Project Reference: 636285
H2020-MG 2014-2015
Innovations and Networks Executive Agency
Project Duration: 1 May 2015–31 April 2018

Information Management System Framework
D3.1

Authors

*** Zaharah Allah Bukhsh, Timo Hartmann, Irina Stipanovic**

*Corresponding author: z.allahbukhsh@utwente.nl

Date: 10/02/16
Dissemination level: (PU, PP, RE, CO): PU
This project has received funding from the European Union's

This project is funded by the European Union

DOCUMENT HISTORY

Number	Date	Author(s)	Comments
1	24/01/2016	Irina Stipanovic (IS)	IR sends the first version of the document to Vijay Ramdas (VR) for review
2	06/02/2016	Vijay Ramdas (VR)	VR sends the reviewed document back to IS
3	10/02/2016	Irina Stipanovic (IS)	IS sends the document to Ken Gavin (KG) for review
4	10/02/2016	Ken Gavin (KG)	KG sends the reviewed document back to IS
5	10/02/2016	Irina Stipanovic (IS)	Final report complete
6	04/11/2016	Irina Stipanovic (IS)	Fix a formatting error at page 9, Final report complete.

Executive Summary

The DESTination RAIL project aims to support railway infrastructure management by developing a number of novel techniques and systems for identifying, analyzing, predicting and remediating the critical rail infrastructure. As its final goal, DESTination RAIL aims to integrate the different techniques into a decision support system that will allow for the holistic management of all assets of a rail network. To allow for such integration, the decision support system needs to be supported by a common data repository that collects the outputs, but also the required inputs of the different identification, analysis, prediction, and remediating techniques and systems. This report proposes an information model to allow for such integration.

The Information Model (IM) was developed following a three step iterative approach. In a first iteration, an initial draft version of the IM was designed based upon the initial DESTination Rail project proposal. This draft version incorporated all the data items as described by the different project tasks. In a second iteration, the IM was then adjusted based on information found in the seminal literature about railway engineering and modelling. Noteworthy knowledge source used was the RailML standard to model railway infrastructure objects, time tables, and driving stock. Additionally, another important source of information during this step was the book by C. Esveeld, Modern Railway Track (2001). In a third iteration, interviews were conducted with a number of project partners of the DESTination rail project. During these interviews the second version of the IM was used as an input to trigger constructive and important feedback. Adjusting the IM with respect to the feedback led to the final version of the IM as presented in this report.

The IM presented in this document includes objects to store structured information about railway infrastructure, rail load management, rail failure sources, rail risk assessment methods, and rail maintenance methods.

Because of the systematic development effort summarized above, the structure and the chosen level of detail is based upon the needs and requirements of the DESTination RAIL project partners. The IM has already been used in the process of requirement elicitation to stimulate active discussion between project partners. Additionally, the developed IM will be used as a blueprint for the implementation of the final information management system into the developed decision support system.

Table of Contents

1 Introduction

According to the health and safety report of the European Rail Agency published in 2013 “train is the safest mode of transport” (Agency, 2013). The fatality risk of a train travelling passenger in the EU is three times lower than that of the passenger travelling in bus / coach (Sedghi, 2013). While the safety record of railway travel is highly encouraging, a number of high profile railway failures have occurred recently. Notable incidents were, among others, the derailment of a Swiss train in August 13th, 2014 due to rainfall (see Figure 1)[1], the damage of the mainline at Dawlish, UK due to a coastal flood in February 2014 (see Figure 2)[2], or the collapse of the Malahide viaduct near Dublin, Ireland in August 2009 due to bridge scour. These incidents highlight the effect of climate change and aging infrastructure on railway safety. The events, in turn, signal the challenges faced by EU rail infrastructure managers with respect to the maintenance of their existing assets.

Figure 1: Derailed train caused by landslide in Swiss Alps, in August 2014

Figure 2: Destroyed railway line by storm, Dawlish, UK, in February 2014,

To address these pressing needs, the Horizon 2020 funded project DESTination RAIL aims to support infrastructure managers by developing a number of novel techniques and systems for identifying, analyzing, predicting and remediating rail infrastructure. The project also intends to design an integrated decision support system that can combine all the developed techniques and systems. To fulfil the needs with respect to data input and output, and interoperability between the different techniques and systems, Task 3.1 of the project is concerned with the development of an Information Management System (IMS) to hold the required data relating to the individual railway assets for a specific network to support decision making.

[1] http://www.dailymail.co.uk/wires/ap/article-2723919/11-injured-train-cars-derail-Swiss-Alps.html

[2] http://www.bbc.com/news/uk-26042990

This document introduces an Information Model (IM) as a basis for the IMS. The IM has been carefully designed based upon the information in the initial call text, a focused literature and state of the art review, and a series of interviews with key project partners of the DESTination RAIL project. The designed IM is intended to serve as the structure for designing database containers that organize, store, and retrieve data, including translation into useful information for a specific railway network with the aim of supporting all the innovative techniques and systems that will be developed in the DESTination RAIL project.

The report is structured as follows: The next section briefly defines the IM and provides the basic considerations for designing an IM. Section 3 introduces the approach for designing the IM. Section 4 then provides a detailed description of the complete IM for railway infrastructure that we have developed. Final sections provide discussion, conclusion and future work directions.

