From Analysis of Information Needs towards an Information Model of Railway Infrastructure

3 authors:

Zaharah Allah Bukhsh
University of Twente

Timo Hartmann
Technische Universität Berlin

Irina Stipanovic
University of Twente

From Analysis of Information Needs towards an Information Model of Railway Infrastructure

Zaharah Allah Bukhsh[1], Timo Hartmann[2], and Irina Stipanovic[1]

[1]Department of Construction Management and Engineering, University of Twente, Enschede
[2]Chair of System Engineering, Institute of Civil Engineering, TU Berlin
z.allahbukhsh@utwente.nl, timo.hartmann@tu-berlin.de, i.stipanovic@utwente.nl

Abstract. Railway is a tightly coupled network, where the operations are directly effected by the condition of rail infrastructure. With the advancement of ICT, a railway network exploit various computerized systems for efficient railway monitoring, maintenance and operations. However, these systems suffer from number of limitations, mainly, the data related to each asset type (e.g. Track, Bridge, etc) are stored in separate database management system. Such scattered and isolated nature of data present the island of information, while making it impossible to perform the sound decision analysis. In this paper, we propose a nework wide information model of railway infrastructure that structure the railway object, specify their properties and identify their inter-relationships. The presented information model supports the railway monitoring, maintenance and operations by providing the layout of railway infrastructure. Structuring data in the form of railway assets, railway risk assessment, railway load management, railway maintenance, and railway failure will provide a solid base to railway stakeholders, e.g. infrastructure managers, to take informed decisions based on data properties.

1 Introduction

Railway network depends on information and communication technology (ICT) for day-to-day operations and infrastructure maintenance. Many computerized information solutions have been developed and deployed in the railway agencies to support the efficient railway operations, monitoring and maintenance. Immense amount of data is stored in the railway systems to keep track of assets' condition, network performance, maintenance records, etc . However, the data related to each network aspect are usually stored in separate database management systems (see Thaduri et al. (2015)), which expose the challenge of data integration (Galar et al., 2012). Such unavailability of data make it difficult to perform the sound decision analysis for cost-effective railway monitoring, maintenance and operations.

This research effort is a part of DESTination RAIL project, which seeks to develop a decision support tool based on network wide information management system. One of the main focus in the development of information management system is to understand the end-users data needs for the decision analysis. To deal with the challenges of data isolation and to comprehend the data needs, we choose to exploit the information modeling approach to design the information management system of railway infrastructure. Among other techniques, information modeling structures the problem domain and provide its semantic interpretation. An information model provides a neutral data model, with varying level of detail, which provides support in data exchange and data maintenance. Thus, the objective of this paper is to design the detailed information model of railway infrastructure, which cover monitoring, operation and maintenance aspects. An iterative design approach has been followed, where the draft version of an information model is used in the process of requirements elicitation to stimulate the active discussion between project partners. Furthermore, semi-structured interviews, directed by a questionnaire, were conducted in order to examine and outline the specific data requirements of project partners. The main contribution of this work lies in presenting an information model, which not only represent the railway objects, but draft the monitoring, operations and maintenance aspects of the railway in the form of railway assets, railway risk assessment, railway load management, railway maintenance and railway failure. The final information model will be used as a blueprint for the development of a network wide information management system. Moreover, we believe that the presented information model is comprehensive and generic enough to be used for the development of railway infrastructure-related data applications.

The rest of the paper is structured as follows: the state of the art of information modeling in ACE domain is outlined in Section 2. Our development approach to design the information model of railway infrastructure is provided in Section 3. The detailed information model is discussed in Section 4. Finally, Section 5 and Section 6 provide the discussion and conclusions, respectively.

2 Related Work

To support the interoperability between softwares, development of information modeling and sharing protocols is trending in ACE domain. An information model 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 (Turk, 2001; Kent and Hoberman, 2012). As explained by Zamanian and Pittman (1999), “information models formulate the concise description of the real-world artifact that can be communicated and processed effectively”.

