

32nd CIRP Design Conference

A Conceptual Model-based Digital Twin Platform for Holistic Large-scale Railway Infrastructure Systems

Shiyang Zhou^a, Stefan Dumss^a, Rebecca Nowak^b, Rainer Riegler^a, Ozan Kugu^a, Martin Krammer^c, Manfred Grafinger^a

^aTU Wien, Institute of Engineering Design and Product Development, Leurgasse 6 / 307, 1060 Vienna, Austria

^bVRVis Zentrum für Virtual Reality und Visualisierung Forschungs-GmbH, 1220 Vienna, Austria

^cVirtual Vehicle Research GmbH, Co-Simulation and Software Group, Inffeldgasse 21a, 8010 Graz, Austria

* Corresponding author. Tel.: +43-1-58801-30723 ; fax: +43-1-58801-30798. E-mail address: shiyang.zhou@tuwien.ac.at

Abstract

Railway is crucial for a sustainable transportation system. Digitization of the railway infrastructure system has a proven impact on improving its resilience and efficiency while reducing the overall cost. The increased digitization of railway infrastructure systems provides a solid basis for the realization of a digital twin (DT), which can continuously merge different sources of data and diverse models for monitoring, diagnostics, and prognostics to optimize railway performance. However, the railway infrastructure system is sophisticated and consists of a series of subsystems, including turnouts, tunnels, vehicles, etc. Current researches mainly focus on the DT application of the single subsystem, but the comprehensive communication and interactions between diverse railway infrastructure subsystems remain a largely under-explored domain. Aiming to fill the gap, this research introduces a conceptual architecture to develop a DT platform to seamlessly integrate and automatically link digital models and data of different subsystems into a holistic large-scale railway infrastructure system. The model cooperation frameworks, communication approaches and visualization toolbox are presented in this work. One typical use case of the platform will also be implemented and discussed. It is expected that this research will shed new light on the sustainable development of the holistic large-scale railway infrastructure system.

© 2022 The Authors. Published by Elsevier B.V.

This is an open access article under the CC BY-NC-ND license (<https://creativecommons.org/licenses/by-nc-nd/4.0>)

Peer-review under responsibility of the scientific committee of the 32nd CIRP Design Conference

Keywords: Intelligent railway infrastructure system; Digital Twin; Industry 4.0; System Design;

1. Introduction

Railway has played a decisive role in building an environment-friendly and energy-efficient transportation system in the past several decades. However, despite its proven safety and efficiency, the railway infrastructure system, including tracks, turnouts, tunnels, and bridges, still suffers from many problems, which are illustrated in Figure 1. In total, the European countries allocated more than 25 billion EUR annually for the maintenance and renewal of the railway system [20].

To cope with these potential crises, railway digitization has been deployed to enhance the availability, resilience, and performance of the holistic large-scale railway infrastructure systems. Spanish National Railway, Renfe, and the cross-channel operator, Eurostar, have already successfully applied methods for digital predictive maintenance in practice and benefited significantly from it [13]. DT plays a transformative role in how

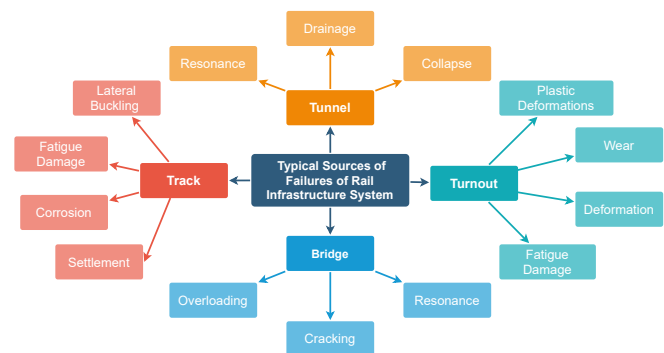


Fig. 1. Typical sources of failure of the railway infrastructure system

we operate intelligent railway systems. The modern DT system, assisted with the Internet of Things (IoT), cloud computing and big data analytic technology, can build a virtual platform with high fidelity for operational monitoring, diagnosis, and prognosis of the complex railway infrastructure system. For exam-