2 Information models

An IM can be defined as a simplified representation of a small, finite subset of the world. As such, each of the objects within a model corresponds to some real or abstract object that might exist in the world or within the state of mind of a group of persons or an individual (Kent, William and Hoberman, 2012; Turk, 2001). While representing such objects a number of decisions need to be made about the level of detail to represent each object. For example, the degree of decomposition of the objects to be represented, how to represent the relationships among objects, and also how to deal with changes to objects over time (Kent, William and Hoberman, 2012). How these decisions are made does not depend on any natural law, but is largely determined by the designers of a specific information system. These decisions could, for example, be driven by the kind of information required for the specific purpose of the information model and considerations about how that will be maintained while in use (Kent, William and Hoberman, 2012).

Each IM can ultimately only represent a certain aspect of the world using a certain categorization and a certain level of detail (Turk, 2001). Therefore, before an information model can be developed and the decisions about the specific form of IM should be discussed, the universe of discourse (Turk, 2001) and a purpose for the model need to be identified. After identification, the particular considerations about the required qualities for the model to support the purpose at hand can be made.

Therefore, the development of an IM should be considered as a continuous design process which needs to iterate through steps of requirements analysis, system design, implementation, as well as, monitoring and maintenance (Teorey, 1999). As an overarching requirement, the universe of discourse of the IM presented here is related to the management of rail infrastructure accounting for possible risks and network effects. The IM then needs to consist of objects, relationships, rules, constraints and operations in a way that describe the railway management domain purposefully (Lee, 1999). To develop such a domain-specific information system requires in-depth domain knowledge and this was gained while designing the IM. Section 3 describes the development process in detail.

3 Development approach

At the outset, to design the IM presented in this document we adopted a conceptual, specification driven perspective. We did not try to develop an IM for exactly modelling and mapping all details of a real railway infrastructure network. Instead, we focused on software design requirements as they are related to each of the techniques and systems that will be developed in the DESTination RAIL projects. Objects of the IM can describe physical entities of the railway infrastructure, such as track or bridge, or conceptual entities related to railway asset management tasks, such as, failure reasons. In addition to representing the physical objects, the IM also represents important attributes of these objects and relationships between the objects. To keep the IM simple, we have omitted operational details, such as train schedules or railway stocks, for now. The Unified Modelling Language (UML) has been used for supporting the model development process and for representing the model (Lee, 1999).

The design effort itself underwent three iterations. During each iteration, we identified and refined objects and their properties related to the universe of discourse at hand. We started the design iterations by developing an initial version of the IM capturing all the concepts and objects that were mentioned in the original DESTination RAIL project proposal. In a second iteration, this initial IM was then enriched through acquiring knowledge of railway infrastructure from various other sources. First and foremost, we consulted one of the state of the art railway engineering textbooks Modern Railway Track (Esveld, 2001). Additionally, we also consulted RailML an existing standard for describing railway infrastructure assets, railway time tables, and railway building stock. Identification of important domain related terms led to the specification of objects, their properties, their inter-relatedness and their overall operations within a second version of the IM. Based on acquired domain knowledge, a second draft of the IM was developed.

To further improve the IM with respect to the needs of the DESTination RAIL project, semi-structured interviews were organized with twelve key project partners from nine different institutes. All of these participants are responsible for different work packages and tasks of the DESTination RAIL project and, hence, possess knowledge of different requirements for the IMS. We used the draft IM developed through the two iterations described above to provide an example to the interviewees, to trigger thoughts and to invite critical review. We also used a semi-structured questionnaire (see Appendix B) to guide the interviews. The questionnaire consisted of questions related to railway infrastructure and the data requirements of interviewee. We conducted the interviews over Skype and telephone. Interviews lasted between 30 to 60 minutes. Appendix C provides additional detail about the information setup and participants. The interviews led to numerous comments on the conceptual IM from the second iteration and we used these comments to update the IM concept from the second iteration to arrive at the final IM described in Section 4.

4 Information Model

The final IM of railway infrastructure designed is divided into the following domains (Figure 3): railway assets, including their condition details, maintenance history, failure sources, and failure history. The following subsections describe each of these domains in detail.

