavily depends on the tools in use and would go far beyond of the scope of this work. The execution of this service offering requires several steps on the product offering side and is intended to demonstrate the possibilities for querying multiple, different data from different providers for a specific track section. Several interactions with menu selections and the interactive map function are highlighted. A key aspect is also to show how quickly the customer can filter the area of interest.

Here, the consumer not only buys the needed "raw data" for the simulation from data providers on the market place but also the (automated) execution of a complex simulation for calculating ballast degradation, which requires this "raw data" as input data. The algorithm itself does not run directly in the marketplace but in a separate Rail4Future simulation environment developed by Virtual Vehicle (`www.rail4future.com`).

The simulation itself runs within a Functional Mockup Unit. A generated pipeline is triggered which starts the FMI master. The master reads the data and executes the simulation and then writes out the simulation results.

After the customer has completed the purchase, the marketplace will exchange the relevant data with this environment, initiate the simulation there and finally make the processed data available to the customer again on the marketplace. See Figure 5 for a mock up view on the service.

Nevertheless, there are many more examples of simulations ranging from mechanical simulations (for example vehicle-track interaction), thermal simulations (behavior of infrastructure under changing temperature) and many more. Many tools which are used (for example Matlab or

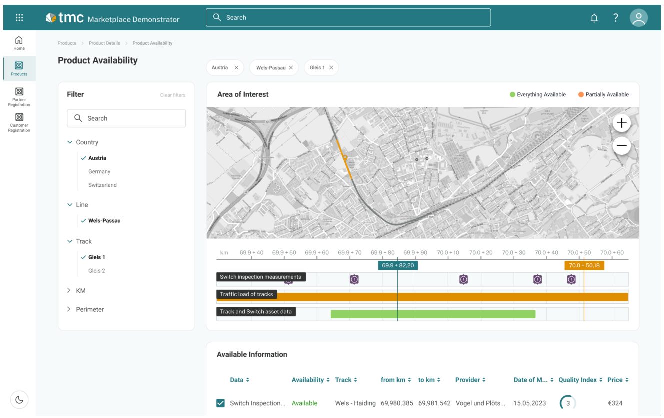

Figure 5. GUI View of the Ballast degradation Use Case

Simpack) provide export capabilities for Modelica standards like FMI or SSP which makes it an obvious choice regarding exchange and co-simulations of models.

4 Conclusion

Managing Railway infrastructure needs a lot of know-how data and coordinated joint effort. Modern standards and initiatives like GAIA-X, FMI, DCP and SSP provide the foundation to enable the railway industry for a modern, comfortable and standardized workflow, and provides non-experts with a place where they can easily buy services which can be used without much effort.

In this work we demonstrated how such a market place can look like and a practical example how this can work. A demonstration shall be ready in 2026 and is open for collaboration. The ERJU offers the ideal platform for the joint development of further features and the promotion and on-boarding of additional partners. Infrastructure op-

DOI
10.3384/ecp218901

erators and rail transport companies in particular are invited to participate and get involved. Ultimately, they are the ones who benefit from accelerated developments, reduced costs and higher availability and can ultimately pass this on to their end customers.

Acknowledgements

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or Europe's Rail Joint Undertaking Neither the European Union nor the granting authority can be held responsible for them. The publication was written at Virtual Vehicle Research GmbH in Graz and partially funded within the COMET K2 Competence Centers for Excellent Technologies from the Austrian Federal Ministry for Innovation, Mobility and Infrastructure (BMIMI), Austrian Federal Ministry for Economy, Energy and Tourism (BMWET), the Province of Styria (Dept. 12) and the Styrian Business Promotion Agency (SFG). The Austrian Research Promotion Agency (FFG) has been authorised for the programme management.

References

Ahmann, Maurizio et al. (2022). "Towards continuous simulation credibility assessment". In: *Modelica Conferences*, pp. 171–182.

ERJU (2023). Web Site. Accessed: https://rail-research.europa.eu/about-europes-rail/europes-rail-reference-documents/ 2025-04-18.

FMI-Working-Group (2020). *Functional Mock-up Interface for Model Exchange and Co-Simulation*. https://github.com/modelica/fmi-standard/releases/download/v2.0.2/FMI-Specification-2.0.2.pdf. Accessed: 2021-04-12, V 2.0.2.

Franklin, Michael, Alon Halevy, and David Maier (2005). "From databases to dataspaces: a new abstraction for information management". In: *ACM Sigmod Record* 34.4, pp. 27–33.

Gomes, Cláudio et al. (2018). "Co-simulation: a survey". In: *ACM Computing Surveys (CSUR)* 51.3, pp. 1–33.

Hällqvist, Robert et al. (2021). "Engineering domain interoperability using the system structure and parameterization (ssp) standard". In: *Modelica Conferences*, pp. 37–48.

Modelica Association Project DCP (2019). *DCP Specification Document, Version 1.0*. Linköping, Sweden: Modelica Association. URL: http://www.dcp-standard.org.

Otto, Boris (2022). "A federated infrastructure for European data spaces". In: *Communications of the ACM* 65.4, pp. 44–45.

Otto, Boris et al. (2018). "Ids reference architecture model. industrial data space. version 2.0". In.

Ozan Kugu Rebecca Nowak, Stefan H. Reiterer et. al. (2023-10). "An FMI- and SSP-based Model Integration Methodology for a Digital Twin Platform of a Holistic Railway Infrastructure System". en. In: *15th Modelica Conference 2023*, pp. 717–726. DOI: 10.3384/ecp204. URL: https://ecp.ep.liu.se/index.php/modelica/issue/view/83/86 (visited on 2024-01-18).

Reiterer, Stefan H and Michael Kalab (2021). "Modelling deployment pipelines for co-simulations with graph-based metadata". In: *International Journal of Simulation and Process Modelling* 16.4, pp. 333–342.

Reiterer, Stefan H, Clemens Schiffer, and Martin Benedikt (2022). "A Graph-Based Metadata Model for DevOps in Simulation-Driven Development and Generation of DCP Configurations". In: *Electronics* 11.20, p. 3325.

Siska, Veronika, Vasileios Karagiannis, Mario Drobics, et al. (2023). "Building a dataspace: Technical overview". In: *Gaia-X Hub Austria*.

Stefan H. Reiterer and Clemens Schiffer (2021-09). "A Graph-Based Meta-Data Model for DevOps in Simulation-Driven Development and Generation of DCP Configurations". en. In: *14th Modelica Conference 2021*, pp. 411–417. DOI: 10.3384/ecp21181411. URL: https://ecp.ep.liu.se/index.php/modelica/article/view/219 (visited on 2022-07-26).

Stefan H. Reiterer, Mario Schwaiger and Clemens Schiffer (2023-10). "Graph-Based Meta-Data Model for DevOps: Extensions to SSP and SysML2 and a Review on the DCP standard". en. In: *15th Modelica Conference 2023*, pp. 159–166. DOI: 10.3384/ecp204. URL: https://ecp.ep.liu.se/index.php/modelica/issue/view/83/86 (visited on 2024-01-18).