ple, Ahmadi et al. [1] proposed an approach to implement DT for operational monitoring and controlling in electric railway power systems. In another example, Kaewunruen et al. [9] suggested a Building Information Modelling (BIM) technology-based DT for the life cycle cost estimation and maintenance visualization of the railway turnout systems. These works demonstrate that railway infrastructure subsystems can derive great benefit from existing numerical models and analytical methodologies. However, the railway infrastructure system is sophisticated and consists of a large quantity of subsystems, which operate cooperatively and interactively to determine the state of the holistic system. In currently developed DT systems, little attention was paid to efficient interactions and communication between diverse models of the infrastructure subsystems. Due to the ever-rising complexity of these subsystems, a novel DT platform that can seamlessly integrate and automatically link digital models and data of individual subsystems into a holistic large-scale system is needed urgently.

A conceptual model-based DT platform, denoted as the Rail for Future Platform (R4F Platform), is presented in this paper to fill this gap. Referring to the Reference Architectural Model Industrie 4.0 (RAMI 4.0) [8], it can combine all elements of the holistic railway infrastructure system in a layered model, in which complex interrelations can be broken down into smaller and simpler clusters. With the R4F Platform, different railway infrastructure assets can be automatically and efficiently integrated, and a continuous flow of reliable and meaningful information throughout the entire life span of different railway infrastructure subsystems can be ensured. Moreover, R4F Platform may also lay an essential basis for the design of a virtual assessment platform as a universally applicable solution for digitized railway systems in the future.

This paper is organized as follows. In Section 2, related researches on DT and its application in railway infrastructure systems are reviewed. Section 3 analyzes the mandatory and expected requirements for the proposed platform, and presents an overview of the R4F Platform. Furthermore, a use case for the prediction of rail track geometry is implemented. Finally, Section 4 presents the conclusions and future work.

2. Related Work

NASA's Apollo space program first introduced the twin concept, then the theory of digital twin was introduced by Grieves [7] in 2003. He proposed a digital twin prototype for a Product Lifecycle Management system, and pointed out that the DT needs to focus on the data connection between the physical entity and the virtual model in the application. After that, the digital twin technology has experienced exponential growth. In [14], the 8-dimension model for the development and operation of digital twins for technical systems and services is described. Furthermore, the six design elements of digital twins are specified, which include concepts of *digital master* and *digital shadow* as preliminary stages to digital twin.

The DT application in the railway infrastructure system is still in its infancy. Kaewunruen et al. [9] has built the world's first 6D DT-aided model for life cycle management of a railway turnout system based on BIM technology by integrating data across the whole life cycle into a big data-sharing plat-

form. However, the development of data-related functions such as smart maintenance was not mentioned. Kampczyk et al. [10] also applied DT to status monitoring for railway turnout systems. They analyzed and verified the health condition of the railway turnout systems with the cyclic data provided by an established DT model and used it to connect to other intelligent devices to share the information. Ye et al. [17] has successfully applied the DT to smart maintenance of railway subsystems such as bridges. They carried out real-time structural health monitoring of two bridges in Staffordshire, UK, by linking the measurement data to the DT model. Zhang et al. [19] developed a DT-based framework for fault diagnosis and real-time health status monitoring of railway point machines by analyzing the electric current waveform and magnitude. These works mainly focus on the single railway infrastructure subsystem. In order to build a holistic infrastructure DT system, one needs to integrate the models and data from different subsystems properly.

