Rail4Future

Project title:	**Resilient Digital Railway Systems to enhance performance**
Start date:	**01/04/2021**
Time duration:	**42 months**
Project number:	**882504**
Announcement:	**8. Ausschreibung COMET Projekte 2019**

Deliverable D1.1.3

Multi Domain Model Description

Due date	31.08.2022
Submission date	16.02.2023
Submitted by	TU Wien MIVP

Version	Date	Edited by	Description
0.1	31.01.2023	Ozan Kugu	Creation
0.1	31.01.2023	David Kern	Complements

Deliverable released

Dr. Manfred Grafinger

by Area Scientific Manager

Deliverable released

Dr. Filip Kitanoski

by Project Scientific Manager

1 Executive Summary

In this deliverable, a multi domain model is shown and described, which connects all simulations (multiphysics, machine learning) with its relevant data and the required semantic to interconnect all characteristics of the model. As the name suggests, it includes more than one domain, which are used to integrate and interoperate different assets, belonging to different subsystems such as track, vehicle, turnout, bridge, with each other in a DT platform called Rail4Future. Besides, the whole model includes various software tools, sub-processes, sub-environments and hardware components, which basically help to build the virtual environment and make it functional for the asset integration and processing purposes. Moreover, it is very important to design the integration process in order to make all the digital sub-models and data ready for their visualization to the end user, which is also handled in this document. Furthermore, semantic can be added to the data to promote interoperability between different stakeholders and distributed data sources for the DT platform in order to overcome the data exchange difficulties. These difficulties come from differences in stakeholders, heterogeneous data sources, data models and sharing ways. Ontologies are one possible solution of adding semantic to data. The ontology enables the user to discover all the information that exists to a specific asset and describe this asset in a machine readable way. Another advantage of the ontology enables a reasoner to ensure the integrity of the simulation through automated inference of the simulation topology.

This deliverable provides an overview about the multi domain model as a conceptual IT system architecture, which is possibly used to apply the use cases by integrating and interoperating their belonging digital sub-models and data into the R4F platform with help of software, hardware, integration approach and ontology engineering.

2 Table of Content

3 Abbreviations and Acronyms

Abbreviations / Acronyms	Description
MBS	Multi-Body Simulation
SysML	Systems Modeling Language
MIVP	Maschinenbauinformatik und Virtuelle Produktentwicklung
R4F	Rail4Future
DT	Digital Twin
VTI	Vehicle Track Interaction
RLT	Remaining Lifetime
ViF	Virtual Vehicle Research GmbH
IT	Information Technology
SCM	Source Code Management
DBMS	Database Management System
UI	User Interface
RDF	Resource Description Framework
RDFS	Resource Description Framework - Schema
OWL	Web Ontology Language
NOR	non-ontology Resources

4 Problem Description / Objectives

4.1 Problem Description

For the digitalization of the holistic large-scale railway infrastructure system, it is important to virtualize the train drive with subcomponents such as rail vehicle, track, turnout, tunnel, which is to be integrated into the R4F platform and the virtual subcomponents are then to be interacted with each other inside the platform in a proper way. In order to succeed it, the importance to design the IT system infrastructure by implementing software and hardware cannot be denied, so that the asset integration and processing in the virtual environment of the railway infrastructure system can be ensured to provide visual and user-friendly service to the end user at the end. On the other hand, there are data exchange difficulties, which come from differences in stakeholders, heterogeneous data sources, data models and sharing ways. The ontology can overcome these difficulties to promote interoperability between different stakeholders and distributed data sources for the R4F platform by adding semantic to the data.

4.2 Objectives

In this deliverable, first, it is aimed to demonstrate an asset integration methodology by proposing a multi domain model, which consists of different software tools, sub-processes, sub-environments, hardware and semantic to build the IT system architecture of the R4F platform. The assets are also provided from asset owners and providers as digital sub-models and datasets, which are then to be adapted into the platform with the ontology from semantic to interconnect model characteristics, input parameters and outputs inside the multi domain model to apply different use cases such as MBS vehicle dynamics, VTI, RLT calculation of a steel bridge, model reduction and machine learning.