Few problem-specific information models have been suggested in the literature with respect to railway domain and in general for the ACE domain. Bosschaart et al. (2015) proposed an UML class representation of railway interlocking data in order to provide a common knowledge/data base of machine readable data formats. To reduce the cost and inefficiency of data exchange, UML classes are mapped to RailML. RailML is a XML based exchange format that ease the data exchange process among the railway applications (Kolmorgen and Huerlimann, 2005). Borrmann et al. (2014) followed the information modeling approach to design the track alignment and procedural geometric models. These data models are further used to facilitate the data sharing and synchronous modeling via collaboration platforms. To deal with challenges of railway data integrations from various data sources, Wang et al. (2012) proposed a neutral meta-data dictionary model. The proposed meta-data dictionary model is used for the development of XML-oriented three-dimensional data models where data objects are grouped into various types. Vossebeld and Hartmann (2014) proposed to use an information model to support part of safety assessment task for tunnels, instead of managing comprehensive dossiers with tunnel critical information. Lucas et al. (2013) developed a comprehensive object-oriented model and an ontology to help facility managers in managing the patient safety events and facility operations in an efficient manner.

There is increasing focus on the design of information and data models to facilitate efficient data exchange and to define data inter-relatedness. As noticed with aforementioned studies, information models are mainly contexts-specific and usually act as an initial step in the development of a computerized solution. Hence, it is important that information models are well developed, cover the problem domain sufficiently and are understood by their perspective end-users. In this paper, we also opted to consider the end-users’ perspective for the development of information model for the domain of railway asset management, in order to understand information needs for decision analysis and to communicate the system design. Furthermore, the resulting information model will be a useful addition to the literature of railway engineering as no general railway infrastructure model was found during the literature search.

3 Information Modeling Approach

To design an Information Model (IM), we adopted a conceptual, specification-driven perspective. The presented IM is not aimed to exactly model and map the railway infrastructure network. Instead, it is targeted to specify and understand end-users' data needs by identifying the railway network objects and their relationships. Objects of the IM can be understood as physical entities of the railway infrastructure, such as tracks or bridges, or conceptual entities related to railway asset management tasks, such as risk types, failure reasons. In addition to identifying objects, the developed IM also represents important attributes of these objects and relationship between objects. To keep the IM simple, operational details, such as train schedules or railway stocks have been omitted. The Unified Modeling Language (UML) is used for supporting the model development process and for representing the model (Lee, 1999).

The design effort underwent three main 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 a first initial version of the IM capturing all the concepts and objects that were mentioned in the project proposal of DESTination RAIL. In a second iteration, this initial IM was then enriched by 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 consulted RailML an existing standard for describing railway infrastructure assets, railway timetables, 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. This led to the second version of IM.

To further improve the IM, semi-structured interviews were organized with twelve key project partners from nine different institutes. These participants belong to various railway agencies and research institutes and possess considerable experience working in the particular railway domain either as a researcher or engineer. These participants will also be the first end-users of the IM. The structuring of railway objects in IM is crucial, since all the innovations and methods developed in DESTination RAIL will depend on it. We used the draft IM developed out of the first two iterations to provide an example to the interviewees and to trigger thoughts and critical reviews. Moreover, we develop a questionnaire to guide the interview process. The questionnaire consists of questions related to railway infrastructure and asks questions about the data requirements of each participant. Interviews were conducted over Skype and telephone that lasted between 30 to 60 minutes. Overall, the interviews led to numerous comments on the conceptual IM from the second iteration. We used these comments to update the IM concept from the second iteration to arrive at the final IM described in the Section 4.

4 Information Model for Railway Asset Management

Railway infrastructure is comprised of physical objects, e.g. tracks, rolling stocks, etc. along with their functional capabilities and features. For the end-to-end railway management and smooth operations, three main aspects, i.e. monitoring, operation and maintenance of railway objects are required. By railway monitoring, current conditions of objects are assessed and required measures are taken. To store the data gathered from condition monitoring, properties of railway objects, their location and their geometrical details should be defined. We design an IM of railway infrastructure assets, which will store the data about the railway objects along with their current conditions, provided in Section 4.1. To assess the conditions and need of maintenance, an IM to support the risk assessment is introduced in Section 4.2. Operation of railway is core aspects in the railway network management, which consist of traffic control, timetable management, etc. Considering the defined scope of DESTination RAIL, we have modeled only railway load management aspect to capture the affect of rolling stocks on the network in Section 4.3. With respect to maintenance, we designed an IM to support railway maintenance in order to capture maintenance types, schedules, costs and affects on railway operations, presented in Section 4.4. Additionally, we model the IM of railway failures in order to keep a record of failure sources in the network, provided in Section 4.5.