Model integration and synchronization in DT systems have been pioneered by several investigations. Eckhart et al. [4] proposed a passive state replication approach to solve the synchronization problems in which the DT system was fed with a stimulus constituted by the inputs of the physical system. However, there are some limitations, as their approach could lead to the loss of some essential data of the physical system during the passive state replication process, which may affect the fidelity of the DT system. In addition, Seongjin et al. [18] proposed a novel structure of the DT platform named universal Digital Twin (uDiT). They designed an object management group and data distribution-based communication middleware that can easily support any entities to share data in different formats and applied it to an autonomous driving system.

As is seen from the reported work, DT has already been applied in smart maintenance of some railway infrastructure subsystems. However, comprehensive communication and effective interaction between diverse railway subsystems have rarely been studied. This presented study aims to fill the gap. The R4F Platform - the first of its kind - is designed, which may integrate the data and models of the different subsystems automatically to build a fully connected and digitized holistic infrastructure system.

3. R4F Platform Design

In this section, we analyze the requirements of the platform, based on which we designed a conceptual model-based digital twin for the holistic large-scale railway infrastructure: R4F Platform. Finally, a use case of the R4F Platform for prediction of track geometry is presented.

3.1. Requirement Analysis

We carried out a questionnaire among our industrial partners (OEBB, Wiener Linien, etc.). Based on the results of the questionnaire, the following requirements for preparing, processing and evaluating the proposed R4F Platform get defined. The requirements are based on the demands of a single DT as well as the prerequisite of an efficient model and data integration process. To distinguish the importance of different requirements, each of them is marked with either a letter

(M), which means “mandatory”, or a letter (E), which means “expected”.

Compatibility (M): The compatibility refers to the platform’s ability to access, integrate and analyze the data and models from various subsystems. Standardized file formats and vocabulary should be defined so that any subsystem can be mapped, dynamically interacted with, and well connected in the platform. Compatibility is an essential characteristic of the proposed R4F Platform, as it can directly influence the platform’s efficiency and consistency.

Synchronization (M): The platform will require the physical entities of different integrated subsystems to be characterized and managed along the same time-axis with a unified data format. These data include the geometry, state, attribute, and internal mechanism of the subsystems, which form a digital virtual mapping of the real-time state of the physical entities.

Reliability (M): The proposed platform requires a reliability assessment system, in which the system can supervise the states of the components of the platform and make corresponding adjustments. With the help of the reliability assessment system, the platform can run smoothly despite disruptions.

Fidelity (M): The fidelity of the platform describes the proximity of the integrated model and the physical entity. It is required that the integrated model maintains a high degree of proximity of the geometry, state, phase, tense, etc. High fidelity is the prerequisite for the development of its future functions.

Expandability (E): The proposed platform should have sufficient expandability to add and integrate new sub-models. It should also allow modifying or replacing existing sub-models and functionalities. Lacking expandability can cause problems for the maintenance and reusability of the platform.

Interactivity (M): The interactivity refers to not only the interaction between the users and the platform, but also the interaction between the designers and the platform, as well as between the physical models and the platform. A closed loop should be built based on the above-mentioned interactions, which may help the platform be updated and improved.

Real Time (E): The platform should make a prompt response to real-time input data to analyze the emergent problems and give corresponding solutions. The platform tracks the changes as versions and ensures consistent behaviour of simulations in their lifecycle.

3.2. Overview of the Platform

In the following section, a conceptual model-based DT structure for the holistic large-scale railway infrastructure grounded on the requirements mentioned in Section 3.1 will be presented. The proposed platform is built based on the demands of our European industrial partners in the R4F project. Thus, it is designed to be operation-oriented and aims to be compatible with the Technical Specifications for Interoperability (TSIs) from the European Union Agency for Railways (ERA) [3]. The proposed R4F Platform framework can be summarized into six layers: asset layer, integration layer, infrastructure layer, function layer, visualization layer and interaction layer (see Fig. 2).