The main purpose of the deliverable is to prove the qualification of the proposed multi domain model as an asset integration methodology, which is then to be implemented into the R4F platform and therefore supposed to help to virtualize the railway infrastructure system in a proper way. At the end, the end user can benefit from the platform for condition monitoring, diagnostics and predictive maintenance purposes in the train operation area.

5 Significance for the overall Project

First, it is important to have the virtual version of the subsystems, belonging to the holistic large-scale railway infrastructure system, to digitalize the system in the R4F platform. Then, the digital assets are to be integrated and interoperated with each other in a proper way. To make it happen, it is important to provide a useful and understandable system architecture for the asset integration and processing by using and implementing different software tools, technologies, semantic and hardware in the platform. As a result, all the provided assets could be then well-prepared to visually present the holistic large-scale railway infrastructure system to the end user in a virtual environment under consideration of user-friendliness, easy-to-use and simulation reliability at the end.

This deliverable gives an overview and important aspects about the multi domain model to the reader, which is then to be applied to the R4F platform as basics for the railway digitalization in future. It will also help to adapt the railway infrastructure system to the virtual world in a proper way, which would significantly reduce the lifecycle costs of the railway infrastructure system as a huge benefit of the R4F project.

6 Description

The multi domain model of the R4F platform is based on an IT system architecture, which is described in the conference paper of Zhou et.al. [1] in details and consists of different software tools, sub-processes, sub-environments and hardware. All the subcomponents of the model are then to be communicated and interacted with each other inside the platform, so that the assets including digital models and dataset from different use cases can be integrated and interoperated with each other to visually represent the holistic large-scale railway infrastructure system to the end user.

In this section, first, the typically used software tools and hardware components are shortly mentioned, which help to make the multi domain model. Second, the integration process is shown and demonstrated with different concrete software tools, sub-models, dataset and sub-processes. Lastly, the semantic modeling and ontology development processes are mentioned in details, where the ontology can describe different assets in the multi domain model, therefore helps to reduce the data exchange complexity and ease the information discovery for different stakeholders about the assets belonging to the railway infrastructure system in the R4F platform.

6.1 Software and Hardware

In this subsection, first, the main useful software tools such as Jenkins, Git and MySQL are specified. After that, the typically used hardware components are shortly mentioned, which are used to work with the software tools.

Jenkins is open-source, user-friendly and tool-independent software with useful plugins and used to provide the Jenkins pipeline technology, which is durable, pausable and enables the continuous integration of different digital sub-models and dataset into the platform as an automation engine. Besides, Git provides GitLab as version control system, which helps the asset providers and integrators to work with different codes, input and output files belonging to the sub-models and dataset of different use cases collaboratively. All the files can also be uploaded into a Git repository and Jenkins pipeline can directly be connected to the repository of the Git server through the SCM system as discovered before. In addition, MySQL is implemented in the integration environment as a relational DBMS, which can extract, archive and store raw data from the sub-models and their belonging input and output files, so that the whole dataset can be prepared for the visualization to the end user.

To use the software tools, they have to be installed to a computer first, where the user can work with the tools directly. The computer is a typical PC as hardware, consisting of a working station, monitor, mouse and keyboard. The working station is supposed to ensure high-performance for large-scale simulations, where tracks of overall hundreds of kilometers are to be passed by a rail vehicle, as a concrete automation engine. The monitor, mouse and keyboard surely help the user to work with the software tools user-interactively as a whole.

6.2 Integration Process

In this subsection, the integration process of the multi domain model to integrate all the above mentioned models and data into the R4F platform is described in details. First, the whole integration process of different use cases is demonstrated as shown in Fig. 2. Additionally, Fig. 3 shows the typical integration approach for the VTI use case followed by the ViF.

