Capacity for Rail

Towards an affordable, resilient, innovative and high-capacity European Railway System for 2030/2050

Data notation and modeling

Date: 10 March 2016

Public deliverable D 3.4.1

This project has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement n° 605650

Collaborative project SCP3-GA-2013-60560
Increased Capacity 4 Rail networks through enhanced infrastructure and optimised operations
FP7-SST-2013-RTD-1

Lead contractor for this deliverable:

- Technische Universität Dresden

Contributors

- University of Birmingham, UoB
- Network Rail, NR
- SYSTRA
- ADIF
- Deutsche Bahn, DB,
- Trafikverket, TRV,
- University of Newcastle, UNEW
- TRL Limited, TRL

Project coordinator

- International Union of Railways, UIC

www.Capacity4rail.eu

EXECUTIVE SUMMARY

Deliverable 3.4.1 of the CAPACITY4RAIL project focuses on data formats and models for data exchange used in the railway sector with considerations of approaches in other transport modes. The focus is on open data formats that have the potential to substitute proprietary data formats in the future. It analysis three usage scenarios, where data exchange is and will be important to guarantee effective usage of railway capacity:

- Consistent cross industry infrastructure data;
- Effective usage of multimodal transport systems;
- Real-time operations across organisational and member state borders.

For each use case, visions for 2020, 2030 and 2050 are outlined, the feasibility of relevant data formats, models and concepts are presented, and current gaps are demonstrated.

The deliverable concludes by making some recommendations on priority areas for data modelling work in the CAPACITY4RAIL project:

- Interaction of IM asset data sets with OpenStreetMap data in a round-trip process;
- Upgrade of ON-TIME RTTP regarding railML 3 / UIC-RailTopoModel, and proposing it to the railML community;
- Incorporate the consolidated findings of SP4 on sensor data into the upgraded RTTP;
- Comparison of Schematron- and Ontology-based approaches for railway data verification;
- Development of ontologies supporting Linked open data from specific formats such as railML and NeTEx;
- Demonstrating the developed ontologies in typical use cases, oriented at the stories of this document.

The question, how the data sets in the proposed data formats and models shall interact in order to enable scenario-oriented software solutions, will be answered in the deliverable D3.4.2 in form of architecture recommendations.

TABLE OF CONTENT

INDEX OF FIGURES

INDEX OF TABLES

ABBREVIATIONS AND ACRONYMS

Abbreviation / Acronym	Description
ADL	Adaptive Zuglenkung (Adaptive train control)
API	Application Programming Interface
ATM	Automatic Ticketing Machine
AWT	All Ways Travelling (EU-founded project)
CEN	European Committee for Standardization
CER	Community of European Railway and Infrastructure Companies
CSV	Comma Separated Value (text file format)
EDIFACT	Electronic Data Interchange For Administration, Commerce and Transport
EIM	European Rail Infrastructure Managers
ERA	European Railway Agency
ERIM	European Railway Infrastructure Masterplan
FSM	Full Service Model Initiative
GML	Geographic Markup Language
GTFS	General Transit Feed Specification
HABD	Hot Axle Box Detection
HMI	Human-Machine Interface (train driver display)
ICT	Information and Communications Technology
IDM^{VU}	Infrastruktur-Daten-Management für Verkehrsunternehmen
IEEE	Institute of Electrical and Electronics

Abbreviation / Acronym	Description
	Engineers
IFOPT	Identification of Fixed Objects in Public Transport
IM	Infrastructure Manager
IU	Interoperability Unit (of ERA)
LOD	Linked Open Data
MS	Member State
NaPTAN	National Public Transport Access Node
NASA	National Aeronautics and Space Administration
NEPTUNE	Norme d'Échange Profil Transport collectif utilisant la Normalisation Européenne
NeTEx	Network and Timetable Exchange
NRE	National Railway Entity
PIS	Passenger Information System
PMM	Perturbation Management Module
PRM	Passenger with Reduced Mobility
PT	Public Transport
PTO	Public Transport Operator
OSM	OpenStreetMap
OWL	Web Ontology Language
RaCoOn	Railway Core Ontology
railML	Railway Markup Language
RDF	Resource Description Framework
RDO	Railway Domain Ontology
RCT2	Rail Combined Ticket 2
RINF	Register of Infrastructure
ROC	Railway Operations Centre
RTPI	Real-Time Passenger Information

Abbreviation / Acronym	Description
	System
RTTP	Real-Time Traffic Plan
SKOS	Simple Knowledge Organization System
SOA	Service Oriented Architecture
SPARQL	Protocol and RDF Query Language
SWRL	Semantic Web Rule Language
TEN-T	Trans European Transport Network
TAF	Telematics Application for Freight
TAP	Telematics Application for Passengers
TCC	Train Control Centre
TMS	Traffic Management System
TSI	Technical Specification for Interoperability
UIC	International Union Of Railways
VDV	Verband Deutscher Verkehrsunternehmen
WP	Work Package
XML	Extensible Markup Language
XSD	XML Schema Definition

1. BACKGROUND

The CAPACITY4RAIL project is a EU FP7 funded industry lead initiative which aims to answer the question *"How to obtain an affordable, adaptable, automated, resilient and high-capacity railway;*
for 2020, 2030 and 2050?"

CAPACITY4RAIL will provide an overall increase in railway capacity by developing a holistic view on the railway as a system of interacting technical components driven by customer demand. It is structured into sub-projects (SP) with interacting work packages (WP) as presented in Figure 1-1.

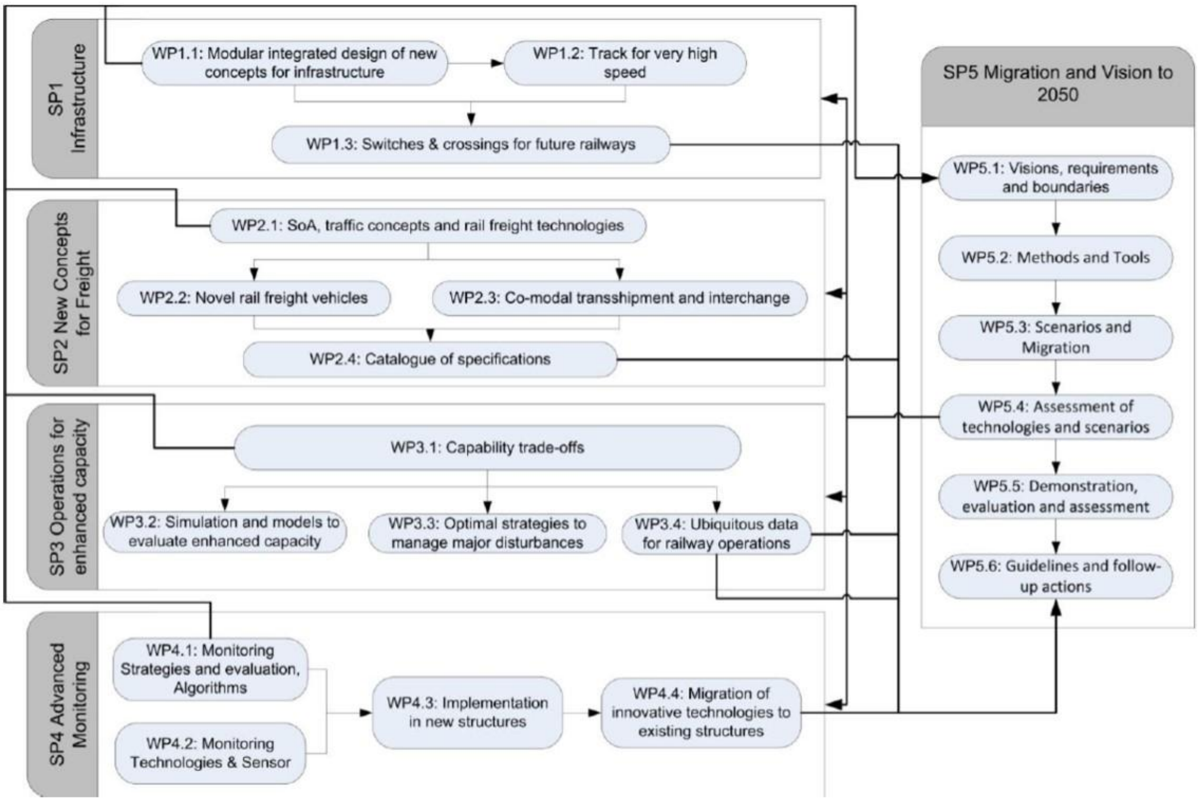

Figure 1-1 CAPACITY4RAIL Structure breakdown and interactions

SP3 "Operations for enhanced capacity" delivers approaches that help planners to understand and prioritise system capabilities and decide on optimal strategies to:

- Increase overall system capability;
- Respond dynamically to planned and unplanned changes and
- Support real-time punctuality management.

Guidance documents will include strategies for incident and emergency management including recommendations for the management of extreme weather situations. A roadmap for automation of traffic management systems following the concept in Figure 1-2 will enable the rail industry to meet the challenges of the future, such as high-speed rail freight and greater levels of transhipment between rail and other modes. (CAPACITY4RAIL, 2013)

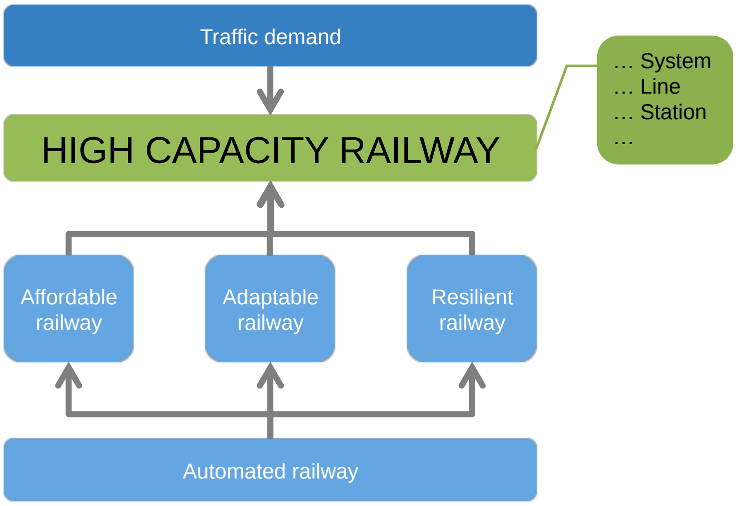

Figure 1-2 Roadmap for automation of traffic management systems

WP3.4 “Ubiquitous data for railway operation” will enable the railways to harness and effectively use large and diverse sources of data to extract meaningful information and knowledge to support operational strategies.