Asset Layer

Based on the available models and data from our industrial partners, the asset layer is proposed to include all the existing assets to be fed into the platform. And it is possible to incorporate other potential assets. Three kinds of assets are mainly considered due to their different nature: physically-based numerical models, measurement data and data-based machine learning models. The measurement data derives from various measuring systems in the railway ecosystem. Since the railway ecosystem is long-lived, all measurements have to be contextualized with the time of measurement. In general, providing context and metadata to the measurement data allows for more extensive applications. Besides, measurement data only covers the past and can not predict the future. Numerical and machine learning models based on measurement data can enrich the input assets and, at the same time, save time and cost. More importantly, these models can also be used for prediction. The prediction can be based on the historical measurement data or a completely new subsystem model. Both are approximations and therefore have to include the model’s fidelity so that the platform can calculate the total fidelity out of the sub-simulations. The model’s fidelity will be controlled by the Data Owner and Asset Provider based on the feedback from the Asset Integrator, which will be exemplified by the following use case in Section 3.3. However, protocols of how they are built and used have to be included in a way that the platform can interpret them. A standard model like SysML is supposed to be applied, as SysML fulfills the international standard ISO/IEC/IEEE 42010 for the systems architecture description. It is widely used for model requirements, behavior, structure and parametric of complex systems, which may include software, hardware, data or data other components of human-made and natural systems [5].

Integration Layer

The integration layer manages the import and integration of models to the R4F Platform. Together with the infrastructure layer, it instantiates and executes simulations that are defined at the function layer. This approach is also known as system simulation. In particular, when different domain-specific solvers, models, and tools are integrated into one common simulation, this is known as cooperative simulation, or in short, co-simulation [6]. Co-simulation methods were advanced significantly during the last decade, including modeling, coupling algorithms, and (structural) integration methods [2, 15]. The Functional Mock-up Interface (FMI) represents the most important interoperability standard. It is maintained as a Modelica Association standard, and its most recent version is FMI 2.0.3 [12]. FMI defines a software interface for co-simulation, effectively contributing to interoperability between models and tools. Its companion standard, the Distributed Co-Simulation Protocol (DCP) [11], is a compatible platform and communication medium independent standard for the integration of models or real-time systems into simulation environments. Therefore, we propose to use FMI as a standard in the platform for model integration. The physically-based models can be integrated with FMI. However, it may need an additional effort, as the platform has to provide a data integration functionality separated but interoperable with FMI.