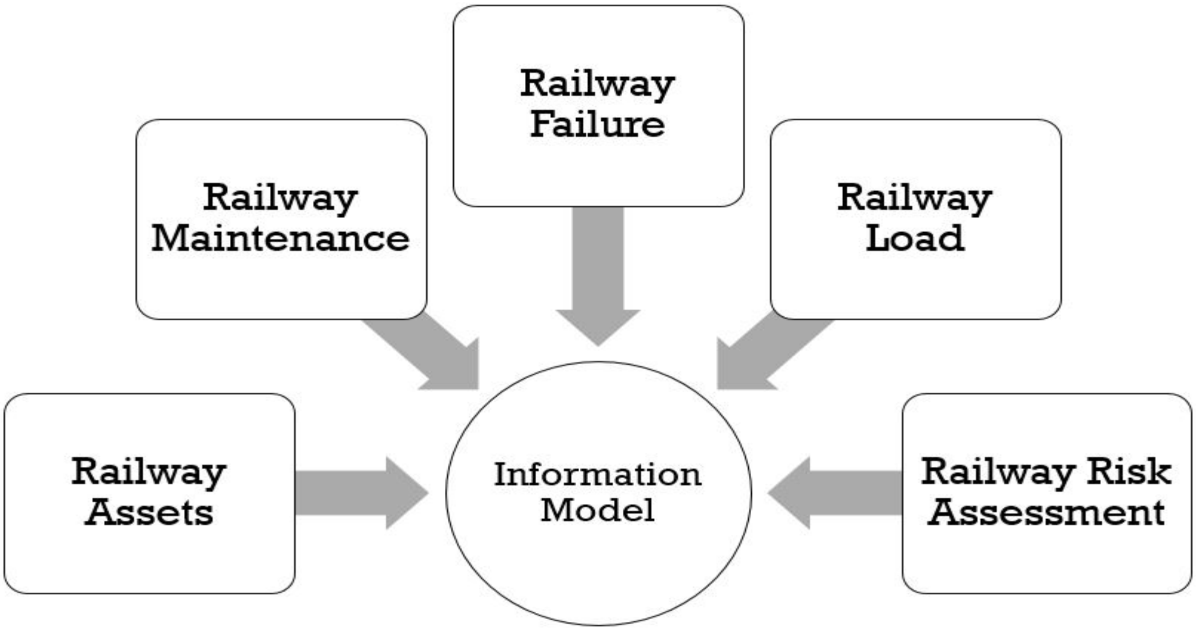


Figure 3: Overview of the IM domains

4.1 Railway infrastructure assets

Figure 4 shows the objects of the IM representing the main components of rail infrastructure. The *rail infrastructure* is composed of *track*, *platforms*, *signaling* and *electrification system*. With respect to project needs, a number of other aspects are connected to rail infrastructure. The logical connections among the components are represented by UML class relationships (Nishadha, 2012). Each rail infrastructure component has one or more associated *construction techniques* that can be used to store different maintenance and replacement methods for each of the components. Similarly, each infrastructure element can have zero or more *hotspots*. Hotspots can be defined as those parts of the railway that are vulnerable to fail. To understand possible failure sources, one or more *risk assessment method* can be associated with each component as well. Moreover, to allow for the representation of different failure reasons for each component, the IM allows the tracking of possible problems with a class named *failure source*.