As the first deliverable from WP3.4, D3.4.1 will focus on current state of the art in data exchange on European railways, and also considering touch points with other transport modes — a key consideration when optimising the use of available capacity. One area of particular interest is the community led efforts developing open data and data models. These projects, which commonly have significant amounts of volunteer time invested during their development have the advantage of a broad range of perspectives on the domain being modelled (due to wide usage) and will be reported here with a view towards potential industry adoption, particular for none safety critical tasks.

Data architectures proposed by the project will be presented in the second deliverable (D3.4.2) and open data will serve as a cross cutting theme (Figure 1-3).

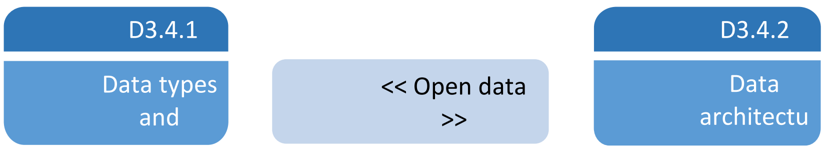

Figure 1-3 WP3.4 Deliverables D3.4.1 and D3.4.2

Data exchange / management practices between stakeholders within the rail domain still lag behind those in other large-scale infrastructure industries, such as the oil and gas industry. Despite this, in recent years progress has been made around modelling key business concepts such as the infrastructure and service provision to customers (train running information etc.) In order to drive the development of automated data exchange within the sector, the C4R project must look beyond the next "logical" commercial step that will be handled by the industry itself, and instead focus on the longer-term goals of wider modal interactions and the provision of linked open data on network state.

Three important usage scenarios for ubiquitous data in railway operations are analysed regarding state of the art data types and models with a perspective to incorporate open data:

- Consistent Cross Industry Infrastructure Data;
- Multimodal Transport Systems;
- Real-time operational data.

Each usage scenario is illustrated in chapter 2 with the help of a storyboard drawing visions for 2020, 2030 and 2050. Certain data exchange formats and models, that may support the visions, are described in chapter 3 ensuring the reader has the necessary background to support the analysis that follows. The analysis section begins with a mapping to the storyboards for keeping the track throughout the document. In chapters 4 to 6 the alignment between the storyboards and the data exchange models and formats is presented, showing where existing work can support the future needs The conclusions in chapter **Erreur ! Source du renvoi introuvable.** give a short overview for data types and models, the following deliverable D3.4.2 will be based on that outcomes for defining the needed architecture. The structure of the document is summarised in Figure 1-4.

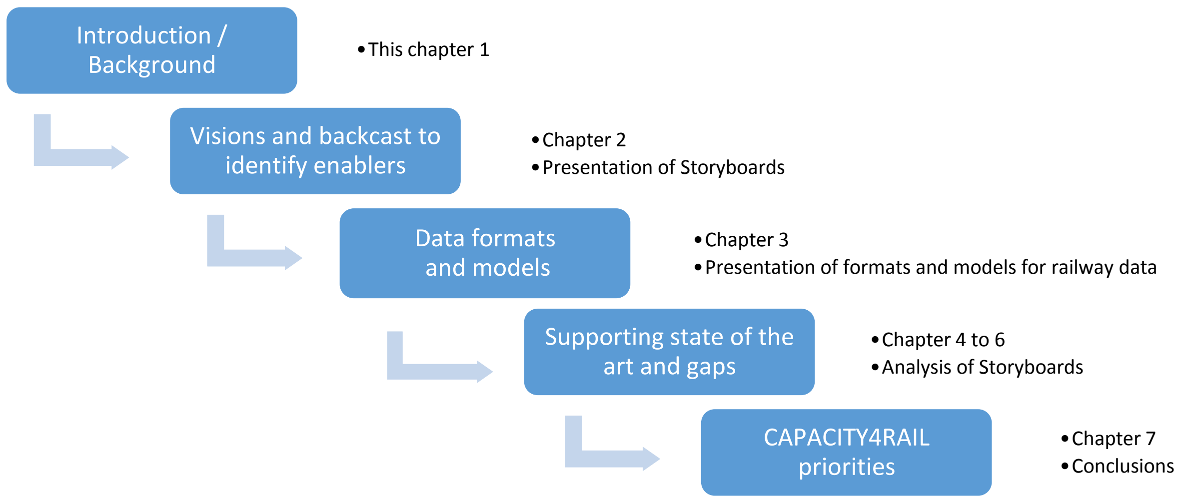

Figure 1-4 Structure of Deliverable 3.4.1

2. STORYBOARDS FOR UBIQUITOUS DATA IN SUPPORT OF RAILWAY OPERATIONS

In this chapter three usage scenarios in the railway operations domain are introduced, that significantly rely on ubiquitous data. Visions for the usage of data within each scenario are presented at the 2020, 2030, and 2050 horizons. Backcasting has been used to ensure the scenarios are focused on delivering the long-term (2050) vision, and are not restricted by the expected development path of current systems and technology.

In doing so, the deliverable follows the CAPACITY4RAIL slogan *"towards an affordable, resilient, innovative and high-capacity European Railway System for 2030/2050"*.

2.1 CONSISTENT CROSS INDUSTRY INFRASTRUCTURE DATA IN SUPPORT OF PLANNING, SIMULATION AND OPERATIONS

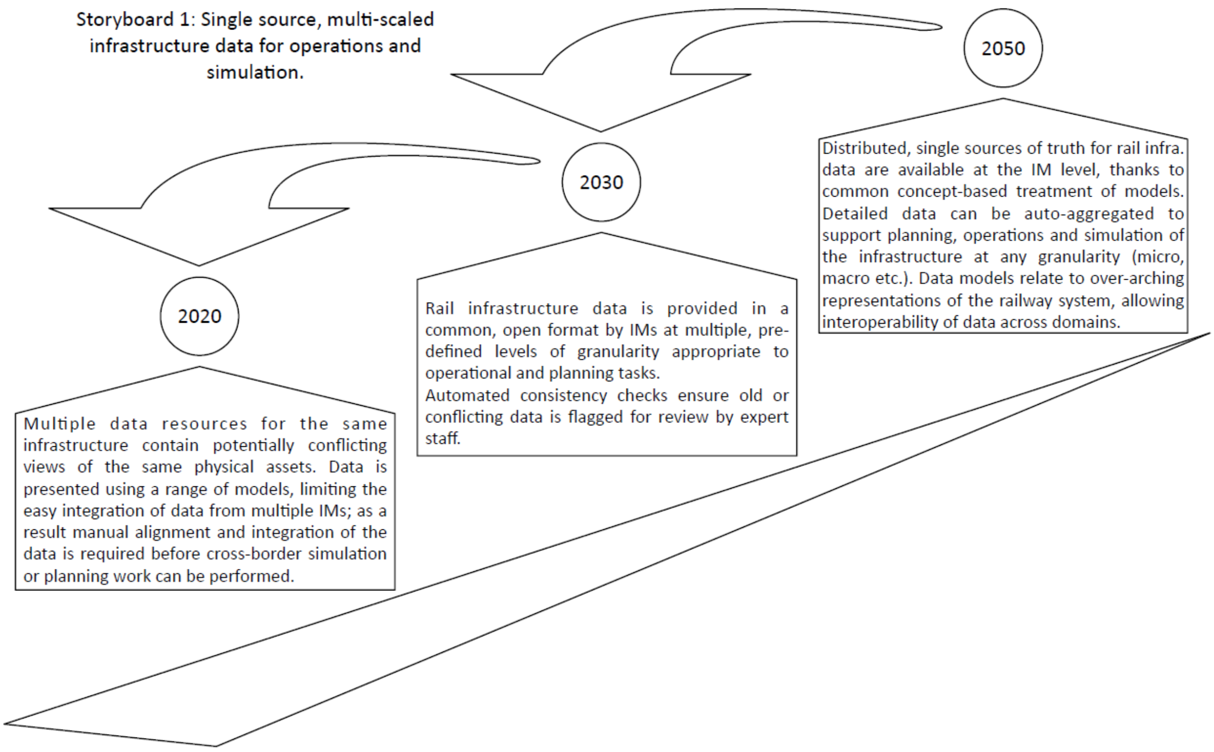

Figure 2-1 Storyboard 1 – Infrastructure data for operations and simulation

Figure 2-1 is a backcasting diagram showing how improved cross industry information exchange will support the increase of capacity on the existing infrastructure. This will be achieved through the delivery of more timely and accurate information to tools for planning, simulation and operations, both in national and international services. A uniform data exchange format, that is able to reflect the user needs in a future-proof way, builds a solid foundation for reducing implementation costs and enabling a broad acceptance.

Currently, different data formats for railway infrastructure data are used to exchange information between different applications, different companies or even different divisions of the same company. Besides infrastructure manager (IM) specific data formats, there exist some promising initiatives to unify data exchange formats across Europe within the railway sector and in the whole transport sector as well. These initiatives are driven by both national interest groups, like the German format IDM[VU] and by legislation mandated by the European Union, like INSPIRE and RINF. Open source initiatives based on a free cooperation of professional and unpaid developers, like railML and OpenStreetMap, are also important players in this area.

Independent from the current application perspective, infrastructure data sets have to deal with different topological granularities, which are defined in Table 2-1 (ERIM Workgroup, 2014), in order to comply with several requirements from planning, simulation and operations.

Table 2-1 Topological Granularities – general abstraction levels

Level	Scale	Description
Corridor	Very small	Primary routes within a network, e.g. rail freight corridor
Macroscopic	Small	A generalised view of the mesoscopic level, e.g. multiple tracks within a line appear as a single line
Mesoscopic	Intermediate	A generalised view of the microscopic level
Microscopic	Large	Track level information at the highest level of details

Figure 2-2 illustrates these levels of granularity as used in this document:

- National as well as cross-border infrastructure data sets shall seamlessly integrate at each of these levels. IM borders are treated the same way as state borders.
- Corridors mostly correspond to TEN-T corridors for certain services, such as freight and passenger or conventional and high-speed transport, including their important stations for fulfilling the service.
- Macroscopic level equals to the typical national network of lines, where a line comprises of one single track or two parallel tracks for each direction. Operational points are considered from junctions to large stations including smaller stations or just stop points. Connections

between lines may be deduced from linking the same operational point ignoring the ability to traverse.