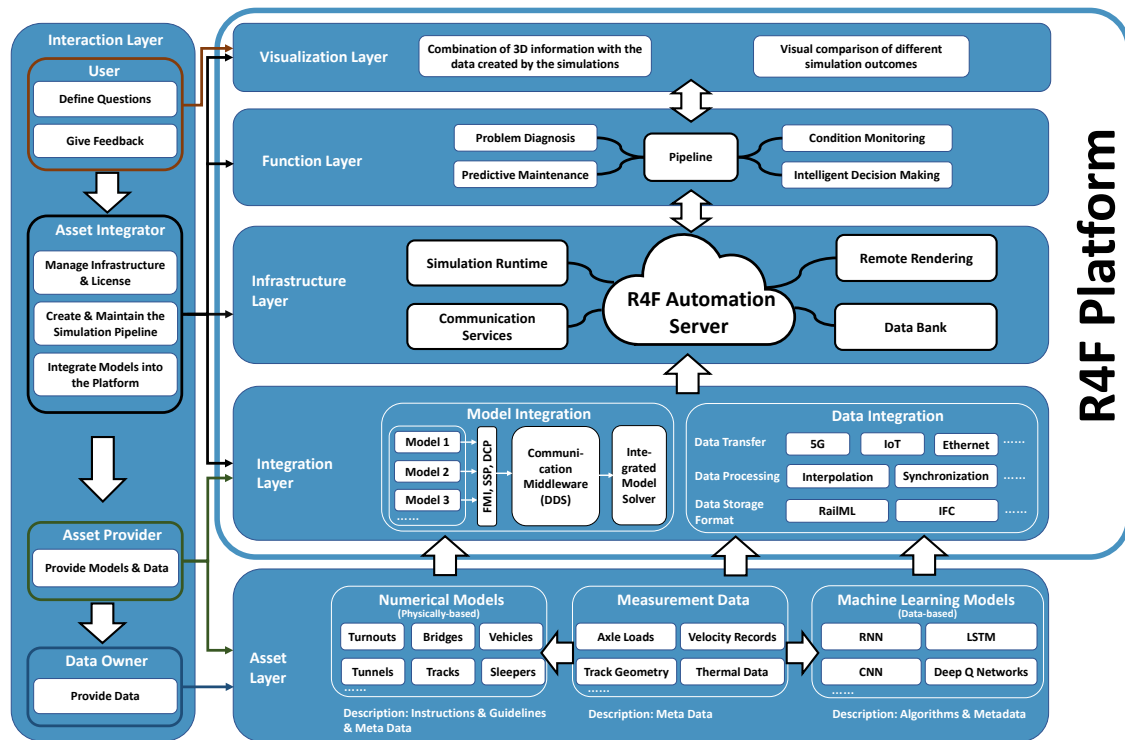


Fig. 2. The R4F Platform landscape

This functionality will provide interfaces for simple data transformation, data transfer between models, data transport between platforms, and storage management of the data. In the data integration section, due to the immense complexity of the holistic railway system, it is hard to deliver all real-time data and synchronize them in reality. However, the platform should be able to respond as quickly as possible to new assets and initiate new calculations. After that, the raw data will be processed, such as data interpolation, data interception, data synchronization, etc. The processed data will be stored in a standard data exchange format like Industry Foundation Classes (IFC) format and railway mark-up language (RailML) format to ensure efficient and consistent railway modeling and the compatibility of the platform. Independent from the type of structural integration, the integration layer enables data or model providers to keep track and manage their contributions. That means they maintain full ownership and responsibility.

Infrastructure Layer

The infrastructure layer, or execution and communication layer, is responsible for running simulations on the underlying hardware and controlling the simulation pipelines defined by the asset integrators and providers. It relies on the interfaces defined at the integration layer, the pipelines defined in the function layer, and the infrastructure configuration provided on its own layer. There are various ways to run sub-simulations: i) Simulations are containerized with an FMI interface and are run as a container on the platform as the pipeline is called; ii) They are not containerized but within an FMI interface and run in the same company's ecosystem; iii) They are data-model based with or without an FMI and are run within the same

company's ecosystem; iv) They can only be run externally via another R4F Platform by connecting to the other infrastructure layer. Besides, the infrastructure layer also provides a remote rendering function for the visualisation process, a Key Performance Indicator (KPI) System for the reliability assessment, and serves as a data bank.

Function Layer

In the function layer, the expected usages of the R4F Platform are listed and defined. The pipelines, as mentioned earlier, and the input variables are managed within this layer. While the integration layer is responsible for managing assets, this layer is responsible for managing the whole simulation flows defined in the pipelines. The most critical aspect of this layer is the ability to integrate new simulation pipelines, provide life cycles for them and query the data and models by timestamps to ensure the synchronization requirement. Input to this layer is provided by the user through the visualization layer or directly by the asset integrators. For details, see the interaction layer.

Visualization Layer

The visualization layer provides users with a holistic system for exploring and understanding the data and simulation results available in the R4F Platform. Spatial (3D) and non-spatial (numeric, quantitative) information will be combined with techniques such as multiple coordinated views using linking, brushing and overlay visualizations. The variations in the scale of railway infrastructure can be addressed with techniques like semantic zooming, where different data is shown for different zoom levels qualitatively. A visual programming approach

for simulation steering has been used in the context of flood simulation [16]. Based on this approach we want to allow domain experts to use visual programming for simulation design and configuration, providing easy-to-use dashboards for other users. Specialized visualization like heat maps and parallel coordinates in combination with overviews, details on demand, zooming, and filtering will allow users to analyze the data visually. They will also be able to compare different simulation runs and study the sensitivity of simulation parameters.