In Fig. 2, the typical integration process by applying the above mentioned software tools is demonstrated, which includes graph database, different digital sub-models and datasets of different use cases and automated model visualization additionally. The graph database provides optimal searching functionality by extracting many different sub-model characteristics from the provided digital assets by the Jenkins pipeline and then generating the pipeline with different sensibly linked graphs. The research on the semantic modeling is already done to find out a solution to utilize all the information from the graph database with the functionality (see section 6.4 for further information). The other in Fig. 2 shown components such as FMU, SCM, JSON to provide interface between the software tools and digital assets are going to be mentioned in details in the next deliverable 1.1.4 (M20).

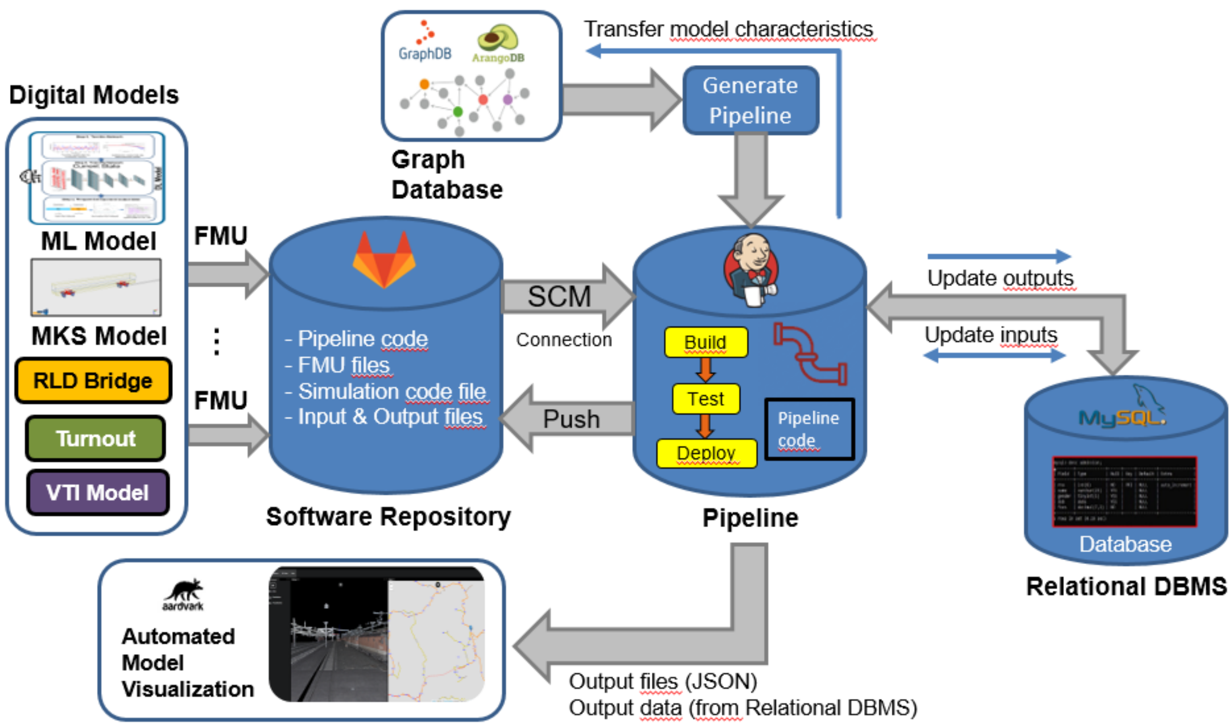

Fig. 2: The demonstrated integration process of different assets. (© TU Wien MIVP)

Additionally, the first implementation of an integration approach into the R4F platform was previously demonstrated for the VTI use case by the ViF as shown in Fig. 3. Like in the previous figure (Fig. 2), the graph generation from the ArangoDB graph database is used to ensure optimal searching inside the use case for the end user. Besides, the Jenkins server is applied to work with the Jenkins pipeline technology as well, which is also responsible for the simulation workflow of the use case. Moreover, it is planned to automatically generate SysML models including the whole simulation process of use cases as a graph, which is then to be transformed into a Jenkins pipeline to ease the execution of simulations belonging to the use cases in future. After the whole asset integration and processing, all the relevant outputs are to be written into a visualization platform called Aardwark, which is provided by the VRVis, and the input parametrization is to be ensured through UI at the same time, so that the use cases can be presented to the end user in a user-friendly virtual environment.