Figure 1 shows the objects of the IM representing the main components of rail infrastructure. The *RailInfrastructure* is composed of *Track*, *Platform*, *Signaling* and *ElectrificationSystem*. Among others, *Track* is the most integral component of the rail infrastructure. The objects related to rail infrastructure are mainly associated with the physical objects (i.e. track, platform, etc) of the railway network. For example, any object of infrastructure can be regarded as *Hotspot*, as it is defined as a part/component/place of the network that is vulnerable to fail. To understand the needs of maintenance, one or more types of *RiskAssessmentMethod* can be associated with each component as well. Moreover, to allow the representation of different failure reasons of each object, the IM allows to track possible problems with a class named *FailureSource*. Similarly, IM can store the data about the spots that are undergoing with certain maintenance operations through *SpotMaintenance*. A *RailInfrastructure* component has one or more associated *ConstructionTechniques* that can be used to store different maintenance and replacement methods for each of the components in the network. Be noted that all of these information models are related to each other in terms of exchange of data. For the sake of discussion and to keep the models readable, we have discussed each model separately in the following sections.

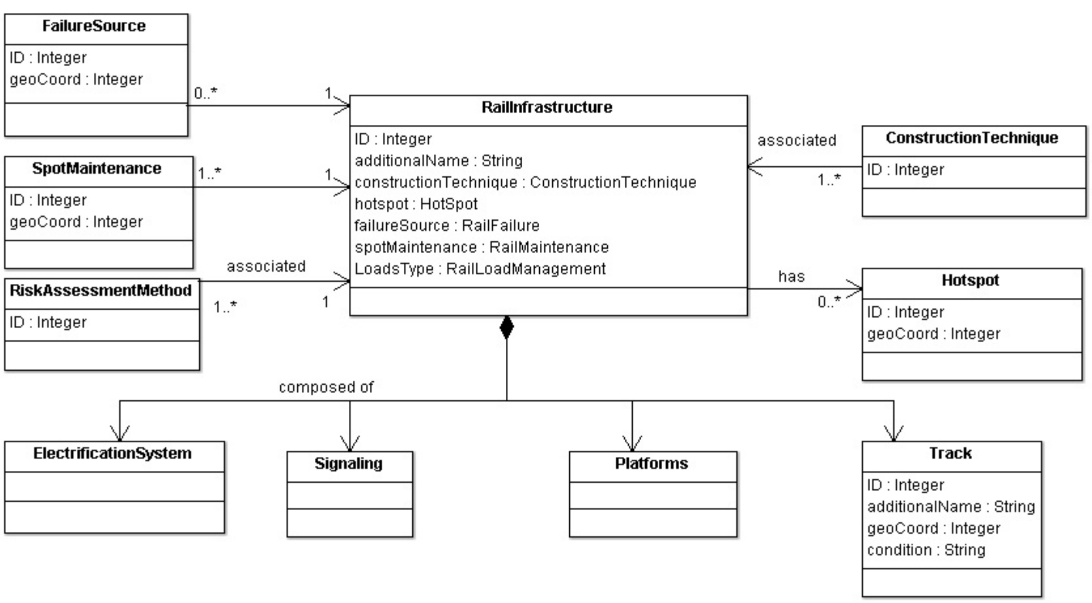

Figure 1: Information Model of Rail Infrastructure

4.1 Model of Railway Infrastructure Assets

A track consists of a number of sub-components as illustrated in Figure 2. Components of tracks are divided into *superstructure* and *substructure*. *Superstructure* includes all those components that are above the foundation. While, *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*. The *substructure* of the track is composed of cutting, embankment, drainage and subgrade. The IM represent the related properties of all these components. 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.

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. Definition of objects along with their properties will be helpful to store the data about these physical components. As shown in Figure 2, a track can have zero or more *SwitchesAndCrossings*. 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.

Along with geometrical and location properties, each object has a condition property which will be used to store the current conditions of asset. This condition property will be updated frequently as a result of monitoring.