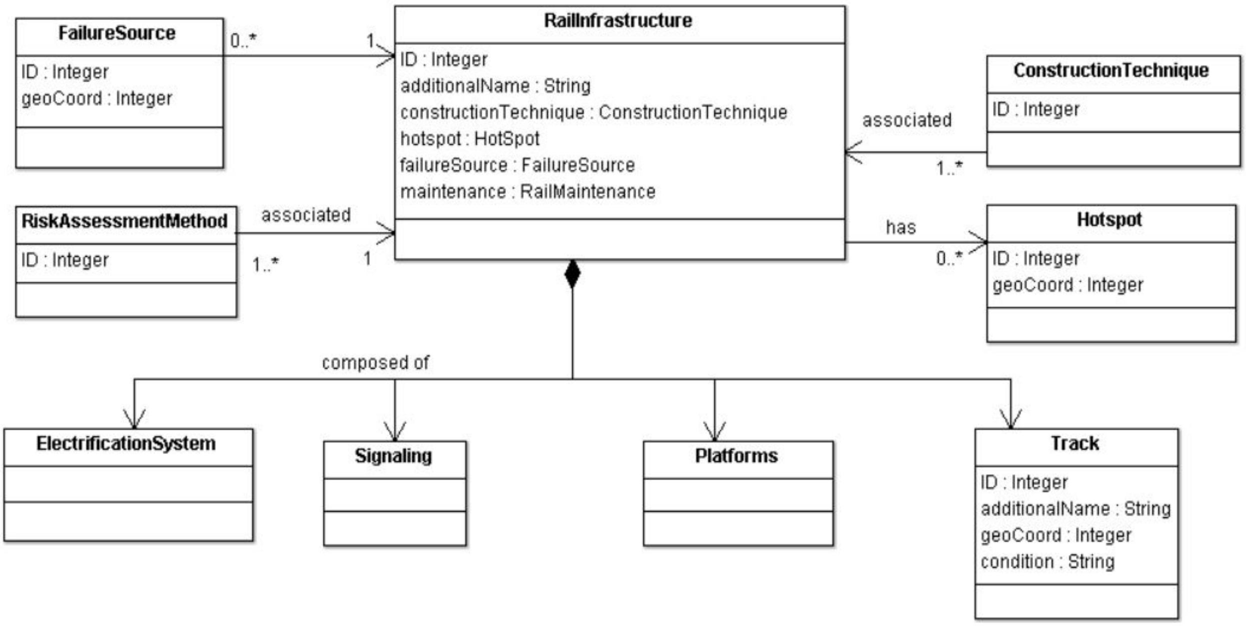


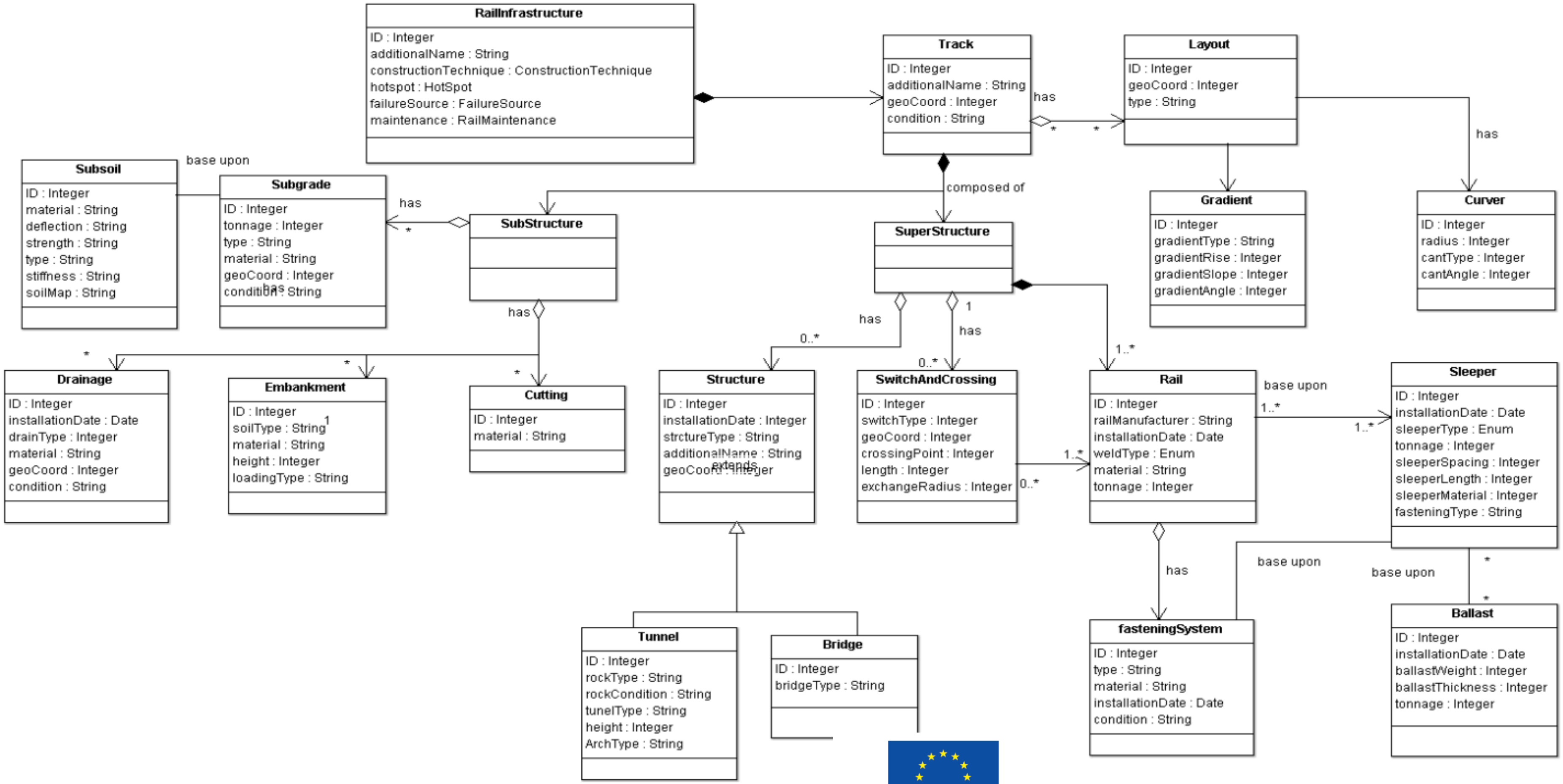
Figure 4: Information Model (Rail Infrastructure)

Among the components, *track* is the integral component of the rail infrastructure. A *track component* is further subdivided into a number of sub-components as illustrated in Figure 5 Components of *tracks* are divided into *superstructure* and *substructure*. *Superstructure* includes all those components that are above the foundation and the *substructure* provides the details of underlying support components.

Every railway track has a specific *layout*. We have defined *layout* using objects for track *curves* and *gradients*. *Curves* represent the curve measurement of the track in degrees of its radius. To support the high speed and minimal friction over the curve, the outer rail is usually raised so that the outer rail and inner rail have different elevations. This difference is named cant. Each track curve is defined to have different cant types and cant angles, reflecting the passing train speed. Finally, the *track gradient* is the relative elevation of two rails along the track. We intend to store the gradient rise, gradient slope and raised angle of each track gradient in the system.

The *substructure* of the track is composed of the underlying support components. The IM can include the properties of *cutting, embankments, drainage* and *subgrade*. A *substructure* can have zero or more *cutting* and underlying *embankment*. It is necessary to keep the information regarding the drainage system updated because most track failures occur due to malfunctioning of the drainage system. The *subgrade* is further based on *subsoil*. It is important to be able to specify *subgrade and subsoil* behaviour under various load cases. Considered properties here are *strength, deflection, stiffness, soil type, soil map and soil material*.

Figure 5: Information Model (Track)



This project has received funding from the European Union's

This project is funded by the European Union

The *Superstructure* of the track consists of all those components that are built over the *substructure*. The key component of the *superstructure* is the *rail* itself. A *rail* object is further decomposed into *sleepers*, *ballast,* and *fastening system*. Important information about *rail* is data about rail manufacturer, its installation date, rail material, weld type and its tonnage capability. Similarly, the IM allows detailed representation of properties for *sleeper*, *fastening system* and *ballast*.

As shown in Figure 5, a *track* can have zero or more *switches and crossings*. Each switch or crossing can include additional information regarding its type, crossing point, length and exchange radius. Additional *structures* can be represented, such as, track *tunnel* or *bridge*. Each *track* element can have zero or more of these *structures*.

4.2 Rail maintenance management

Figure 6 shows the main object classes important for *rail maintenance management* acounted for in the IM. As a result of *rail maintenance* work, rail operations may need to be halted or speed restrictions applied. Regarding the particular spot maintenance, a number of important items of information will be stored. The data about the location of the spot, maintenance date, maintenance cost, maintenance type, maintenance duration, maintenance history and imposed speed will be stored. Most important properties are maintenance type, spot history and imposed speed. The maintenance type can be represented as renewal, grinding, tamping, or others. Additionally, the IM allows storage of data about the spot history and all the maintenance activities that have been conducted at the particular spot in the past. As maintenance can affect train operations, it is useful to keep the data about the imposed speed during the maintenance.