- Mesoscopic level consists of the same operational points as for the macroscopic level. Whereas tracks between operational points are defined instead of lines. Connections between tracks are established through operational points, traversing feasibility shall be provided.
- Microscopic level contains tracks, switches and crossings and more detailed trackside facilities, e.g. signals, platforms.
- Nanoscopic level, which would focus on both rails, is not considered in this document.

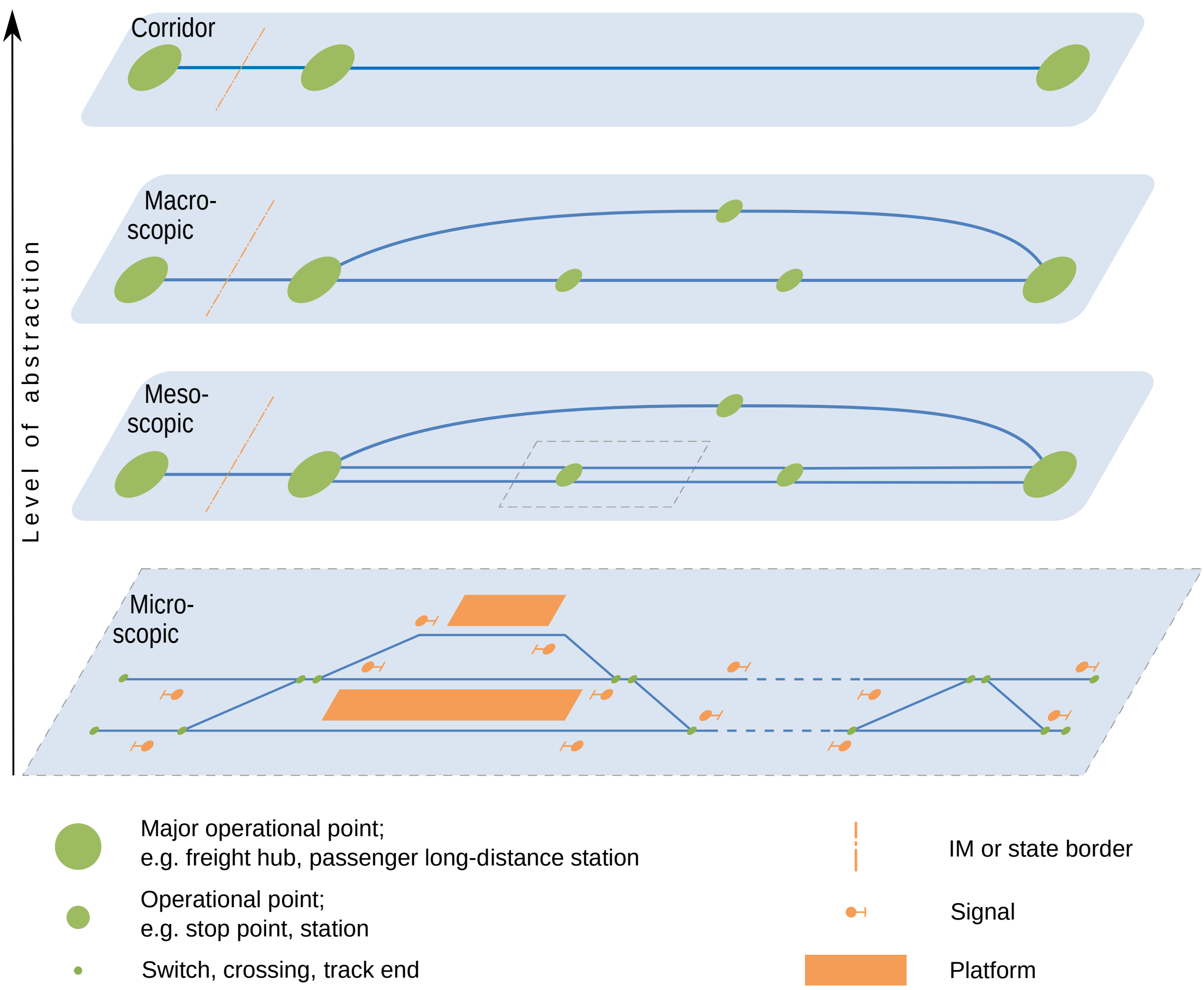

Figure 2-2 Topological granularities of infrastructure

Depending on the nature of the task being performed, input data at any one of these levels of abstraction may be required, as shown in Figure 2-3. Generally speaking, corridor and macroscopic level data provide an adequate base for the planning of long-term and mid-term railway traffic, however mesoscopic data must be considered for more detailed operational planning tasks. Meso- and microscopic levels are appropriate for fine-grained simulation and real-time operation.

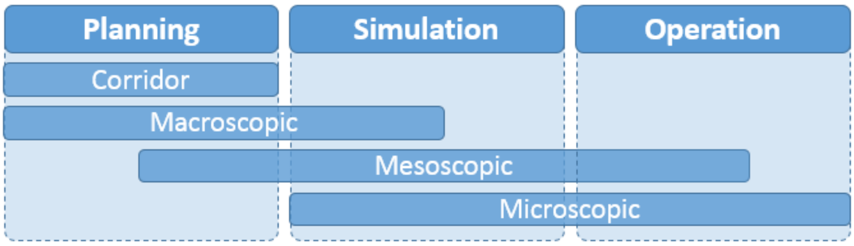

Figure 2-3 Topological granularities for planning, simulation and operation

In general, detailed simulations based on microscopic data sets are very time-consuming and therefore not suitable for real-time traffic controller assistance. Approximated simulations based on macroscopic data are fast enough to be used operationally, but lack conflict detection at the level of track vacancy. A compromise solution using elements of both approaches would lead to fast simulations that work well in all but the most complex capacity scenarios at junctions or busy stations.

A lack of infrastructure data at the appropriate level of granularity to support a given scale of simulation is a relatively common problem in this domain, so a further use case in this area can be found in the provision of abstracted / inferred data based on available information at a different scale to provide approximate simulation results where needed.

Additional benefits arise, if the already available level-specific data sets are joined and compared with data sets at other granularities. Storyboard 3 "Real-time operational data across organisational and member state boundaries" is partly based on this approach. The comparison of data sets from different sources may also lead to the detection of inconsistencies, enabling better data qualities.

The analysis in chapter 3 focuses on data types and models that assure interoperability of data regardless their origin and the capability to aggregate between any granularity. Therefore the following data formats and models are introduced: railML in section 3.2, UIC RailTopoModel in section 3.3, RINF in section 0, INSPIRE in section 3.5, IDM^{VU} in section 3.6 and OpenStreetMap in section 3.15.

2.2 Effective usage of multimodal transport system capacity

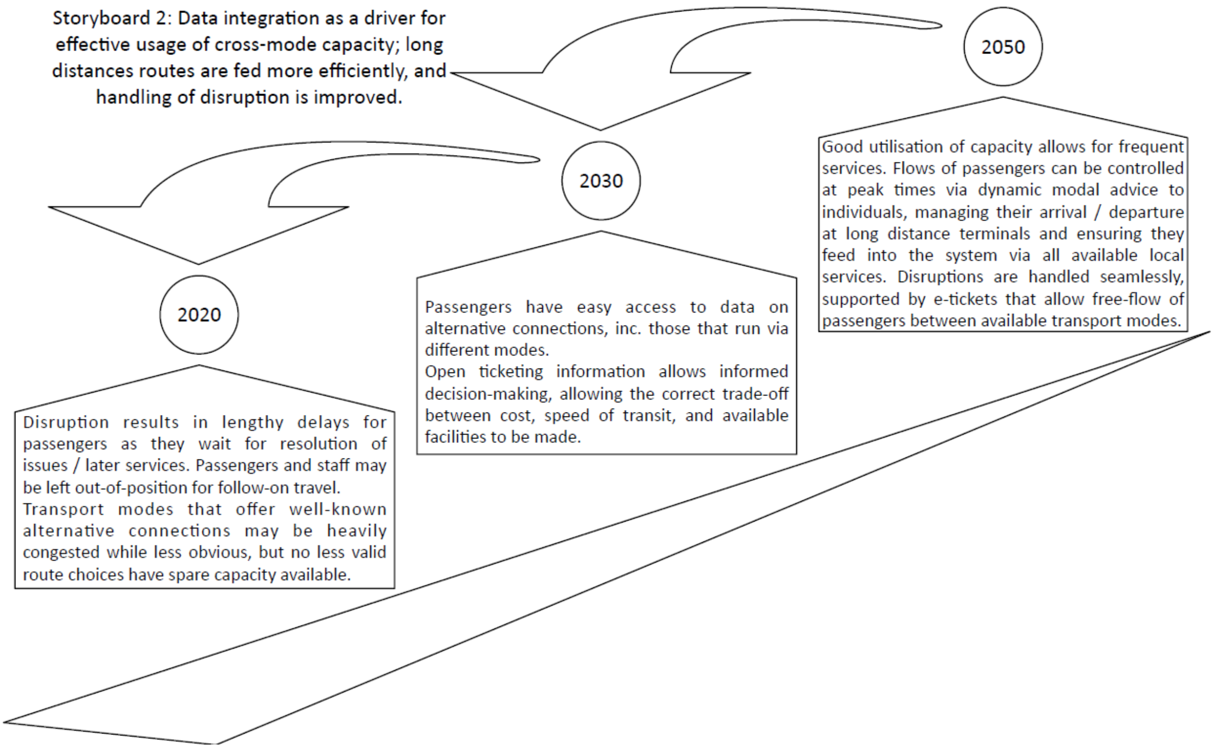

Figure 2-4 Storyboard 2 – Effective usage of cross-mode capacity

The second storyboard focuses on the use of data integration as a driver for more effective use of existing network capacity both within rail and in the wider multimodal transportation system, particularly during disrupted operations (Figure 2-4).

Multimodal transport is characterised by the use of more than one mode within the scope of a single end-to-end journey. By making the best use at available multimodal capacity, C4R will free up rail capacity and encourage modal shift from modes such as shorthand air travel. The set of transport systems considered are shown in Figure 2-5.