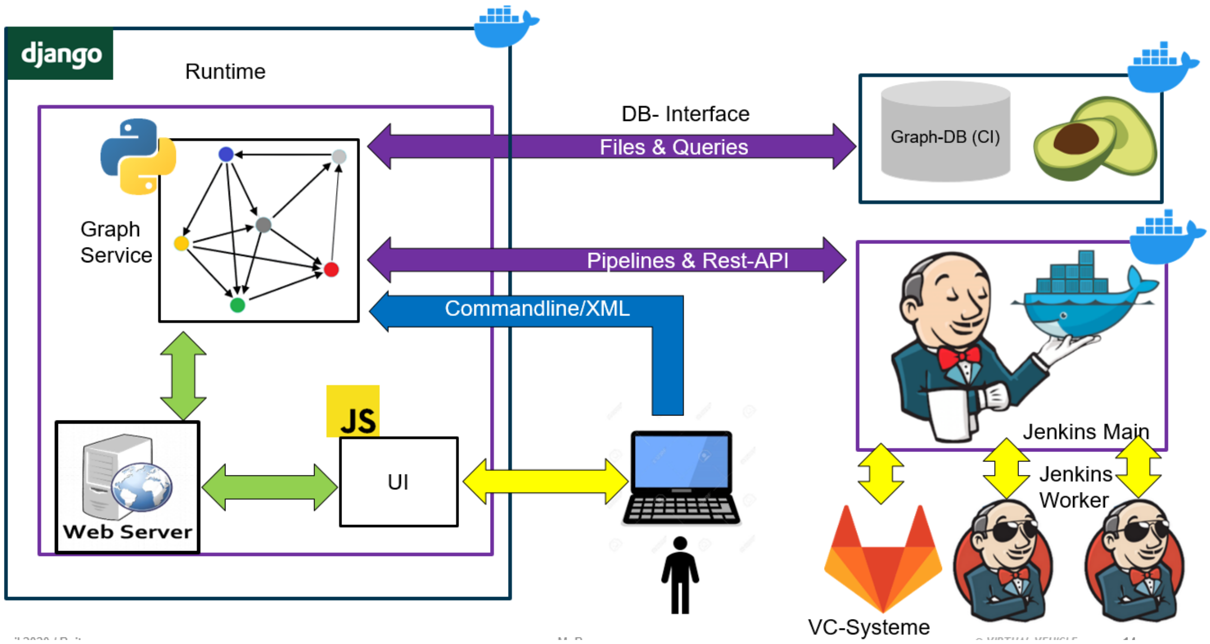

Fig. 3: Integration approach for the VTI use case. (© ViF)

6.3 Semantic Modeling and Ontology Development

Interoperability plays a crucial role in information management as most of our current information needs involve drawing data from multiple sources owned by different stakeholders. Multiple sources or different stakeholders often also mean multiple different interfaces and data models. However, to augment the value of the available data, semantic context information should be attached to it. With this semantic interoperability some different problems arise such as the so-called semantic clash, where two or more concepts or terms being used in a specific context or system have different meanings or interpretations, leading to confusion or misinterpretation. For example, in one context, the term "load" might refer to the weight or force that a structure or component is designed to bear, while in another context, "load" might refer to the amount of energy being used by a machine or system. One possible solution to overcome this problem is the usage of ontologies. The term ontology originates in philosophy and is a branch of metaphysics which is concerned with the nature of existence, including the relationships between entities and categories. In computer science, the term refers to the representation of knowledge as a set of concepts within a specific domain and the relationship between these concepts. The usage of an ontology provides a couple different advantages. With a common, shared vocabulary expanding and integrating new information sources can be added without need of modification to the other sources or their vocabularies. Only new terms and relations need to be added to the shared vocabulary. Another advantage is that implicit knowledge can be made explicit.

For the development of the ontology, semantic technologies such as the RDF, RDFS and OWL are used.