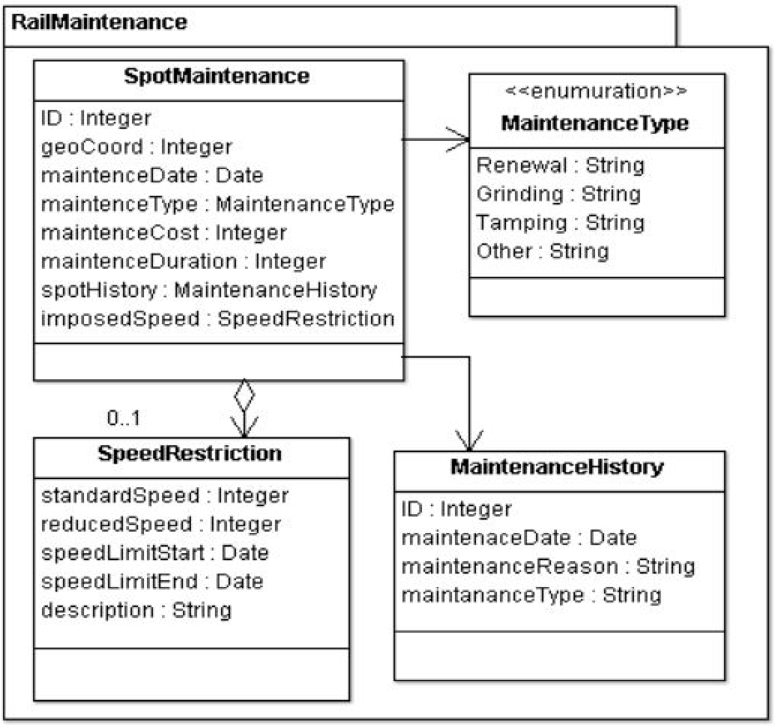


Figure 6: Information Model (Rail Maintenance)

4.3 Rail Failure

Figure 7 shows the main components and data properties related to *rail failure*. *Failure source* is the main object of rail failure management. Information about the location of failure, failure reasons, failure severity, failure cause and failure history can be represented using the IM. Additional information regarding the type of failure and failure cause can be assigned to each rail failure element, for example, to determine how common certain failures are. Similarly, information regarding the history of failure will be useful to take adequate mitigation approaches to avoid failures in future.

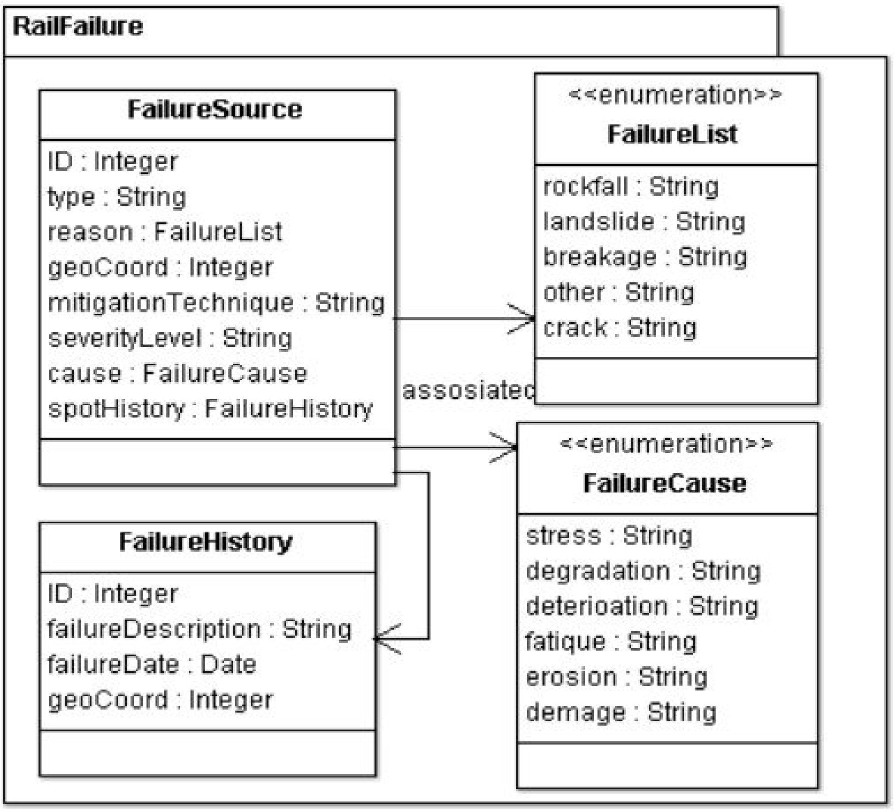


Figure 7: Information Model (Rail Failure)

4.4 Rail load management

Figure 8 shows the main type of *railway loads*. It is important to store load data as the behaviour of substructures and superstructure can vary under different loads. Information regarding various vehicle types, weight of load, axle load, load speed, load length, caused deflection, force and load type would be stored. Several other properties of load can also be stored in the system, if needed.