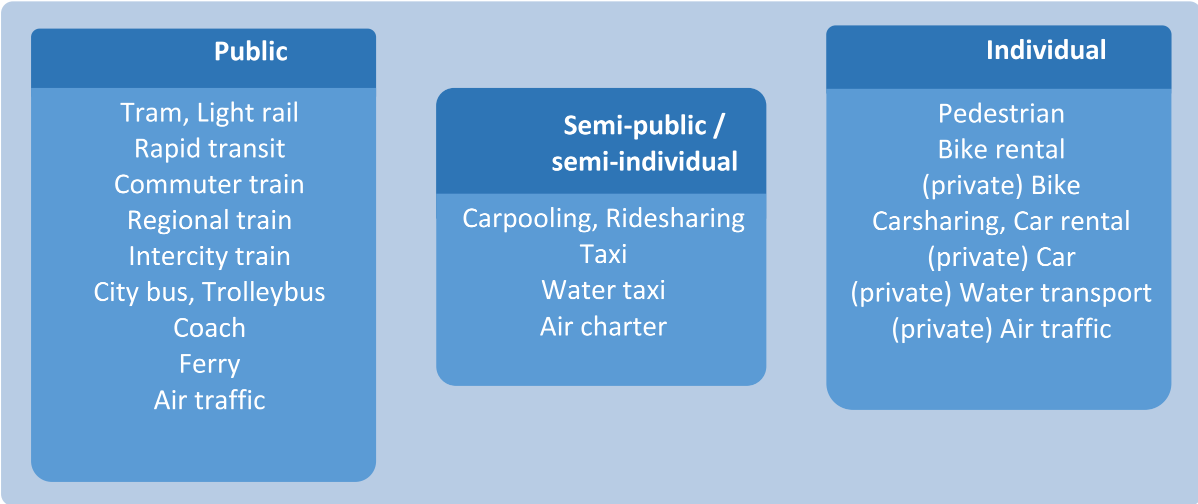

Figure 2-5 Passenger transport modes

The highest impact from increasing railway capacity within the scope of the second storyboard is expected to focus on public transport modes, keeping in mind that individual and semi-individual transport modes play an important role as feeder systems. Air traffic is only considered as a feeder in case of blackout or irregularities.

Beyond the direct scope of C4R, the smooth interaction of data across the outlined modes is a key outcome for the sector as described in the Whitepaper of the European Commission "Roadmap to a Single European Transport Area – Towards a competitive and resource efficient transport system", which specifies ten goals. One amongst them is "Multimodal information services" (European Commission, 2011):

> *[…]*
> *(5) A fully functional and EU-wide multimodal TEN-T 'core network' by 2030, with a high quality and capacity network by 2050 and a corresponding set of information services.*
> *[…]*

A similar objective has been investigated by the *"All Ways Travelling (AWT)"* consortium active from 04/2013 to 01/2016. Appointed by the European Commission, AWT will develop and validate a model for a multimodal pan-European passenger transport information and booking system (AWT Consortium, 2013-2015). In its first phase, "In-depth study of multimodality", AWT prepared a list of parameters that influence the passengers' decision for use of different transport modes (AWT Consortium, 2014):

- Timetable information – accurate and on short call;
- Station information – including transfer and navigation path information;
- Fare information – individual and cross-mode.

It is hoped that the reliable provision of this information subset will encourage modal shift from individual to public transport modes as driven by the availability of flexible information services. These three key functions are taken as guidelines for the analysis of data types and models throughout chapter 5, namely NeTEx introduced in section 3.7, TAP TSI in section 3.12 and GTFS in section 3.13.

2.3 Real-time operational data across organisational and member state boundaries

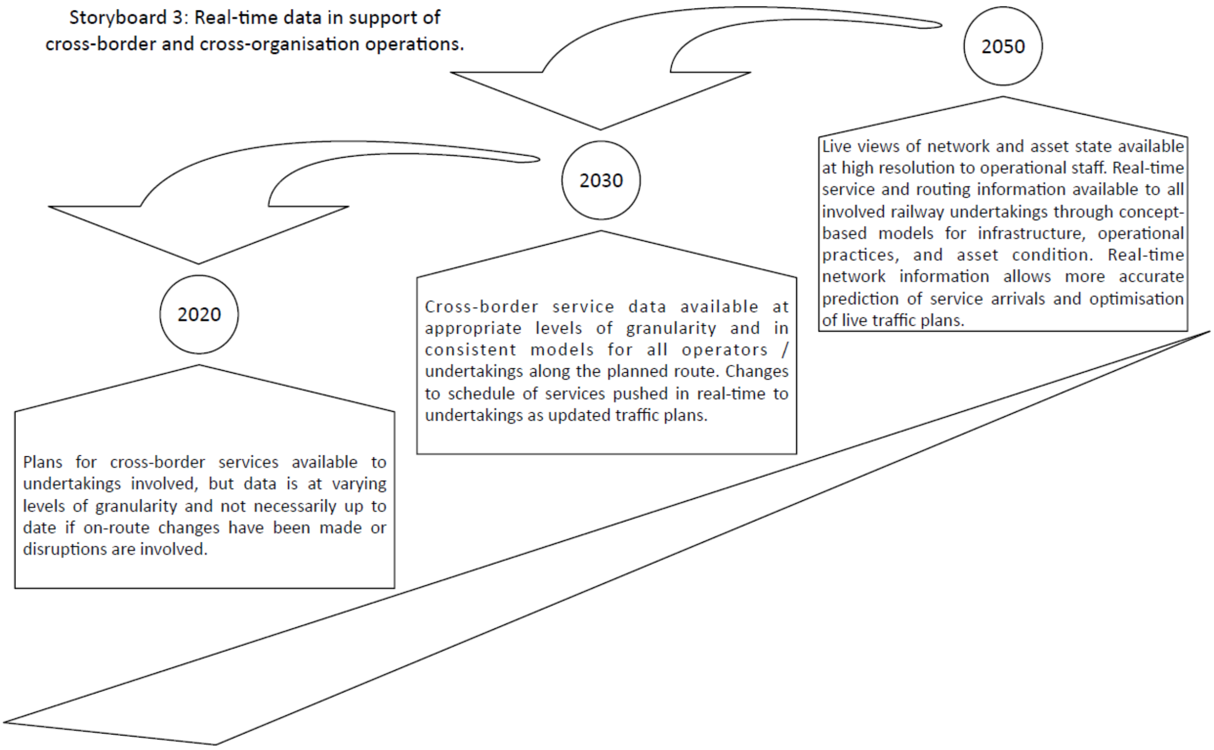

Figure 2-6 Storyboard 3 – Real-time data in support of cross-border / cross-organisation operations

The third storyboard deals with the handover of planned rail services between organisations at operational or state borders and the delivery of timely operational data throughout a journey (Figure 2-6). The availability of accurate, real-time data opens the door for operational optimization and customer information.

In addition to key data from the railway, such as booked train paths and working timetables, data from a number of external stakeholders may also influence the capacity trade-off in the wider multimodal system. In particular a clear understanding of live delays across the system as a whole and the way in which those delays propagate, are of critical importance to the effective use at capacity available.

From the perspective of data modelling, several factors have a heavy influence on this issue. On the one hand, clarifying the level of granularity and standardizing the content and transfer of planned and real-time data will enable the widest possible usage in new services.

Where possible, the provided data sets shall be enriched with all available data in a standardized structure.

By providing consistent data that is more easily integrated with information from other transport modes, the rail industry will be able to maximise on its IT investments and deliver capacity improvements beyond the scope of its own assets (i.e. it will drive the use of local modes as feeders for the national rail network). Finally, real-time and planned operational data shall be available for any interested party in the railway domain.

Figure 2-7 illustrates the evolution of data model and recipients exemplarily on traffic disruption on a railway line:

- 2020: IM informs subsequent RU about the incident within his responsibility zone. Data is exchanged in the IM-RU-specific way. Thus, data format and granularity of data is not standardized;
- 2030: Subsequent RUs along the planned route get real-time operational data, even before they enter the responsibility zone of the IM in charge. All involved partners get the same information in the same format and the appropriate level of granularity including quality indication;
- 2050: IM provides information for any interested party, i.e. RUs along the planned route as well as along routes that join the disturbed route behind the disrupted area. Thus, newly available slots can be used to optimise the network capacity. Comprehensive data are provided in a detailed level of granularity.

At the local level, cross border exchange of information may take place between different stakeholders with the same function, or at operational boundaries such as the interface point between two routes. On larger scales, the same software interfaces and governance processes can be used to exchange data between different infrastructure managers (IMs) across member state borders.

Prospective optimization for a cross-border region through re-scheduling of services is based on current and short-term predicted network capacity states, which rely on sound real-time data. Integrating actual, precise, pre-processed sensor data enables more robust estimations and predictions. Research on sensor data and monitoring of railway infrastructure elements is done in the frame of SP4 of CAPACITY4RAIL, which is complementary to the findings of the current WP.