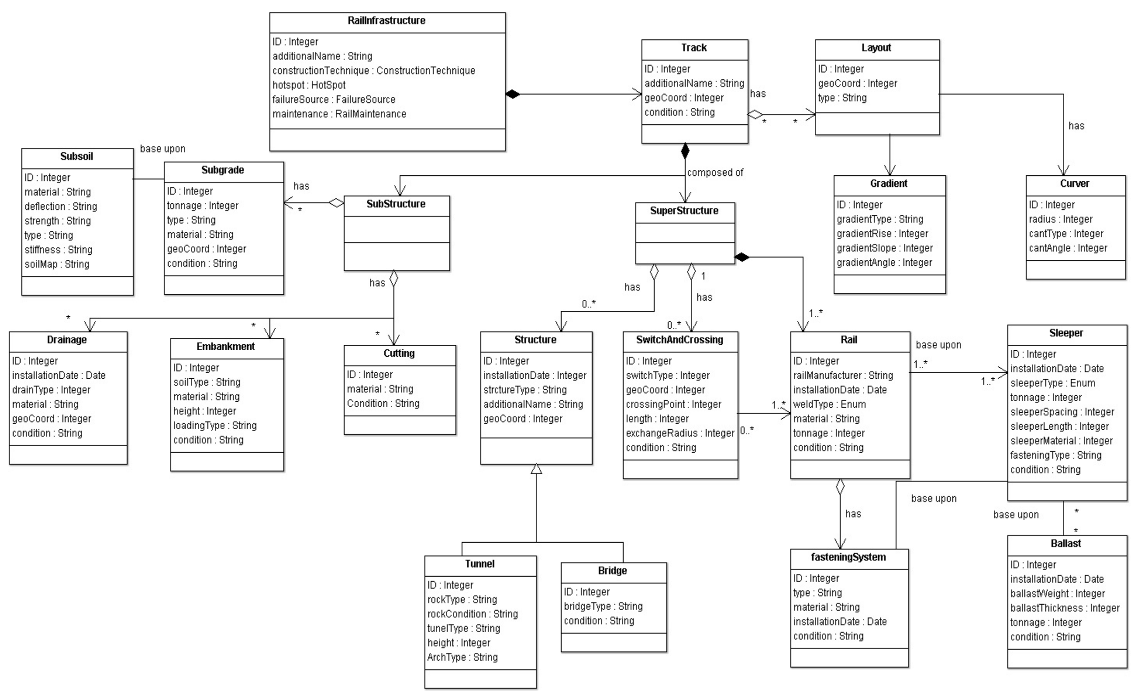

Figure 2: Information Model (Railway Infrastructure Assets)

4.2 Model of Risk Assessment Method

Figure 3 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 classes of *RiskIdentification*, *RiskAnalysis*, *RiskEvaluation* and *RiskTreatment*, each of which can be modeled in IM. The *RiskIdentification* capture the possible risks, their effects, and their analysis techniques based on a particular *RiskSource*. In the *RiskAnalysis* process, the identified risks are further analysed by considering further information of risk factors, risk likelihood, risk consequences, risk levels, or risk causes, among others. Based on the *RiskAnalysis* methods, a risk evaluation method prioritise the risk and suggest treatment requirements. Finally, *RiskTreatment* captures the treatment assessment, treatment selection and treatment priority for the sake of risk mitigation. For a single risk, it is possible to assign a number of risk treatments. IM presented in Figure 3 is able to capture the whole process of risk assessment in a form of defined data entities.

The risk assessment method presented here are of generic nature and can be applied to any infrastructure, system, and organization. The development of risk assessment method specific to railway infrastructure is part of our future work.