Interaction Layer

The interaction layer illustrates the user interaction with the platform and assures the interactivity of the platform. On a simplified view, there can be four major roles which interact with the platform. *Users* are the experts that can interpret simulation results and define actions according to the outcome. They give feedback on what simulations they need and what outcomes are of interest to them. The graphical user interface for this group needs to be simple and intuitive. The *Asset Integrator* picks up that feedback and creates or adapts the simulation pipelines to the user's needs. This may also include the organization and management of parts of the infrastructure like allocation of licenses and interaction with *Asset Providers* to integrate or update models in the platform. Therefore, the *Asset Integrator* has to be an expert within the simulation domain and a manager for the platform. The *Asset Integrator* has access to more complex visual tools like visual programming for adapting simulation pipelines. The *Asset Providers* supply the platform with models and data. They are experts in their domain who develop and maintain models. To do so, they use the data provided by *Data Owners*. *Data Owners* are the owners of the data as legal identity. They may also own the models developed by the *Asset Provider*. This user-centric approach should ensure that only models that are needed and used by the *Users* are integrated.

This holistic railway infrastructure system is sophisticated. The stability and performance of the whole system cannot be determined by each single subsystem but by the interaction of various subsystems. Integrating the subsystems into one holistic system creates added value that exceeds the sum of its parts. Based on existing models and data, a use case that demonstrates the integration of vehicle and track subsystem is shown in the next section.

3.3. Case Study: A R4F Platform Application for Prediction of Track Geometry

Track geometry may directly influence the safety of the entire railway infrastructure system. However, due to the long-term wear and tear of the track, it is hard to predict the lifetime of the track in use. This subsection introduces a smart maintenance process of the railway track as one application of the R4F Platform by studying the interaction between the vehicle and track subsystems with the integration of numerical and machine learning models, as shown in Fig. 3. Starting from the Integration Layer ① the problems and possible solutions of the railway track system are defined by *Users* (Railway Operators). Based on feedback from *Users*, the *Asset Integrator* will create the integration pipelines and provide them to the function layer

and manage corresponding licenses of various software. Afterward, the *Asset Integrator* will demand and check the models and data from the *Data Owner*, who is the legal owner of *Assets*. Three assets are listed in our case with their respective parameters in *Asset Layer* ②. The measurement data refers to the vertical deformation of the track geometry, which will be collected through measurement cars. Then, the measurement data will be fed to the *Vehicle and Track Interaction (VTI)* model and *RNN* model as input. The *VTI* model refers to a numerical model that considers the discrete support of elastic rails. It can use the processed measurement data of the track geometry as input, and it predicts the development of vertical track deformation. The *RNN* model is a machine learning model that can be fed with processed measurement data and keep iterating based on the proximity of the model. The finally trained *RNN* model can make precise predictions about the geometry and life cycle of the track, which may enrich the dataset of the track sub-model. In ③, all the data and models with initialized parameters are integrated within the integration layer into the platform via their respective interfaces. Then the integrating process is generated, and all these assets will be stored in the graph database. Every simulation path must have its own validated graph, which is created by *SysML* models. The graph is then transformed into a pipeline which is responsible for the execution of the defined workflow. After that, simulations are carried out in the virtual machines, and the simulation results are written to a data storage. Based on the user-defined output functions, the simulation results are visualized at the visualization layer in ④ (see Fig. 3).

4. Conclusion and Future Work

Digital Twin brings digital and automation technology to the traditional railway infrastructure system to make it more resilient, reliable, and intelligent. The application of DT in the railway infrastructure system shows a manifold potential, but different models need to be efficiently integrated into a holistic platform for wide use.