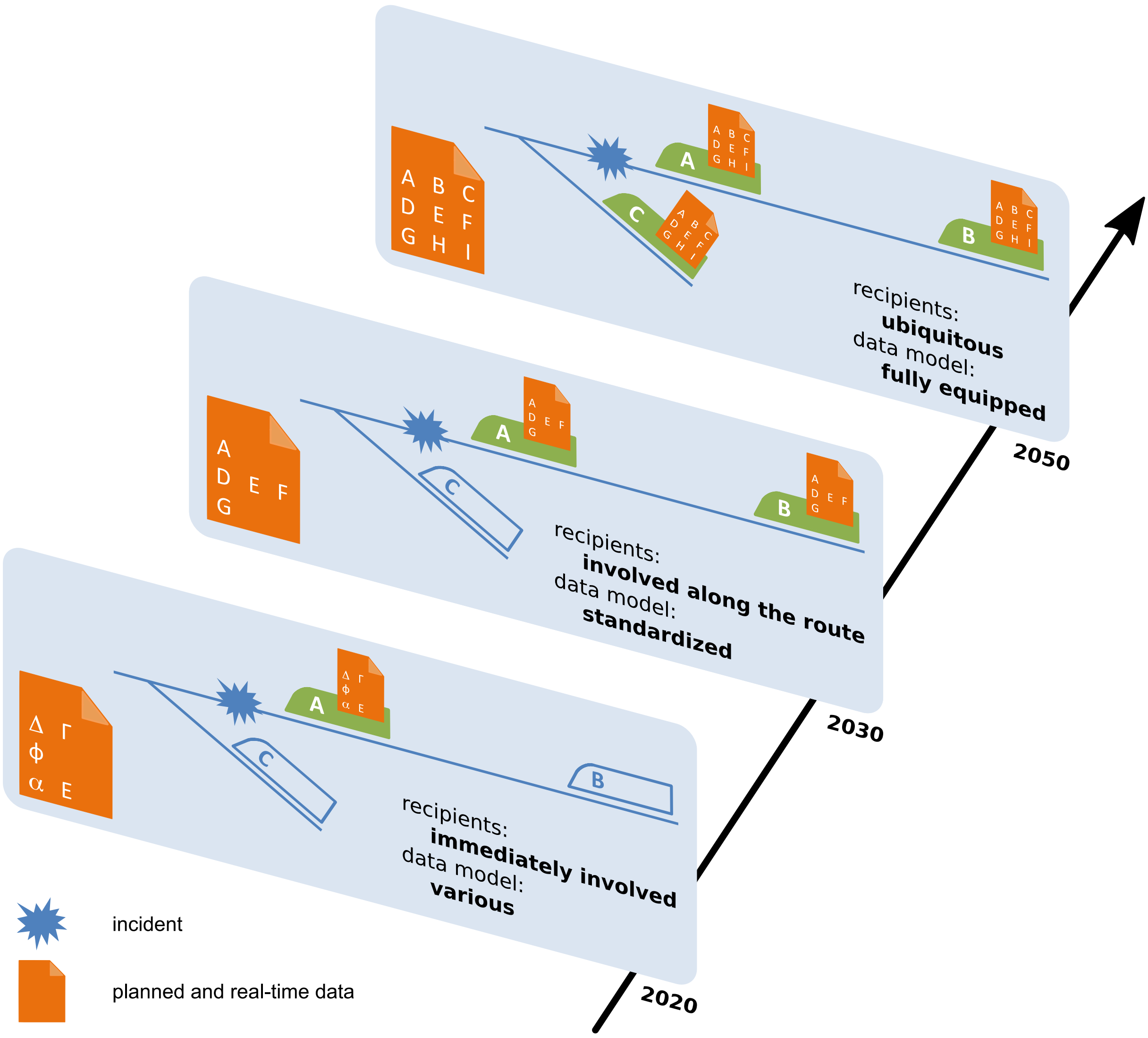

Figure 2-7 Evolution of available data for involved undertakings

Standardizing and enriching real-time operational data, alongside demand from a broader audience of such upgraded information is analysed in detail in chapter 6, which bases on the prior to that introduced data formats and models: railML in section 3.2, UIC RailTopoModel in section 3.3, SIRI in section 3.8, TAF TSI in section 0, ON-TIME RTTP in section 3.10 and GTFS-realtime in section 0.

3. DATA FORMATS, MODELS AND CONCEPTS

The storyboards described in chapter 2 can only realized their full benefit, if the available data models support the exchanges required between different tools in the same or different companies in a complete, unambiguous and verifiable way. The following criteria are taken into consideration during the analysis of certain data formats and models in the next sections:

- Reducing the number of data formats for import and export result in cheaper interfaces through less development effort.
- Standardized data formats for import and export reduce barriers to data exchange and enable easy adoption of software in new markets.
- Well-documented data formats and models minimize problems, which occur by reason of misinterpretation.
- A widespread user community of certain data formats and models results in validation-through-use in different regions and use cases.
- Open Source data formats and models ensure an easy way to receive the specification. Well-known Open Source licenses for them minimize potential property right infringements.
- Traditional international, European and national standardization bodies are similar to classic Open Source communities, where the de jure standards evolve more slowly than usual because of their meeting and voting regulations.

Effective data exchange between the railway domain and other transport modes is encouraged by European and national legislations in order to enhance sustainable public transport services in general. The EU pushes Open Data and exchange formats in the transport sector as well as in other application domains.

The French *Secretary of State for Transport, Sea and Fisheries* commissioned a report concerning “Open Data for Transport in France”, illustrating their intent to comply with the EU requirements. It focuses on technical, economic and legal aspects of open data provision, considering French legislation and standards as well as European. (Jutand, 2015)

Provision of a data classification for public transport services is a central thread of the French report. In Figure 3-1 this data classification is shown as basic structure and overlaid by the storyboard data demands, and shares many commonalities with this document. The storyboards share many data classes. In particular, the topological network plays a decisive role for a successful application of all storyboards within the drafted horizons. Real-time personal usage data is out of scope for this deliverable, not considering possible transport service adjustments from post-processed data sets.

The following sections introduce data formats and models that have potential to deliver the data needs of the storyboards. Readers, who are familiar with them, may skip this chapter and proceed with the analysis of the storyboards from chapter 4 on.

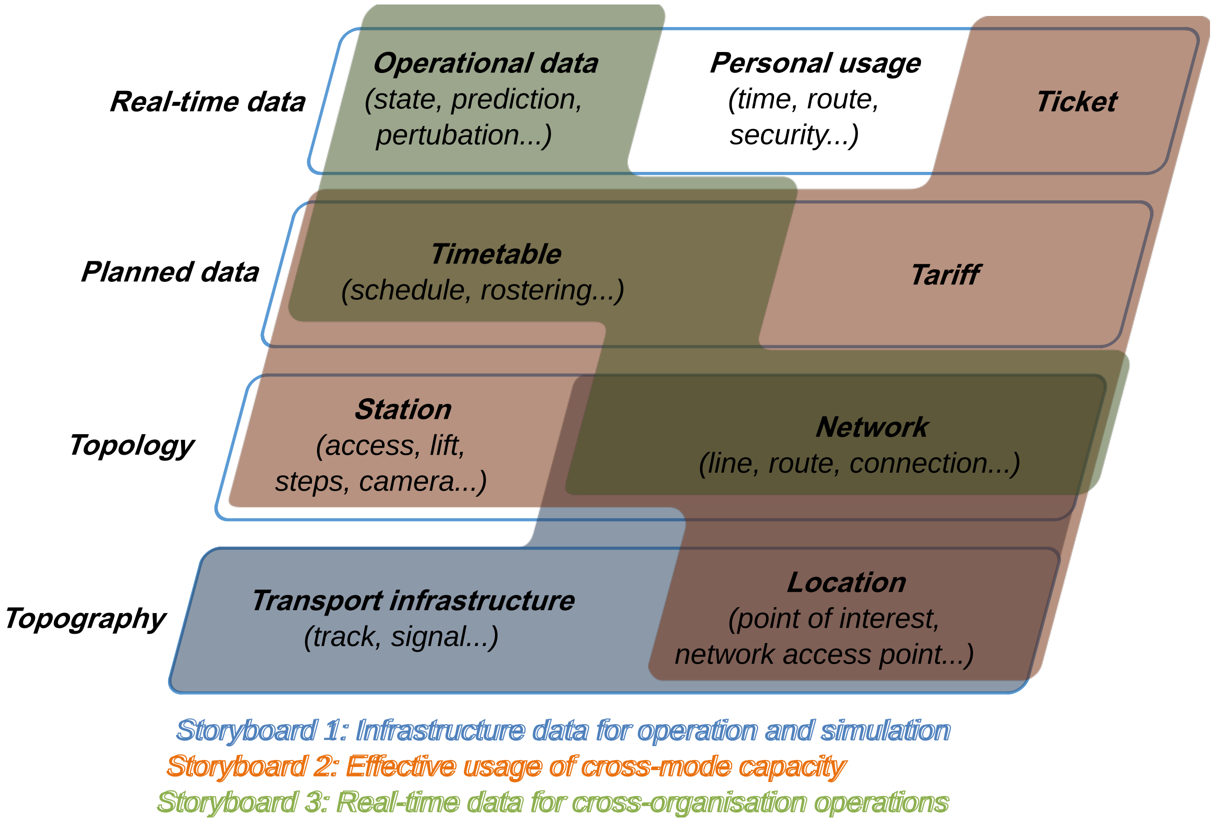

Figure 3-1 Public transport data classification, adapted from (Jutand, 2015)

3.1 Schema based XML formats and domain specific models

Data sets may be exchanged between software tools in several file formats. The eXtensible Markup Language (XML) specified by the World Wide Web Consortium (W3C) defines a format that enables the creation of domain specific sublanguages in a formal way and is therefore widely used, also in terms of a large variety of dialects.

The grammar of an XML file may be constrained with the help of XML Schema Document (XSD), which defines allowed elements and attributes alongside their types, order and multiplicity. The XML tree is built upon parent-child relationships of the elements. Validation of XML files is based on such grammar definitions, if available. Otherwise they are checked for their well-formedness.

XML parsers are widely available for modern programming languages and therefore ease the software development process for importing and exporting data sets in XML files. Domain specific XML specifications enable fast agreements between providing and consuming parties, and availability of XSDs supports quality checking on the data.

XML files are stored as text files, that means, machines and humans may read them. Applying special encodings, the files may contain Western European, Russian or even Chinese characters. No escaping is needed for such use cases.

While XML files themselves encapsulate data, XSDs can be seen as platform specific models (PSM) that define and support understanding of the relationships between domain specific entities and their attributes. Platform-independent models (PIM), such as UML class diagrams, in addition enable a more broaden comprehension and allow for further applications outside of the XML file transfer, e.g. internally in data base solutions.

The following sections introduce key XML formats and their models, impacting rail, which are capable of support storyboards presented in chapter 2.

3.2 RAILML®

railML® may support all three storyboards in different data classes as shown in Figure 3-2.

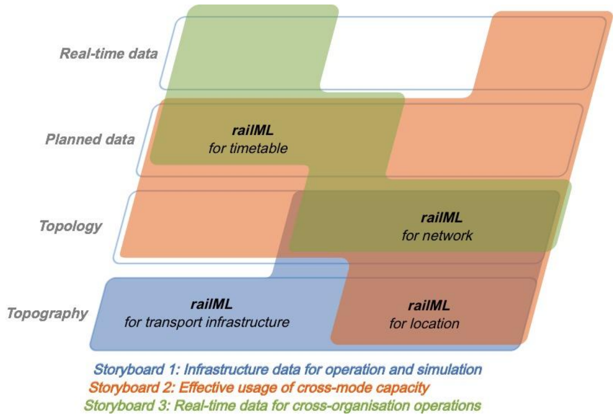

Figure 3-2 railML coverage for data classes and storyboards

railML is a data exchange format developed by a consortium of railway companies, academic institutions and consultancy firms. Formed in 2002, the railML.org project aims to continuously develop this format in order to facilitate its use in a wide range of railway applications.