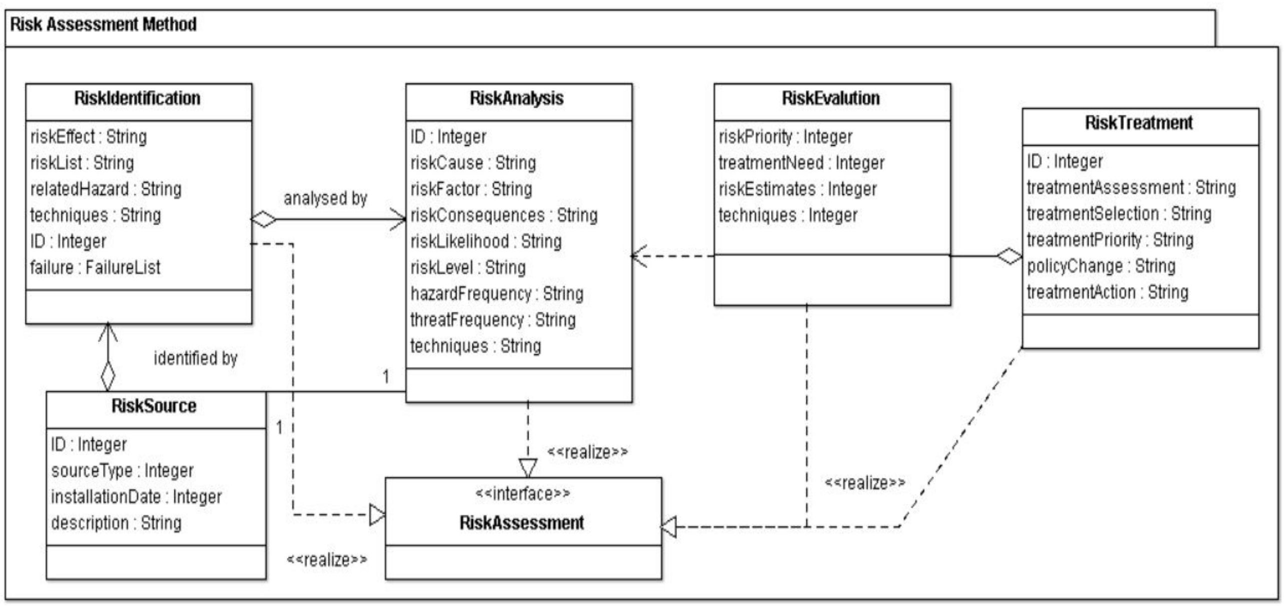

Figure 3: Information Model (Risk Assessment Method)

4.3 Model of Railway Load Management

As discussed earlier, railway load management belongs to the area of railway operations. Data about loads and rolling stock is important to store as it can be used to simulate the behavior of substructures under various loadings. Moreover, information regarding various vehicle types, weight of load, axle load, load speed, load length, caused deflection, force and load type are important properties to be stored. The information model of *RailLoadManagement* is provided in Figure 4. Several other properties of load can also be stored in the system, if needed.

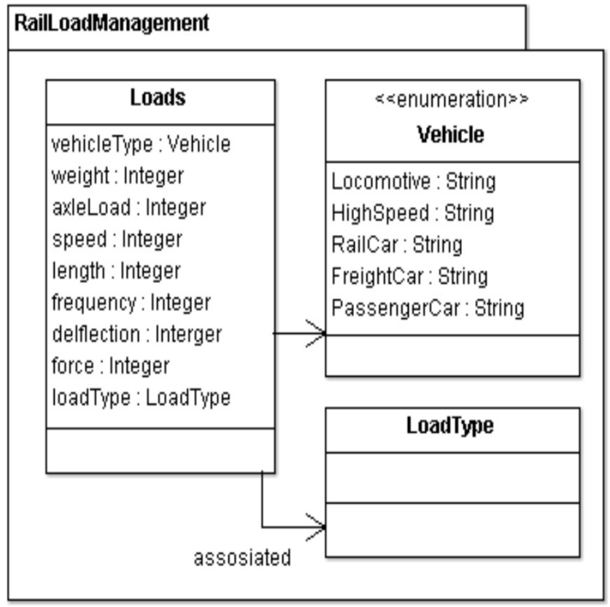

Figure 4: Information Model (Railway Load)

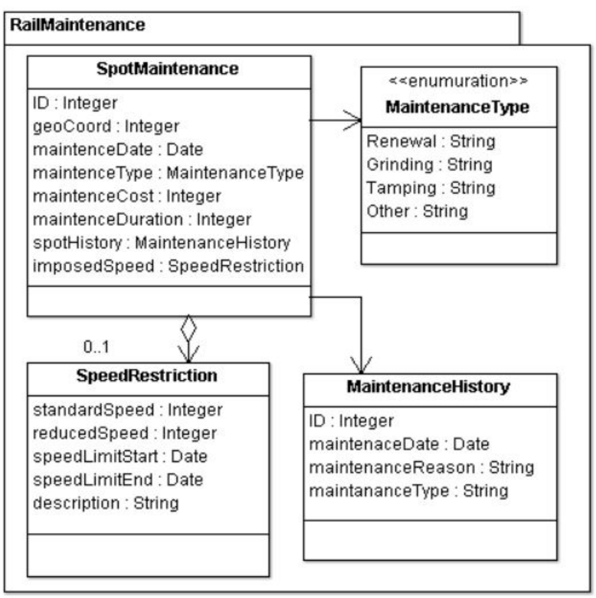

Figure 5: Information Model(Railway Maintenance)