In this paper, a conceptual model-based digital twin platform that aims to integrate measurement data, numerical models, and machine learning models from different railway subsystems into a holistic large-scale railway infrastructure platform is proposed for the first time. The R4F Platform may ensure that reliable and valuable data can continuously flow throughout the whole life span of the holistic system and its subsystems, which may help build a fully connected and digitized railway infrastructure system. We first identify the crucial requirements of the R4F Platform to define its essential characteristics, which ensures that the platform works consistently and reliably. After that, a 6-Layer based architecture of the proposed R4F Platform is presented. The R4F Platform will be exploited for problem diagnosis, predictive maintenance, condition monitoring, and some other cases based on user demand. Furthermore, corresponding visualization processes and techniques will provide added value to the platform users.

The presented study outlines a broad spectrum of conceptual architecture for an entirely virtual validation platform for the integration of model and data of railway infrastructure subsystems. The R4F Platform is still under development, and it plays an instructive and essential role in the R4F project. More use

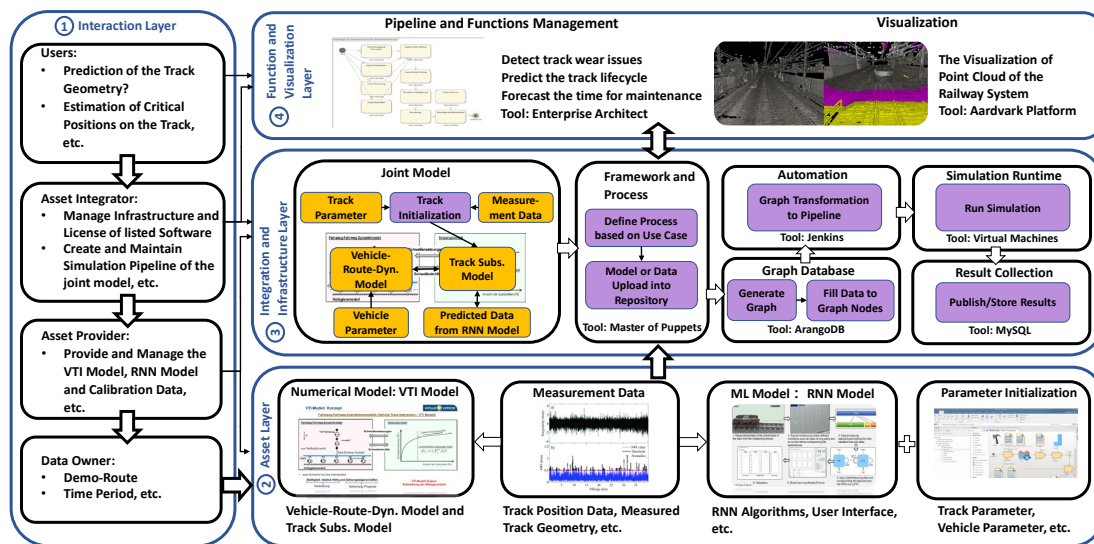


Fig. 3. Flow chart of model integration

cases shall be integrated into this platform to validate its reliability and fidelity in future publications. Visualization research will be conducted according to the principles of user-centered design and improved iteratively based on the properties of available data and the feedback of domain scientists.

Consequently, the establishment of the above-mentioned conceptual R4F Platform is a major part of the recently started R4F research program. It will provide plenty of opportunities for future work, pushing sustainable development of the holistic large-scale railway infrastructure system.

Acknowledgements

The authors would like to acknowledge the financial support of the COMET Program R4F (882504) within the COMET Competence Centers for Excellent Technologies from the Austrian Federal Ministry for Climate Action (BMK), the Austrian Federal Ministry for Digital and Economic Affairs (BMDW), the Vienna Business Agency and the Styrian Business Promotion Agency (SFG). The Austrian Research Promotion Agency (FFG) has been authorised for the program management.