The project started as a partnership between the Fraunhofer Institute for Transportation Systems and infrastructure (FhG-IVI) and the Swiss Federal Institute of Technology's Institute for Transportation Planning and Systems (Nash, Huerlimann, Schütte, & Krauss, 2004), and is currently coordinated by a small independent team. railML.org conferences are held twice a year and supplemented by specialized working group meetings (railML.org, 2012). At time of writing the officially published version of railML is 2.2, with version 3 under development (Table 3-1).

railML is published as a series of XML schemas holding subschemas, each of which encompasses a particular field of railway application:

- Common concepts and objects, sometimes not mentioned separately;
- Timetable (TT);
- Rolling stock (RS);
- Infrastructure (IS), both macroscopic and microscopic;
- Interlocking, from railML 3 on.

Table 3-1 railML development steps, adapted from (Nissi, Jeanmaire, Seybold, & et.al., 2013)

Subschema	railML version / year					
	0.x 2002	1.0 2005	1.1 2007	2.0 2009	2.2 2013	3.x 2015
Interlocking						Ready for daily use
Infrastructure / microscopic				First test, use cases	Elements added	Total reor-ganisation
Infrastructure / macroscopic		First test, use cases	Ready for daily use	Small changes	Elements added	Total reor-ganisation
Rolling stock	First test, use cases	Ready for daily use	Elements added	Small changes	Elements added	Small changes
Timetable	First test, use cases	Ready for daily use	Elements added	Total reor-ganisation	Elements added	Small changes

The current release of the data model (railML 2.2) focuses on information that has proven value in the railway operation planning process, but provides extension points for other topics or data of special interest.

Common schema

The common schema allows the definition of metadata using the well-known Dublin Core vocabulary, e.g. information about the application, which generated the data set. Furthermore author, date, identifier, and 'other' miscellaneous information may be given.

Timetable (TT) and rostering schema

The timetable and rostering schema supports various operational concepts, amongst others the coupling of trains. The rolling stock for a timetable has to be defined in the RS schema. All levels of granularity are supported by the same TT concept, allowing at minimum a name for the train material as well as defining a detailed combination of locomotives and wagons with track effort curves and seat places. This enables both passenger and freight transport services to be defined according to the same schema.

Time restrictions may be defined in a flexible way enabling both fixed and on-demand services. Cascading restrictions allow for general time periods for the whole timetable, further constrained at regular service level. Pre-defining holiday periods enables flexible pre- and post-holiday services that are required by freight trains engaged in multi-day journeys via closed terminals on those special days. Figure 3-3 shows an excerpt of a railML file with holiday and operating period settings for a schedule.

The TT schema is based on a railway infrastructure defined within the IS schema. At minimum operational points have to be provided, e.g. declaring its name. If more information is available, lines with mileages link those operational points. At the most detailed level, a fine-grained microscopic track network at switch level is offered. All those usage models are supported by the same timetable concept containing placeholders for detailed IS references if they are available.

```
<timetable id="tt01">
    <timetablePeriods>
        <timetablePeriod id="ttp01" description="Fahrplan 2015" startDate="2014-12-14" endDate="2015-03-28"
            <holidays>
                <holiday holidayDate="2014-12-25" description="1. Weihnachtstag" />
                <holiday holidayDate="2014-12-26" description="2. Weihnachtsfeiertag" />
                <holiday holidayDate="2015-01-06" description="Heilige Drei Koenige" />
                <holiday holidayDate="2015-04-03" description="Karfreitag" />
                <holiday holidayDate="2015-04-06" description="Ostermontag" />
                <holiday holidayDate="2015-05-01" description="Tag der Arbeit" />
                <holiday holidayDate="2015-05-14" description="Christi Himmelfahrt" />
                <holiday holidayDate="2015-05-25" description="Pfingstmontag" />
                <holiday holidayDate="2015-06-04" description="Fronleichnam" />
                <holiday holidayDate="2015-08-15" description="Mariae Himmelfahrt" />
                <holiday holidayDate="2015-10-31" description="Tag der deutschen Einheit" />
                <holiday holidayDate="2015-11-01" description="Allerheiligen" />
            </holidays>
        </timetablePeriod>
    </timetablePeriods>
    <operatingPeriods>
        <operatingPeriod id="op01" description="Montag bis Freitag" timetablePeriodRef="ttp01">
            <operatingDay operatingCode="1111100"/>
        </operatingPeriod>
        <operatingPeriod id="op02" description="täglich" timetablePeriodRef="ttp01">
            <operatingDay operatingCode="1111111"/>
        </operatingPeriod>
        <operatingPeriod id="op03" description="nicht täglich, 6. Jan bis 27. Mär 2015 Mo-Fr; 24. Aug bis 1
            <operatingDay operatingCode="1111111" startDate="2014-12-14" endDate="2015-01-05"/>
            <operatingDay operatingCode="1111100" startDate="2015-01-06" endDate="2015-03-27"/>
            <operatingDay operatingCode="1111111" startDate="2015-03-28" endDate="2015-05-28"/>
        </operatingPeriod>
```

Figure 3-3 railML file – example "operating period" (Excerpt)

INFRASTRUCTURE (IS) SCHEMA

The infrastructure schema defines properties and structures for railway facilities. Key concepts are:

- Network topology with operational points and lines ('macroscopic' graph in railML);
- Track topology with switches, crossings and track sections ('microscopic' node and edge);
- Operation and control elements, such as signal, balise, axle counter, level crossing;
- Civil engineering structure, e.g. bridges, tunnels;
- Track capability in form of 'change point', e.g. line speeds, gradients, presence and kind of electrification;
- Further concepts to model infrastructure visualisation at simulator software.

RailML 2.2 is intended to define macroscopic graphs the same way as microscopic nodes and edges. This concept built a robust foundation for a long time, where an infrastructure model is exchanged along with an appropriate timetable.

Since railway asset management moves into the focus of railML infrastructure development, the approach had to be changed in order to save both microscopic and macroscopic infrastructure data tightly coupled within one file. This movement is based on the platform-independent UIC RailTopoModel (chapter 3.3) and leads to railML 3 in form of XSDs for real XML file exchange. Final railML 3 is not available so far, but its mockup schemas seam already promising.

ROLLING STOCK (RS) SCHEMA

The rolling stock schema allows the representation of locomotives, multiple units, and passenger and freight wagons at various levels of detail. Factors such as propulsion type, braking ability, and mechanical traction losses can also be captured along with many other attributes allowing trains to be physically modelled in great detail.

As a basic concept, each rolling stock unit that may be coupled outside of a workshop is defined as a single 'vehicle' or as instance of a vehicle family. A composition of 'vehicles' is called a 'formation' – best known as a 'train'. But in the railML context a 'train' is used for the timetabling perspective. The 'formation' is referred from within the timetable for each 'train'.

RAILML 3 DEVELOPMENT

While railML 2.2 is largely mature, railML 3 is undergoing active development in close cooperation with the UIC RailTopoModel (Chapter 3.3) following a decent 'use case' concept (Figure 3-4). The creation of subschemas for the railML model will be driven by use cases, which will be provided by the railML partners and modelled in UML class diagrams.

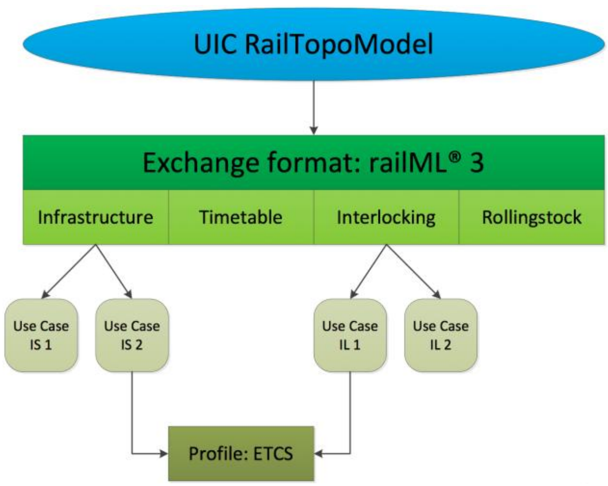

Figure 3-4 Use case concept for railML 3 development (railML.org, 2015)

A use case describes the implementation of railML 3 elements and attributes for a specific field of application. Therefore several software tools may share the same use case, while a single tool will often implement several use cases. This approach encourages stricter interface implementations for the pre-defined use cases, providing better guidance to the programmer and more reliability to the costumer for the daily data exchange between different software tools wile retaining the standard highly flexible for future needs.

Figure 3-5 gives an overview of railway related topics that will guide the railML 3 development. The dark green themes will be incorporated as core competences in railML 3 already, whereas the light green boxes will follow with next versions.

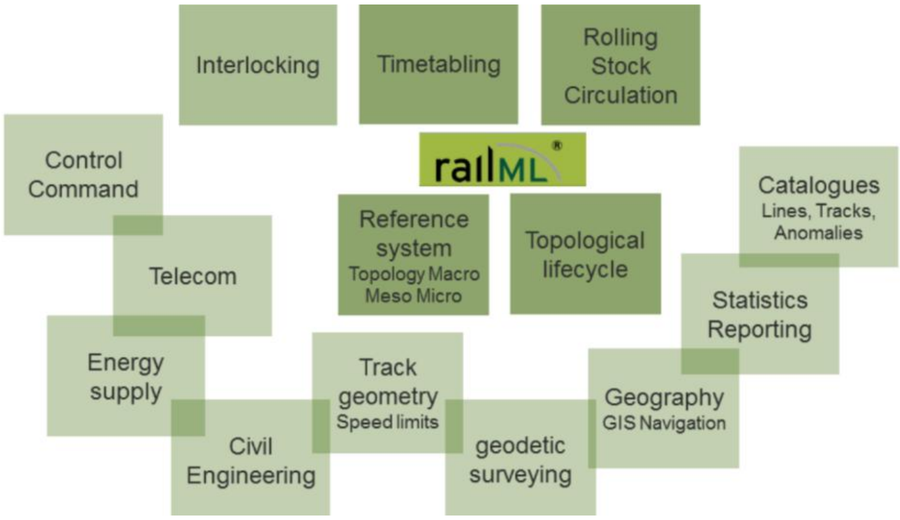

Figure 3-5 railML 3 development topics (UIC, 2014)

Figure 3-6 shows the development strategy for fulfilling the different user needs with railML 3:

- Providing a *base vocabulary* in a clear structure and unambiguous syntax for each topic as detailed as needed; most attributes and elements are optional for highest flexibility;
- Providing *profiles* according to the use cases of railML interest groups constraining needed parts and skipping out-of-scope elements fully relying on the railML base vocabulary;
- *Individual subsets* developed by railML partners for their own purposes using parts of the railML base vocabulary constrained by their specific rules.