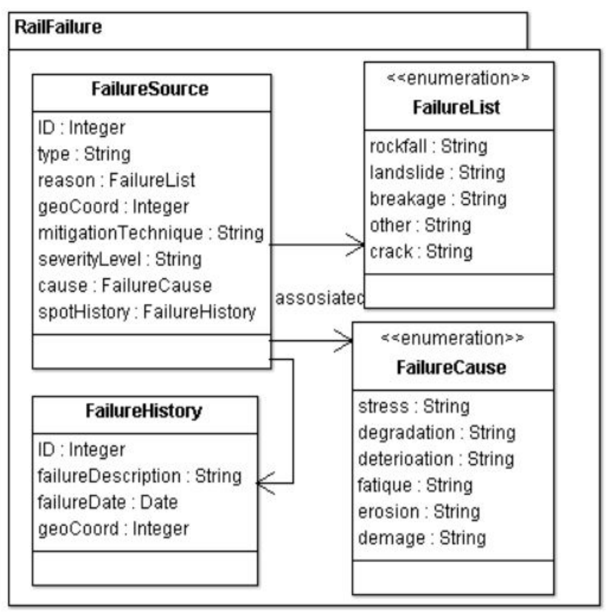

Figure 6: Information Model (Railway Failure)

4.4 Model of Railway Maintenance

Figure 5 shows the main object classes for rail maintenance management accounted for in the IM. As a result of rail maintenance work, the rail operation might need to be halted or speed restrictions need to be applied. Regarding the maintenance of particular section/spot, a number of important information can 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 to store data about all the maintenance activities that have been conducted at a particular section/spot in the past. As maintenance can affect the train operations, it is useful to keep the data about the imposed speed during the maintenance.

4.5 Model of Railway Failure

Figure 6 shows the main components and data properties related to railway failure. Failure source is the main object of railway failure, which has been failed. Information about the location of failure, failure reasons, failure severity, failure cause and failure spot history can be represented using the IM. Additionally, information regarding the type of failure and failure cause can be assigned to each railway failure element to for example determine how common certain failures are. Similarly, information regarding the history of failure will be useful to take the useful mitigation approaches to avoid failures in future.

5 Discussion

The presented information model not only provides a layout of railway infrastructure, but it provides the data support to monitoring, maintenance and operational domain of railway. Data structured by information model of railway assets represents the geometrical specification and inter-relations of core assets, which is further useful in the process of monitoring. An information model of rail maintenance specifies the planned schedule of maintenance as well as record the maintenance history along with maintenance reasons and maintenance type. Moreover, an information model of railway failure supports the identification of failure sources, types of different failures and possible reasons. The data properties of all the physical and conceptual entities specified by information model provide a strong base of data for stakeholders, such as infrastructure managers, to take informed decisions. The nature of decisions could vary from identification of failure sources, to risk assessment of an object in the network and management of load assessment of rolling stock. The presented information model will be used to provide a digitalized framework of data in a form of information management system to project partners. Moreover, we believe that the presented information model is comprehensive and generic enough to support the development of those railway applications where the data/record about the identified objects is required to be stored.

The key concept in the design of information model is to define the optimum level of granularity. High level of granularity provides in-depth and detailed knowledge of objects and their inter-relationship. However, it could result in an information model which is complex to understand and too difficult to modify. While, with the low level of granularity, the chances of leaving out the important objects is high. We mitigated this problem by taking two steps: first, we represent the rail infrastructure into different domains as shown in previous sections. Within each domain, we focused only on main components. Secondly, during our interview sessions, we asked the interviewees to suggest about the inclusion and exclusion of considered objects as well as any modification in the objects inter-relationship. The main focus of the developed information model is to communicate the system design instead of outlining the implementation details. Thus, we mainly focused at the level of granularity that kept our model easy to communicate and understandable.

As with any information model, many concerns can be raised based on the selection of objects, their considered properties, and their interrelationship among them. Thus, it is important for an information model to act as a dynamically updated model, during design and development phase, instead of a static model which reflect the system design. To support the need of continuous modification, the design and implementation techniques for the system development should be flexible enough. It can be noted in the information models, provided above (for instance see Figure 1), that few class attributes e.g. geo-Cord, condition is repeatedly defined for each class, while the concept of inheritance could have been employed for this purpose. Many standardized data models (e.g. CityGML, IFC) use the inheritance to achieve the reusability. We avoid the use of inheritance intentionally as even though inheritance improve the reusability of code but it induces the unnecessary coupling in the system design and implementation (Burn, 2014). With the compelling need to support the dynamic system design and implementation, a prominent shift has been noticed, in database management techniques, from the static database schema definition (relational databases) to dynamic schema definition (NoSQL databases) (Moniruzzaman and Hossain, 2013).