References

- Ahmadi, M., Kaleybar, H., Brenna, M., Castelli-Dezza, F., Carmeli, M.S., 2021. Adapting digital twin technology in electric railway power systems, in: 2021 12th Power Electronics, Drive Systems, and Technologies Conference (PEDSTC), IEEE. pp. 1–6.
- Benedikt, M., Watzenig, D., Zehetner, J., Hofer, A., 2013. Nepce-a nearly energy-preserving coupling element for weak-coupled problems and co-simulations, pp. 1021–1032.
- Brenna, M., Foiadelli, F., Leone, C., Casiraghi, F., 2019. Modeling and simulation of railway network in compliance with technical specifications for interoperability, in: 2019 AEIT International Annual Conference, pp. 1–6.
- Eckhart, M., Ekelhart, A., 2018. A specification-based state replication approach for digital twins, in: Proceedings of the 2018 Workshop on Cyber-Physical Systems Security and Privacy, pp. 36–47.
- Friedenthal, S., Moore, A., Steiner, R., 2014. A practical guide to SysML: the systems modeling language. Morgan Kaufmann.
- Gomes, C., Thule, C., Broman, D., Larsen, P.G., Vangheluwe, H., 2017. Co-simulation: State of the art. arXiv preprint arXiv:1702.00686.
- Grieves, M., 2014. Digital twin: manufacturing excellence through virtual factory replication. White paper 1, 1–7.
- Hankel, M., Rexroth, B., 2015. The reference architectural model industrie 4.0 (rami 4.0). ZVEI 2, 4–9.
- Kaewunruen, S., Lian, Q., 2019. Digital twin aided sustainability-based lifecycle management for railway turnout systems. Journal of Cleaner Production 228, 1537–1551.
- Kampczyk, A., Dybel, K., 2021. The fundamental approach of the digital twin application in railway turnouts with innovative monitoring of weather conditions. Sensors 21, 5757.
- Krammer, M., Schuch, K., Kater, C., Alekeish, K., Blochwitz, T., Materne, S., Soppa, A., Benedik, M., et al., 2018. The distributed co-simulation protocol for the integration of real-time systems and simulation environments, in: Proceedings of the 50th Computer Simulation Conference, pp. 1–14.
- Modelisar Consortium, Modelica Association Project "FMI", 2021. Functional Mock-up Interface for Model Exchange and Co-Simulation, Version 2.0.3.
- Scordamaglia, D., 2019. Digitalisation in railway transport, a lever to improve rail competitiveness, in: EPRS, European Parliament, April.
- Stark, R., Fresemann, C., Lindow, K., 2019. Development and operation of digital twins for technical systems and services. CIRP Annals 68, 129–132.
- Stettinger, G., Benedikt, M., Thek, N., Zehetner, J., 2013. On the difficulties of real-time co-simulation, pp. 989–999.
- Waser, J., Ribicic, H., Fuchs, R., Hirsch, C., Schindler, B., Bloschl, G., Groller, E., 2011. Nodes on ropes: A comprehensive data and control flow for steering ensemble simulations. IEEE transactions on visualization and computer graphics 17, 1872–1881.
- Ye, C., Butler, L., Calka, B., Iangurazov, M., Lu, Q., Gregory, A., Girolami, M., Middleton, C., 2019. A digital twin of bridges for structural health monitoring.
- Yun, S., Park, J.H., Kim, W.T., 2017. Data-centric middleware based digital twin platform for dependable cyber-physical systems, in: 2017 Ninth International Conference on Ubiquitous and Future Networks (ICUFN), IEEE. pp. 922–926.
- Zhang, S., Dong, H., Maschek, U., Song, H., 2021. A digital-twin-assisted fault diagnosis of railway point machine, in: 2021 IEEE 1st International Conference on Digital Twins and Parallel Intelligence (DTPI), pp. 430–433.
- Zhu, H., 2019. Smart maintenance railway infrastructure in Norway. Master's thesis. NTNU.