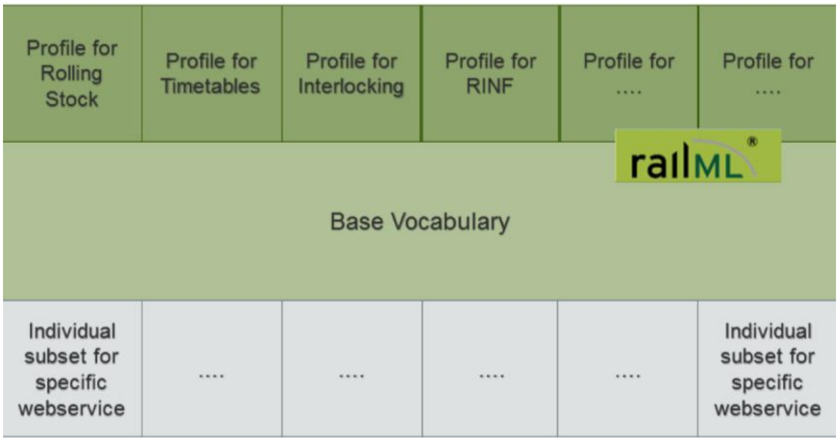

Figure 3-6 railML 3 architecture (UIC, 2014)

USER GROUPS AND USE CASES

railML has a number of user groups, each of which have their own specific requirements. One strong user group takes current railML as an import and export data format for railway simulation systems, such as OpenTrack and HERMES. Another substantial usage of current railML is enabling sound input and output in production software, e.g. Viriato, OpenTimeTable, TPS, FBS, IVU.rail.

The railML.org project is supported by a number of European railway operators, manufacturers, and infrastructure managers; these include Alstom, Deutsche Bahn, SBB, Siemens and Infrabel amongst others.

Table 3-2 Use cases for railML 3 (railML.org, 2015)

'IS' use cases	Responsibility	'TT' use cases (excerpt)
RINF	SNCF Réseau	**Timetable information**
NRE reporting	ÖBB	**Passenger information at stations / at vehicles**
ETCS	Infrabel	**Timetable for competition (call for proposal)**
Speed directory	ÖBB	**Timetable for vehicle working scheduling**
Capacity planning	Jernbaneverket	**Formation data**
Positioning system	DLR	**Operational simulation regarding feasibility**
Interlocking	Deutsche Bahn	**Long-term planning**
Driver advisory system	Network Rail	**Train path ordering and re-allocation**
Infrastructure recording	Bahnkonzept	**Rolling stock rostering / vehicle disposition**
Passenger information	BLS	**TAF / TAP**
Maintenance planning	SBB, BLS	**Intermodal Transport Control System (ITCS)**

The RINF reporting process is defined as one use case for railML 3, see also section 0.

railML will be further analysed focussing on the storyboard requirements in the following chapters:

railML		
Storyboard 1 **Infrastructure data for operation and simulation**	Storyboard 2 **Effective usage of cross mode capacity**	Storyboard 3 **Real-time data for cross-organisation operations**
2020 → chapter 4.2	2020 → chapter 5.2	2020 → chapter 6.2
2030 → chapter 4.3	2030 → chapter 0	2030 → chapter 6.3
2050 → chapter 4.4	2050 → chapter 5.4	2050 → chapter 6.4

3.3 UIC RailTopoModel / RailML 3

UIC RailTopoModel plays a central role for all three storyboards providing a common network reference as shown in Figure 3-7.

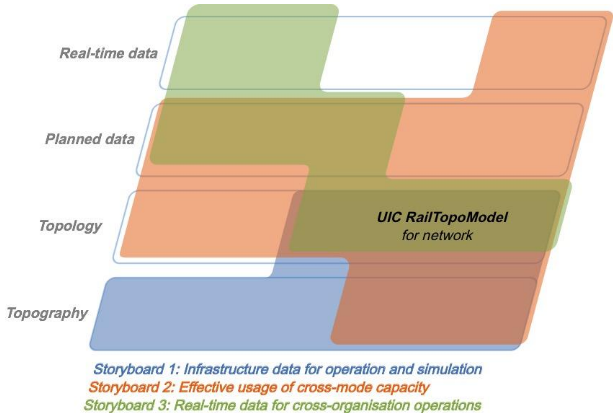

Figure 3-7 UIC RailTopoModel coverage for data classes and storyboards

The UIC RailTopoModel is being pushed by the UIC (International Union of Railways) in close cooperation with the railML® consortium as a base topological model for railway infrastructure. railML 3 (section 3.2) shall provide a reference implementation for UIC RailTopoModel through an XSD schema enabling a standard data exchange. Since 2013 the UIC RailTopoModel conference and the railML.org users group meeting are held twice a year at the UIC headquarter in Paris on consecutive days.

Many operational data formats for railways are legally mandated, e.g. RINF or INSPIRE. The design of each format primarily focuses on its specific requirements, with INSPIRE targeting mesoscopic data and RINF used for the more detailed microscopic data. Nevertheless any data format requires qualified topological data for railway networks. So far, each infrastructure manager (IM) developed specific models and interfaces by its own (Figure 3-8).

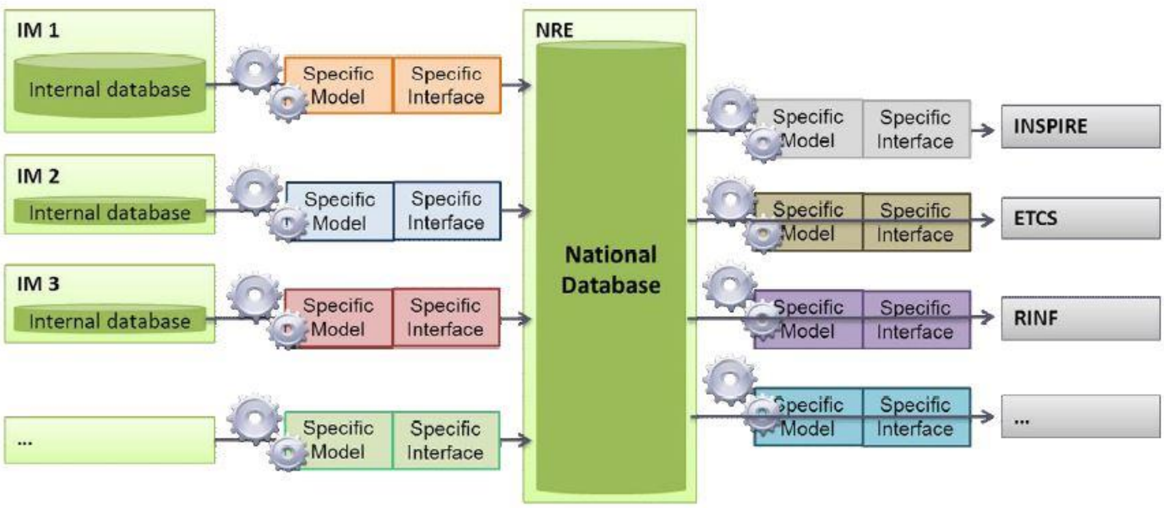

Figure 3-8 Data Models – Current national situation (railML.org, 2015)

The UIC RailTopoModel is focussed on the following concepts:

- Reducing duplication of effort and encouraging collaboration between stakeholders;
- Preventing laboured and repetitive developments in IT;
- Reducing lengthy IT project development phases;
- Enabling innovation;
- Improving compatibility by reducing overlap and redundancy issues.

The model will help the railway sector to become a more competitive market, with the fast and efficient exchange of data between companies, their industrial suppliers, and railway regulation bodies and other authorities (ERIM Workgroup, 2014). Figure 3-9 illustrates this new situation.

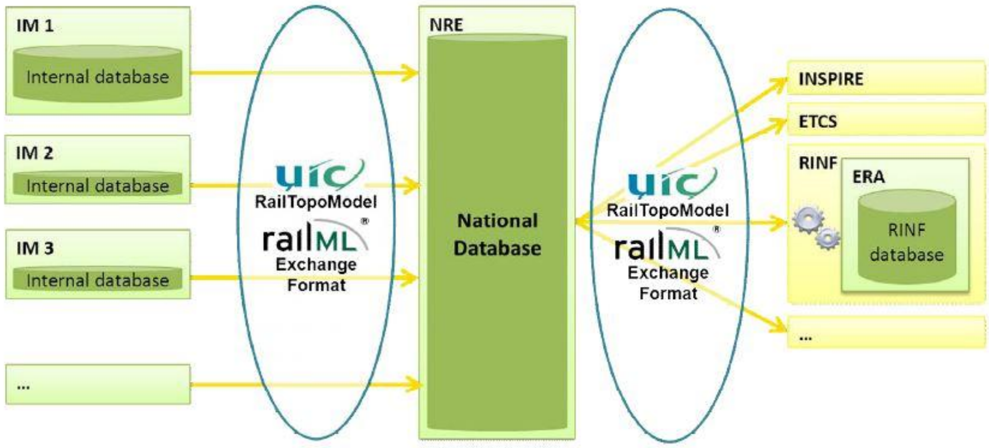

Figure 3-9 Ideal national situation with the UIC RailTopoModel and railML® (railML.org, 2015)

The UIC RailTopoModel guarantees the ability to aggregate from the microscopic through mesoscopic up to macroscopic infrastructure data (see chapter 2.1 for granularities). These granularities may be flexibly combined with each other in one data set, offering flexible data handling for simulations based on detailed microscopic data in stations and less detailed macroscopic data on lines. (Gely, Dessagne, Pesneau, & Vanderbeck, 2008) This general concept may also be used for filling gaps with less detailed open data from different sources in case of missing detailed IM data.