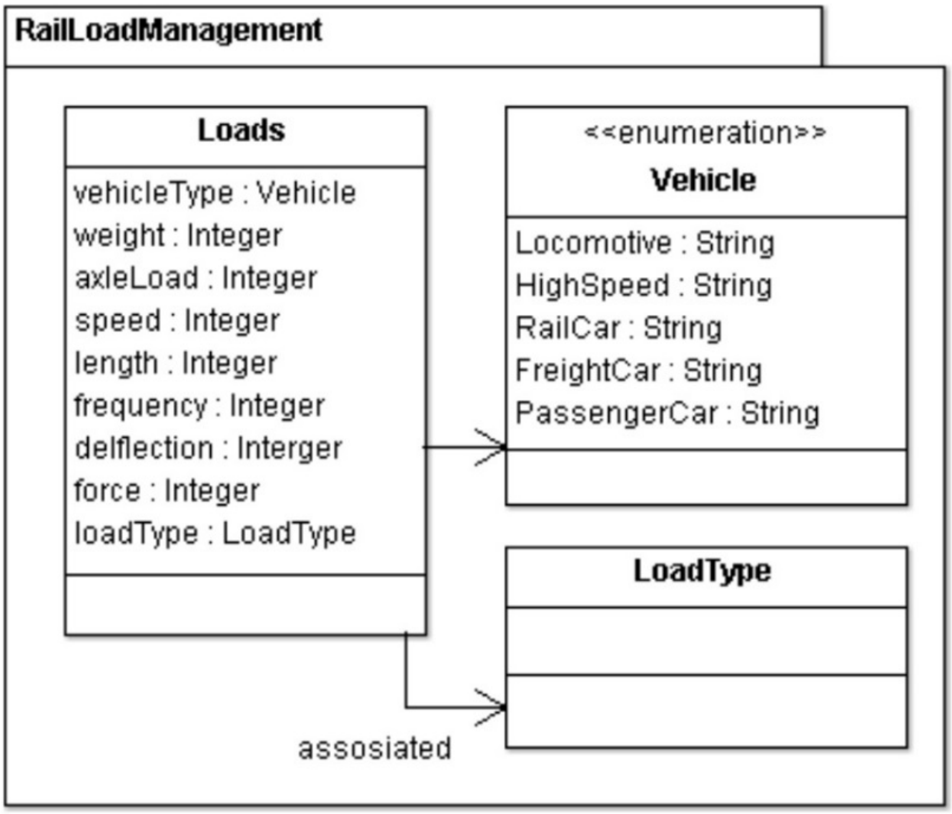


Figure 8: Information Model (Rail Load Management)

4.5 Information Model of Rail risk assessment method

Figure 9 shows how the IM can represent the important concepts of the general risk assessment method adopted from ISO (2009). The risk assessment method consists of risk identification, risk analysis and risk evaluation processes, each of which can be represented. In the risk identification process, the possible risks, their effects, and the analysis techniques can be recorded. In the risk analysis process, further information related to risk factors, risk likelihood, risk consequences, risk levels, and risk causes, among others can be included. Based on the risk analysis and risk evaluation methods, a risk evaluation method suggesting risk mitigation treatment forms part of the IM. For a single risk, it is possible to assign a number of risk treatments.

The risk assessment method presented here is of generic nature and can be applied to any infrastructure, system, and organization. A risk assessment method specific to railway infrastructure will be developed later in the project. It is important to reflect that properties defined for each object can be added and modified even after the system is developed as the requirements from the system might evolve as the project progresses.

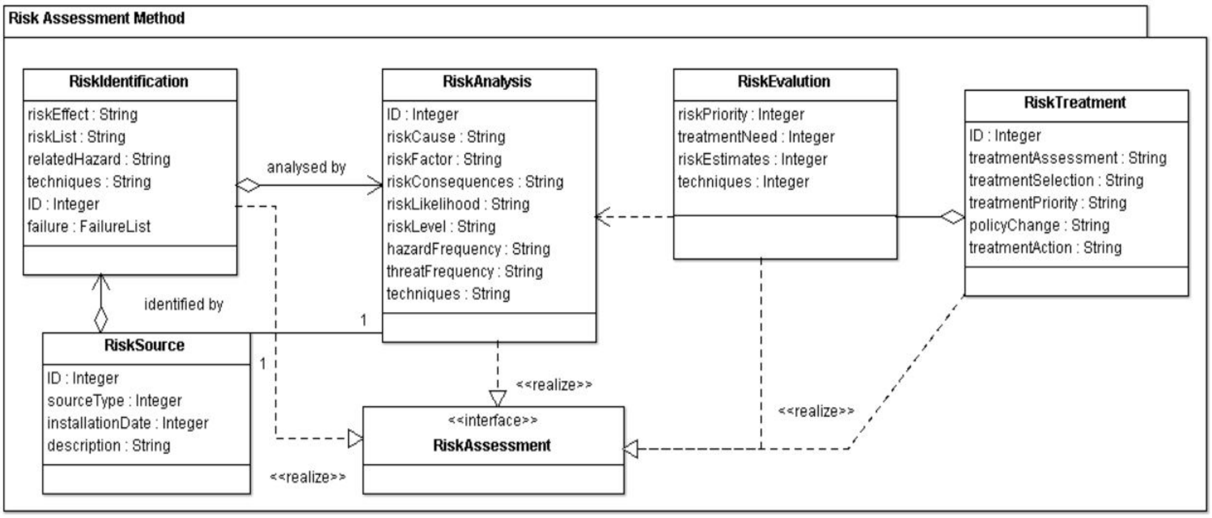


Figure 9: Information Model (Risk Assessment Method)

5 Discussion

The IM described in this report provides a basic layout for the information management system development and, in turn, for the final decision support system. Through the structured design approach that included input from the key partners of the DESTinationRAIL project, we consider the IM is sufficiently comprehensive to integrate the innovations developed within the Find, Analyse, Classify, and Treat work packages of the DESTination Rail project. We also believe that the proposed IM provides a natural way to support decisions of rail infrastructure managers at each level of their network.