The ontology development process adapts elements of the NeOn[1] methodology. NeOn is scenario-driven ontology engineering framework and uses the so-called divide and conquer strategy, by breaking down the general problem into smaller sub-problems. The Framework defines nine different scenarios based on which ontologies can be developed as seen in Fig. 4. A combination of the following scenarios is used to develop the ontology.

- Scenario 1: From specification to implementation.
- Scenario 2: Reusing and re-engineering of non-ontological resources
- Scenario 4: Reusing and re-engineering ontological resources

Each scenario defines activities to guide the development process.

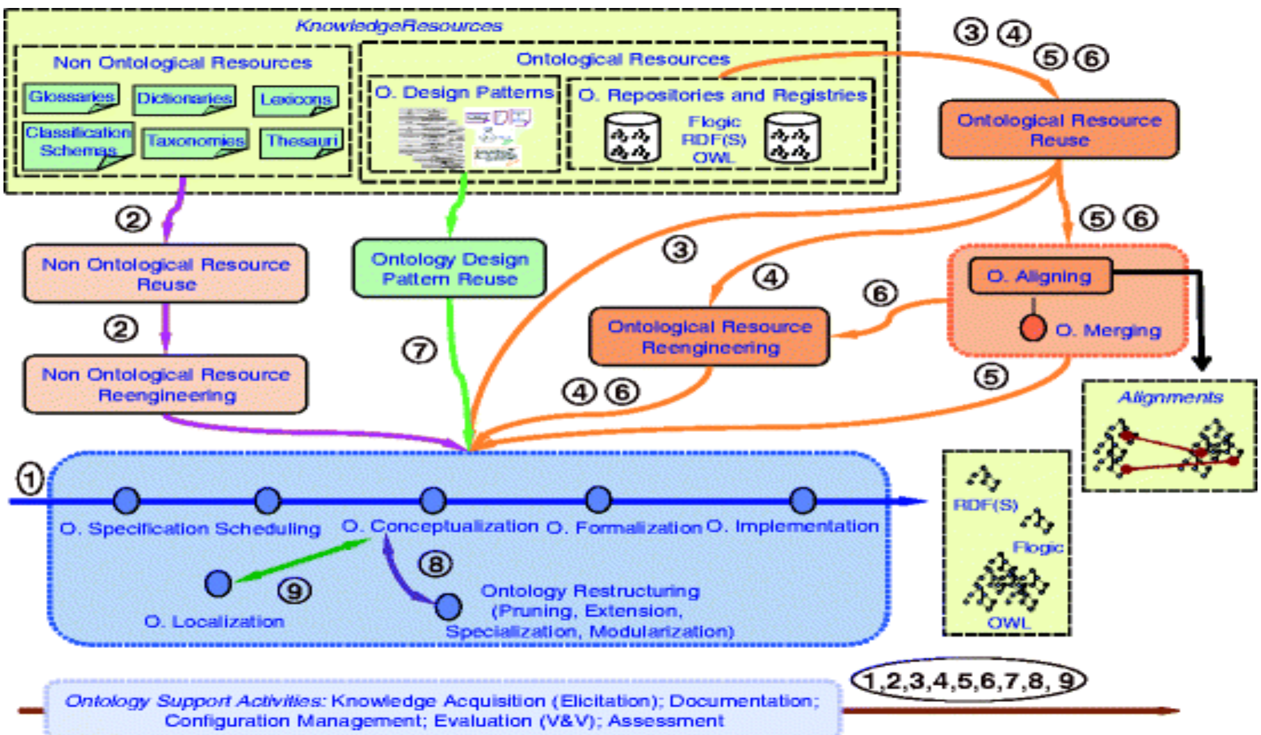

Fig. 4: Scenarios for building ontologies[2].

---

[1] M. C. Suárez-Figueroa, A. Gómez-Pérez, and M. Fernández-López, "The NeOn methodology for ontology engineering," in Ontology Engineering in a Networked World, Springer Berlin Heidelberg, Dec. 2011, pp. 9–34. doi: 10.1007/978- 3- 642-24794-1_2

[2] Ibid.