6 Conclusion

The information model of railway infrastructure developed in the paper is the part of an on-going effort of DESTination RAIL project, which require the development of network wide information management system to mitigate the challenges of data isolation. We followed an iterative development approach to design information model, where the data item from the project proposal is identified, standard data sharing format (RailML) of railway along with seminal railway engineering literature is considered, finally semi-structured interviews with twelve participants from nine different institutes were conducted. Instead of focusing only on railway assets and railway operations, the information model presented here can store data about the all the important aspects of railway maintenance, railway assets, railway failures, railway loads and railway risk assessment. Finally, the resulted information model provides a strong base of data that can be used by infrastructure managers to take well grounded operations and/or maintenance decisions.

The developed information model has served two main purposes: first, it structures the railway infrastructure based on its data properties, which will be further used for the development of information management system. Secondly, it is used as a communication tool to understand the end-users' perspective with respect to data needs and domain knowledge. We believe, the provided information model is generic enough to be used by other railway application where the data about considered objects is required to be stored.

References

Bernardi, S., Flammini, F., Marrone, S., Merseguer, J., Papa, C., and Vittorini, V. (2011). Model-driven availability evaluation of railway control systems. In *Computer Safety, Reliability, and Security*, pages 15–28. Springer.

Borrmann, A., Flurl, M., Jubierre, J. R., Mundani, R.-P., and Rank, E. (2014). Synchronous collaborative tunnel design based on consistency-preserving multi-scale models. *Advanced Engineering Informatics*, 28(4):499 – 517.

Bosschaart, M., Quaglietta, E., Janssen, B., and Goverde, R. M. (2015). Efficient formalization of railway interlocking data in railml. *Information Systems*, 49:126 – 141.

Burn, S. (2014). Inheritance is Bad: Code Reuse. Available at: http://blogs.perl.org/users/sid$_{b}urn/2014/03/inheritance - is - bad - code - reuse - part - 1.html$.

Esveld, C. (2001). Modern railway track. MRT-productions Zaltbommel, Germany.

Galar, D., Gustafson, A., Tormos, B., and Berges, L. (2012). Maintenance Decision Making Based on Different Types of Data Fusion. *Eksploatacja I Niezawodnosc-Maintenance and Reliability*, 14(2):135–144.

Jeff Leighton, BIM Manager, V. (2015). Evolution of Building Information Modeling. Available at: http://www.sbcmag.info/news/2015/nov/contractor-bim-process-not-software.

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

Kolmorgen, V. and Huerlimann, D. (2005). Railml-a standard interface for railway data. *EUROPEAN RAILWAY REVIEW*, (4).

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

Lucas, J., Bulbul, T., and Thabet, W. (2013). An object-oriented model to support healthcare facility information management. *Automation in Construction*, 31:281–291.

Moniruzzaman, A. and Hossain, S. A. (2013). Nosql database: New era of databases for big data analytics-classification, characteristics and comparison. *arXiv preprint arXiv:1307.0191*.

Thaduri, A., Galar, D., and Kumar, U. (2015). Railway assets: A potential domain for big data analytics. *Procedia Computer Science*, 53(1):457–467.

Turk, Z. (2001). Phenomenologial foundations of conceptual product modelling in architecture, engineering and construction. *Artificial Intelligence in Engineering*, 15(2):83 – 92.

Turkaslan-Bulbul, M. T. and Akin, O. (2006). Computational support for building evaluation: Embedded commissioning model. *Automation in construction*, 15(4):438–447.

Vossebeld, N. and Hartmann, T. (2014). Supporting tunnel safety assessment with an information model. In *2014 International Conference on Computing in Civil and Building Engineering*.

Wang, H., Xu, W., and Jia, C. (2012). Metadata-oriented data model supporting railway distributed system integration. *Journal of Software*, 7(4):814–822.

Young, R. R. (2002). Recommended requirements gathering practices. *CrossTalk*, 15(4):9–12.

Zamanian, M. K. and Pittman, J. H. (1999). A software industry perspective on aec information models for distributed collaboration. *Automation in Construction*, 8(3):237–248.