Immense amounts of data on railway operations and asset condition are gathered through a number of means, e.g. by sensors, drones, ground penetrating radars, etc.. Therefore, sensor readings and streams need to be pre-processed before they can be stored in the IMS. The pre-processing of data needs to be performed based on its future use and/or the requirements of techniques that will exploit the available data. The above described storage requirements for dealing with sensor data are not within the scope of the here presented IM.

At this stage, information regarding required data types (e.g. int, double, string, enum) for the implementation of the IM in software is not available. In one of the next steps, we will closely examine data from a railway agency to decide upon the data types to use.

Further, the model's current scope does not include any information about operations of a rail network, such as schedules or rolling stock. Such information (WP4), however, will be important to predict the effects of the risks of failure and maintenance on the overall network. We chose this solution because the existing standard RailML is able to represent schedule and rolling stock information and can interface with existing network simulation software. Therefore, one future development step will be the implementation of an interface between the IM presented here and the RailML standard.

As with any IM, many concerns can be raised based on the selection of objects, their considered properties and their interrelationship among them. Modifications to the model are only to be expected. However, it is important to consider the difference between the proposed IM and its implementation approach. Proposed IM provide a static impression of defined components with number of other design possibilities. Modification or addition of a components or relevant property in proposed IM will lead to several changes. To support the design changes, we are taking into consideration number of advance database implementation techniques e.g. key-value, column store, graph-based, document-based. In contract to tradition database techniques, these advance database techniques allow the flexible schema development, where the schema model can be modified even during development phase. Therefore, implementation of the IM as presented in this report will be open for flexible adjustments. At the same time, all relevant concerns by project partners should be raised as soon as possible, ideally before system development is started.

6 Conclusions and Future Work

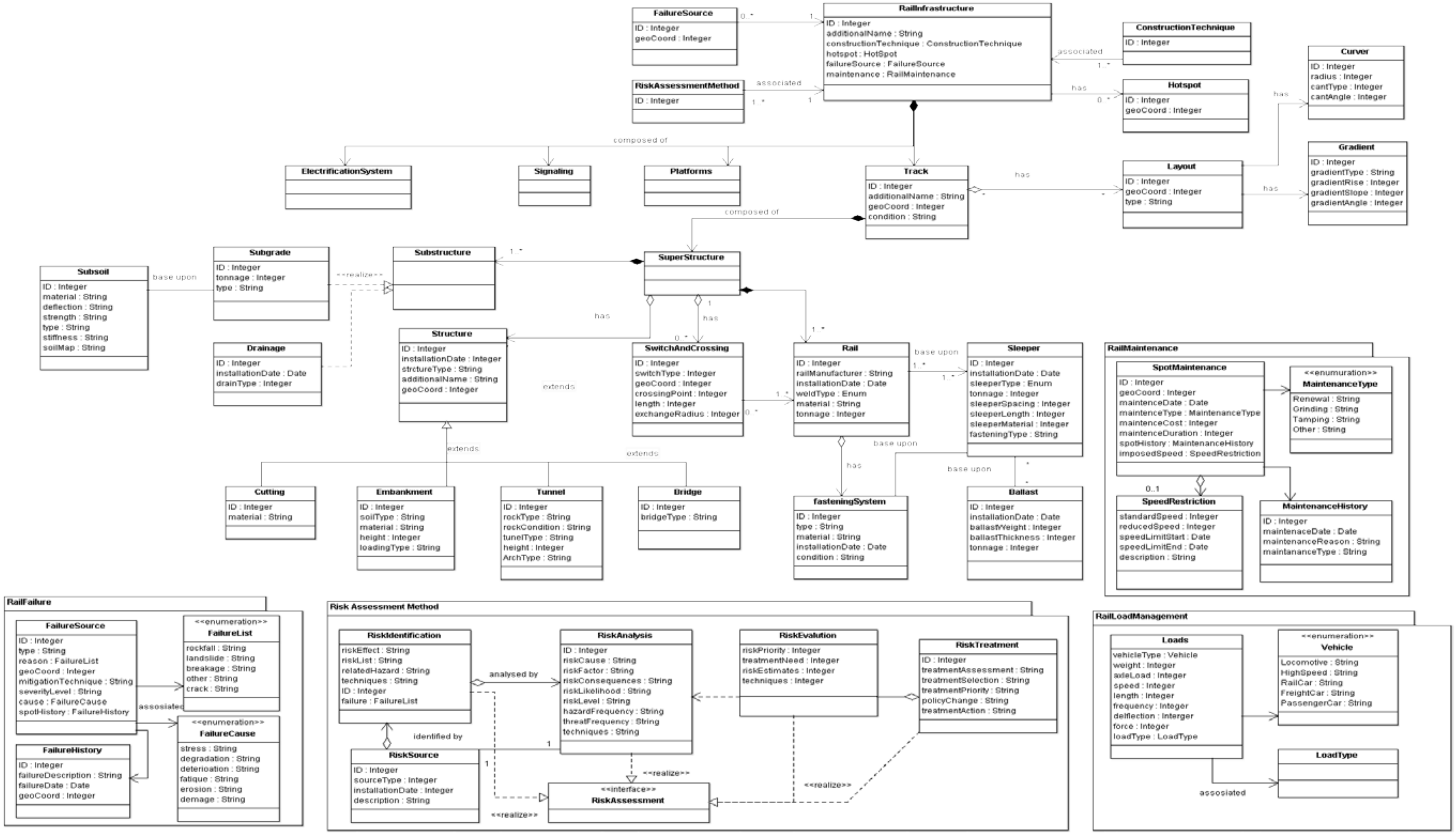
The IM of railway infrastructure provided in this document is part of an on-going effort to better structure, communicate and ultimately develop an information management system for DESTination RAIL project, which can also be used in the future at railway agencies. Instead of just focusing on railway assets and railway operations, the IM presented here can store data about all the important aspects of railway maintenance, railway assets, railway failures, railway loads and railway risk assessment.