Scenario 4 defines activities to reuse and re-engineer existing ontologies. For the current ontology development process some of the following ontologies were reused and re-engineered.

- **QUDT[3]:** presents a unified architecture for the conceptual representation of quantities, quantity kinds, units, dimensions, and data types. It is largely based on the international standard for metric units (SI) but also includes other systems. It allows for Unit conversion from and to base SI standard.
- **FOAF[4]:** is used to describe different Agents and their relations to one another. It allows to model different stakeholders, as well as individual persons.
- **vCard[5]:** extends the possibilities of the FOAF vocabulary to describe Agents and enables to model contact information and addresses.
- **dcterms[6]:** is used to model additional metadata for digital or physical resources.
- **spdx[7]:** provides properties and classes to model communicating software bill of material information, including licenses, copyrights, and security references.
- **DCAT[8]:** is used to describe datasets and data services. Originally this is done in a catalogue, for the present work this is re-engineered to fit the needs.
- **Adms[9]:** enables a version control of the described assets.

Scenario 2 describes the process needed to include NOR into an ontology. NOR used in this work include terms from the FMI Specification. One advantage of the re-use of terms from the standard, is that mapping from FMU-files to the ontology is much simpler.

As a final development step the missing classes and attributes are defined according to the activities described in scenario 1. Here the prefix rff is used to denote those terms. The class RealWorldObject can be as a mapping point to other ontologies such as the ERA[10] project to describe the railway infrastructure in more detail. The final ontology can be seen in Fig. 5.

At its core the Rail4Future ontology defines a SimulationModel concept which is described by its used VariablesGroup as well as the RealWorldObject it describes. Additionally, some metadata is added to the simulation model in order to model the developer/maintainer of that model.

[3] https://qudt.org/

[4] http://xmlns.com/foaf/0.1/

[5] https://www.w3.org/TR/2014/NOTE-vcard-rdf-20140522/

[6] https://www.dublincore.org/specifications/dublin-core/dcmi-terms/

[7] https://spdx.org/rdf/terms/

[8] https://www.w3.org/TR/vocab-dcat-2/

[9] https://www.w3.org/TR/vocab-adms/

[10] https://data-interop.era.europa.eu/era-vocabulary/

Supporting documents, such as a documentation or any document relevant to that simulation model can also be added.

The VariableGroup concept describes a group of variables used in a simulation model. The property 'isConnectedWith' models the connection between an input and an output variable group. This connection is only valid if the two groups have a different causality. A VariableGroup can be published as a part of a Dataset which itself can have multiple different versions, modelled with the adms ontology. Since the publisher of a dataset does not necessarily have to be the publisher of the simulation model, a separate connection to model such cases is included. The Distribution concept represents a specific representation of a dataset. A dataset might be available in multiple serializations that may differ in various ways. This difference can be modelled with the fileformat property form the dcterms vocabulary. Here the media type definition by IANA[11] is strongly encouraged. The spdx:checksum property provides a mechanism that can be used to verify that the content of a file or package have not changed. A distribution of a dataset can be obtained via a DataService, that provides access to one or more datasets. Finally contact information to different agents can be modelled using the FOAF and vCard vocabulary.

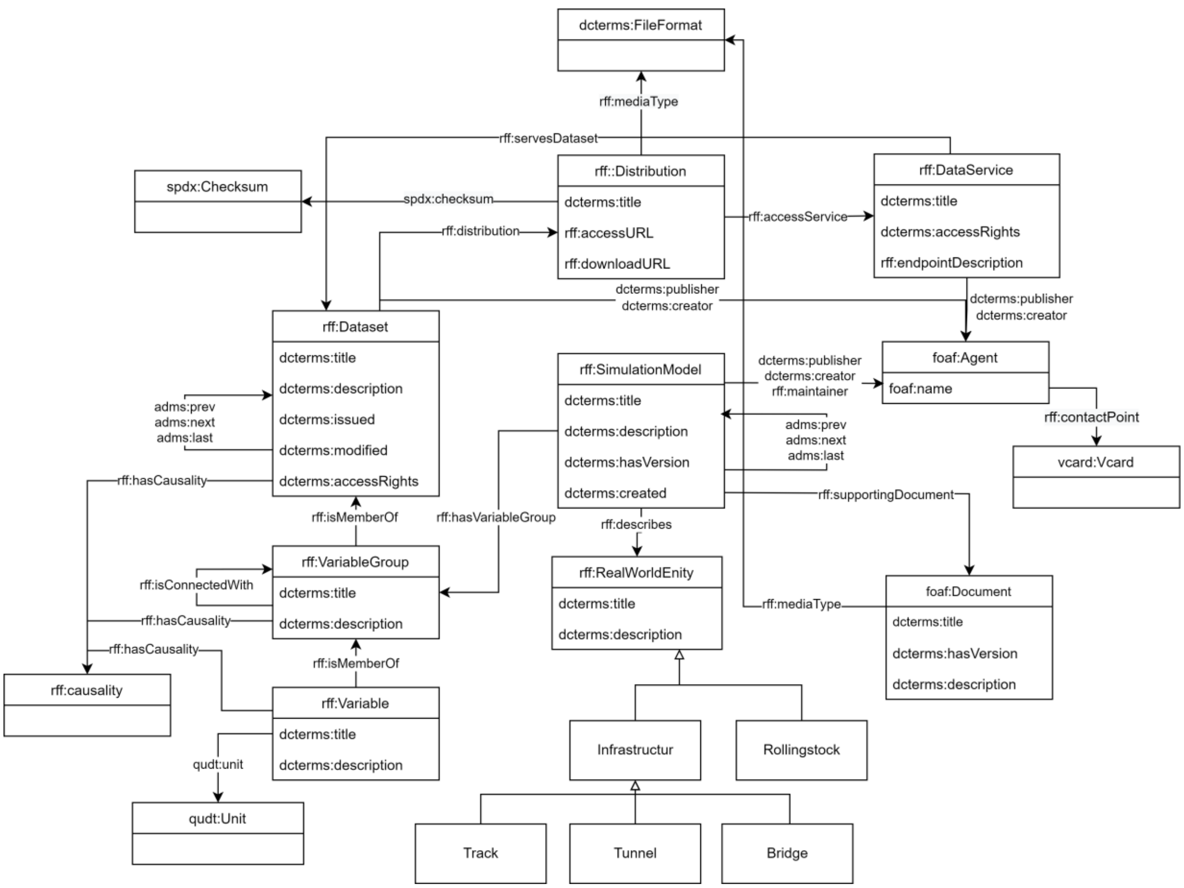

Fig. 5: Class Model of the RFF Ontology.

[11] https://www.iana.org/assignments/media-types/media-types.xhtml

With the described ontology it is possible to describe different assets used in the Rail4Future Domain and make these information machine readable as well as help different stakeholders discover all the information available. It also can be used to help the automation process. In order to assist the automation process, a variable-group-based approach is taken. However, this approach is still in development and needs to be validated.

7 Conclusion

This deliverable aimed to make an overview about a multi domain model demonstrating the IT system architecture of the holistic large-scale railway infrastructure system in a virtual environment. After configuring inputs, outputs, dependencies and physical forms of different sub-models and describing the ontology of the data, their subsequent integration into the R4F platform and interoperation with each other can then be realized based on the designed system architecture, which is also a challenging and elaborate task inside the R4F project. In the next step, some useful interfaces and adapters are to be provided for the task. The integration process is also going to be improved and optimized in the platform with help of the interfaces and adapters, which aims to enhance a continuous and semi-automated integration and deployment of the whole digitalized railway infrastructure system in future.

8 References

[1] Zhou, S., Dumss, S., Nowak, R., Riegler, R., Kugu, O., Krammer, M. and Grafinger, M., 2022. A Conceptual Model-based Digital Twin Platform for Holistic Large-scale Railway Infrastructure Systems. Procedia CIRP, 109, pp.362-367, https://doi.org/10.1016/j.procir.2022.05.263.