The implementation of the IM within an information management system is the next step we will conduct within this task. For this purpose, we are in the process of conducting an experiment to compare various database management systems for their scalability and performance efficiency.

At the same time, since the innovations developed in the DESTination RAIL project are of a dynamic nature, the IM will need to be adapted to changing requirements if required. Therefore, implementation solutions need to be flexible in nature.

We are also working to define the conceptual framework of the decision support tool, risk assessment methods, life cycle cost models and all the innovative solutions/methods that are under development as a part of the project. The conceptual framework will help us to foresee the aspects of interface interactions among these diverse methods, techniques and systems.

7 Appendix A : Information Model



8 Appendix B: Questionnaire

Interview Questions

Please list all physical objects (e.g. Bridges, Tracks, Slopes) of the rail system and its surroundings that are relevant to your task/ model!

Questions about Infrastructure

For each of objects please list their important parts!

What are the functions of the objects within the overall rail system?

How would you describe the location of these objects within the overall rail system?

How do you think these objects are related to each other (in your task/model)?

What other information are important to know about these mentioned objects with respect to your task/model?

Questions about Relevance to IMS

Considering the objects mentioned, what information do you think is most important to keep/record in the IMS?

In what format are these information or (ideally) will these information be collected/available?

Do you have any suggestion/preference for system implementation? Specific tools and technologies?

How do you envision to work with the DESTinationRAIL decision support system?

Please be as specific about the information input and output the IMS would provide and in which format that information would ideally be exchanged.

9 Appendix C: Interview Setup

WP	Partner	Country	Interviewee	Date (2015)	Duration (Minutes)	Medium
Slopes (Soil)	GDG (Gavin and Doherty Geosolutions)	Dublin, Ireland	Luke Prendergast	16th Oct.	25	Skype
Slopes + Past experiences	GDG (Gavin and Doherty Geosolutions)	Dublin, Ireland	Karlo Martinovic	19th Oct	25	Skpye
Slopes+Rocks	UZ (University of Zagreb)	Zagreb, Croatia	Mario Bacic	3rd Oct.	65	Skype
Bridges	UT (University of Twente)	Enschede, the Netherlands	Neda Mostafa, Richard Loendersloot		55	Face-to-face
Track	TU Munich	Munich, Germany	Bernhard Lechner	3rd Nov	40	Skype
Risk assessment framework	ETH	Zurich, Switzerland	Claudio Martani, Papathanasiou Natalia	9th Oct	40	Skype
Switches and crossings	NTNU (Norwegian University of Science and Technology)	Trondheim, Norway	Elias Kassa	19th Oct	40	Skype
Bridges	RODIS	Dublin, Ireland	Mark Tucker, Lorcan Connoly	15th Oct	35	Skype
Design remediation measures	ZAG	Ljubljana, Slovenia	Stanislav Lenart	2nd Nov	25	Skype
Life Cycle Cost Model	TRL (Transport Research Laboratory)	London, UK	Vijay Ramdas, Jonathan Sharpe	2nd Nov	40	Telephone

10 References

Agency, E. R. (2013). *Intermediate report on the development of railway safety in the European Union 2013*.

Esveld, C. (2001). *Modern Railway Track*. (D. Z. Nieuwenhuizen & L. J. van 't Zand, Eds.) (2nd ed.). MRT-Productions.

ISO, I. (2009). 31000: 2009 Risk management--Principles and guidelines. *International Organization for Standardization, Geneva, Switzerland*.

Kent, William and Hoberman, S. (2012). *Data and reality: A timeless perspective on perceiving and managing information in our imprecise world*. Technics publications.

Lee, Y. T. (1999). Information modeling: From design to implementation. In *Proceedings of the second world manufacturing congress* (pp. 315–321).

Nishadha. (2012). Class Diagram Relationships in UML with Examples. Retrieved from http://creately.com/blog/diagrams/class-diagram-relationships/

Sedghi, A. (2013). How safe are Europe's railways? Retrieved from http://www.theguardian.com/news/datablog/2013/jul/25/how-safe-are-europe-railways

Teorey, T. J. (1999). *Database modeling & design*. Morgan Kaufmann.

Turk, Ž. (2001). Phenomenologial foundations of conceptual product modelling in architecture, engineering and construction. *Artificial Intelligence in Engineering*, *15*(2), 83–92. http://doi.org/10.1016/S0954-1810(01)00008-5