UIC RailTopoModel will be further analysed focussing on the storyboard requirements in the following chapters:

UIC RailTopoModel		
Storyboard 1 **Infrastructure data for operation and simulation**	Storyboard 2 **Effective usage of cross mode capacity**	Storyboard 3 **Real-time data for cross-organisation operations**
2030 → chapter 4.3	2030 → chapter 0	2030 → chapter 6.3
2050 → chapter 4.4	2050 → chapter 5.4	2050 → chapter 6.4

3.4 RINF – Register of INFrastructure

RINF may support storyboard 1 providing detailed infrastructure items as shown in Figure 3-7.

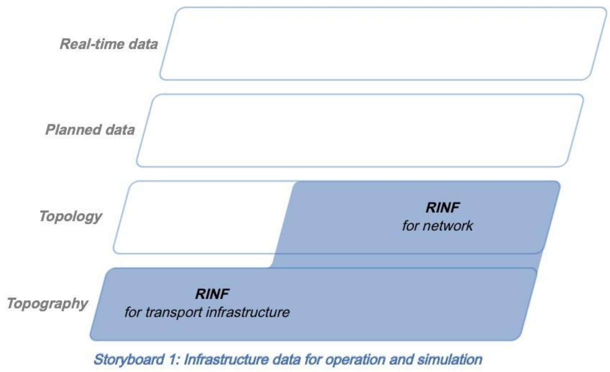

Figure 3-10 RINF coverage for data classes and storyboards

The European register of infrastructure has been introduced on the legal basis of Article 35 of *Interoperability Directive* 2008/57/EC followed by the common technical specifications in a Commission Implementing Decision (RINF Decision) (ERA, 2015).

Purpose of RINF

RINF is intended to contain and provide data about important features of mainline railway infrastructure, and has to be implemented in the context of technical specifications that support interoperability on the railway networks within the European community. TSIs are developed to achieve accepted specifications for both TEN and non-TEN-lines (according

Directive 2008/57) for each structural subsystem such as infrastructure, rolling stock or functional subsystem like operation and traffic management.

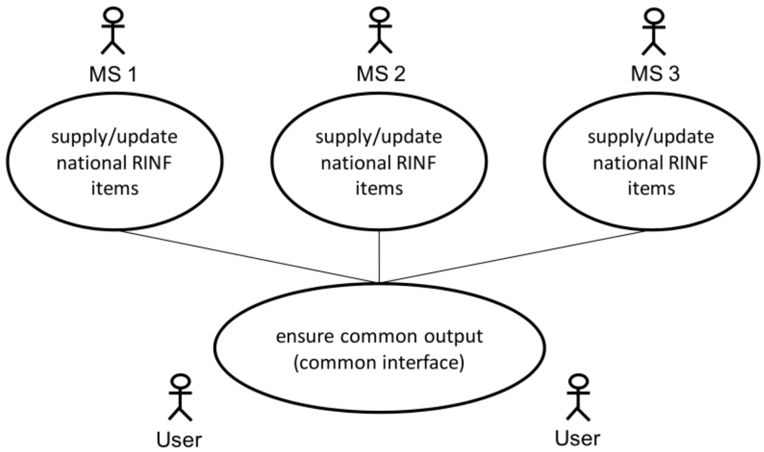

Figure 3-11 Basic RINF System-architecture (ERA, 2010)

According to Figure 3-11, each Member State (MS), represented by a National Registry Entity (NRE), shall supply and update the national RINF items to the centralized part of the RINF system, the common user interface (CUI). The RINF implementation involves the main activities, described in Table 3-3.

Table 3-3 RINF main activities, adapted from (Joint CER/EIM working group on RINF, 2013)

Order	Responsible	Activity
1.	IM	Data collection
2.	NRE	Structuring and integration of IM's data collections to a national data collection for the MS
3.	NRE	Data extract from the national data collection and file transfer to the ERA
4.	ERA	Managing and maintaining the CUI, enabling searches and retrieving RINF information to the users

An RINF XSD, requiring the use of XML files, defines the interface between MS and CUI. There are no regulations about the data exchange format neither between IMs and its NRE nor between CUI and RUs.

Data collection and structuring is based on RINF item specifications (objects and properties). Data structuring is delegated to a national data model. Data extract and file building is designed based on a proprietary RINF file format as currently described in the RINF application guide. The joint CER/EIM working group on RINF has requested that the ERA elaborates on the intended process for achieving this objective [a standardised data exchange format] together with experts from the railway sector.

The approach gives an opportunity for IMs to develop their data extracts and formatting as a reusable investment, not only for RINF purposes, but also for any other future needs of infrastructure data exchanges (this development, currently viewed by IMs as a cost could then be presented as an investment). Moreover, the strength of law for RINF deployment would give the railway sector and ERA a unique opportunity to roll out a Europe-wide universal standard, thus increasing interoperability and productivity. (Joint CER/EIM working group on RINF, 2013)

RINF RAILWAY NETWORK TOPOLOGY

For the purpose of RINF the railway network is considered to be a series of operational points (OPs) connected by sections of line. Further to this:

- A line is a sequence of one or more sections, which may consist of several tracks;
- A section of line is the part of line between adjacent OPs and may consist of several tracks;
- Operational points are locations for train service operations for example where train services can begin and end, change route and where passenger or freight services are provided;
- Stopping points for passengers on plain line are also regarded as OPs;
- Operational points may be locations where the functionality of basic parameters of a subsystem are changing for example: track gauge, voltage and frequency, signalling system;
- Operational points may be at boundaries between MSs or IMs;
- Passing loops and meeting loops on plain line or track connections only required for train operation do not need to be published;
- Sidings are all tracks not used for train service movements.

(ERA, 2010)

Figure 3-12 shows the general concept of the railway network structure of RINF. Elements are coloured after their IMs.

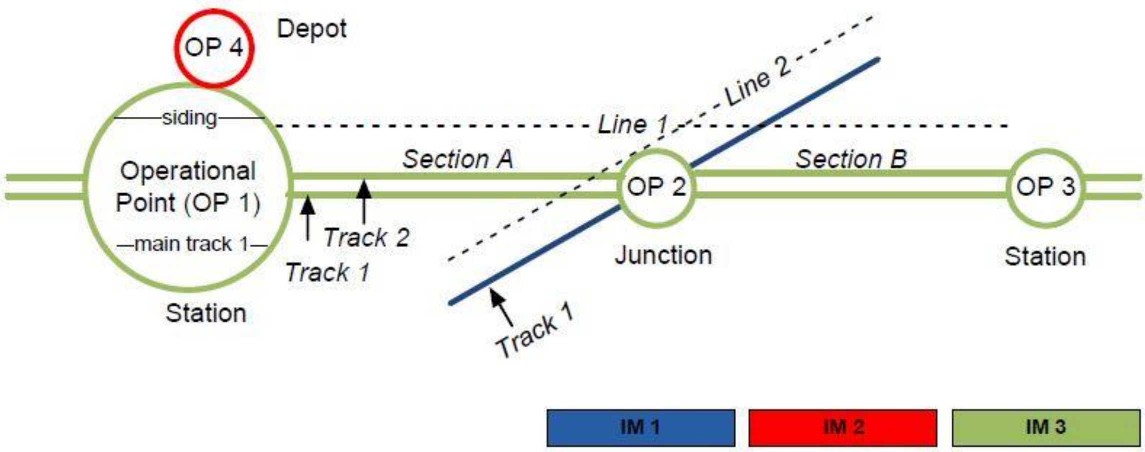

Figure 3-12 RINF structure of the railway network for the register (ERA, 2010)

In summary, a detailed network infrastructure description includes the network topology and 158 other parameters (Nissi, Jeanmaire, Seybold, & et.al., 2013).

RINF will be further analysed focussing on the storyboard requirements in the following chapters:

RINF		
Storyboard 1 **Infrastructure data for operation and simulation**	Storyboard 2 **Effective usage of cross mode capacity**	Storyboard 3 **Real-time data for cross-organisation operations**
2020 → chapter 4.2 2030 → chapter 4.3 2050 → chapter 4.4		

3.5 INSPIRE

INSPIRE may support storyboard 1 in providing a common network structure for different transport modes as shown in Figure 3-13.

The Directive 2007/2/EC of the European parliament and of the council establishing an Infrastructure for Spatial Information in the European Community (INSPIRE) came into force on 15 May 2007 and will be implemented in various stages, with full implementation required by 2020 (EC, 2007).

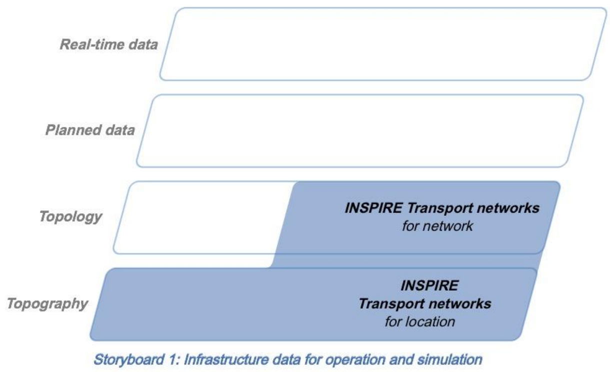

Figure 3-13 INSPIRE Transport networks coverage for data classes and storyboards

INSPIRE is based on a number of common principles (INSPIRE, 2014):

- Data should be collected only once and kept where it can be maintained most effectively;
- It should be possible to combine seamless spatial information from different sources across Europe and share it with many users and applications;
- It should be possible for information collected at one level/scale to be shared with all levels/scales; detailed for thorough investigations, general for strategic purposes;
- Geographic information needed for good governance at all levels should be readily and transparently available;
- Easy to find what geographic information is available, how it can be used to meet a particular need, and under which conditions it can be acquired and used.

The INSPIRE directive addresses 34 spatial data themes needed for environmental applications, wherein "Transport networks" is one theme of Annex I (Figure 3-14).

Figure 3-14 INSPIRE Data Specification – Spatial Data Themes (INSPIRE, 2014)

Member States (MS) have to make their data available according to the implementing rules within two years of the adoption for newly collected and extensively restructured data and within five years for other data in electronic format still in use, see also Figure 3-15. However, the effective implementation of the directive varies from country to country.

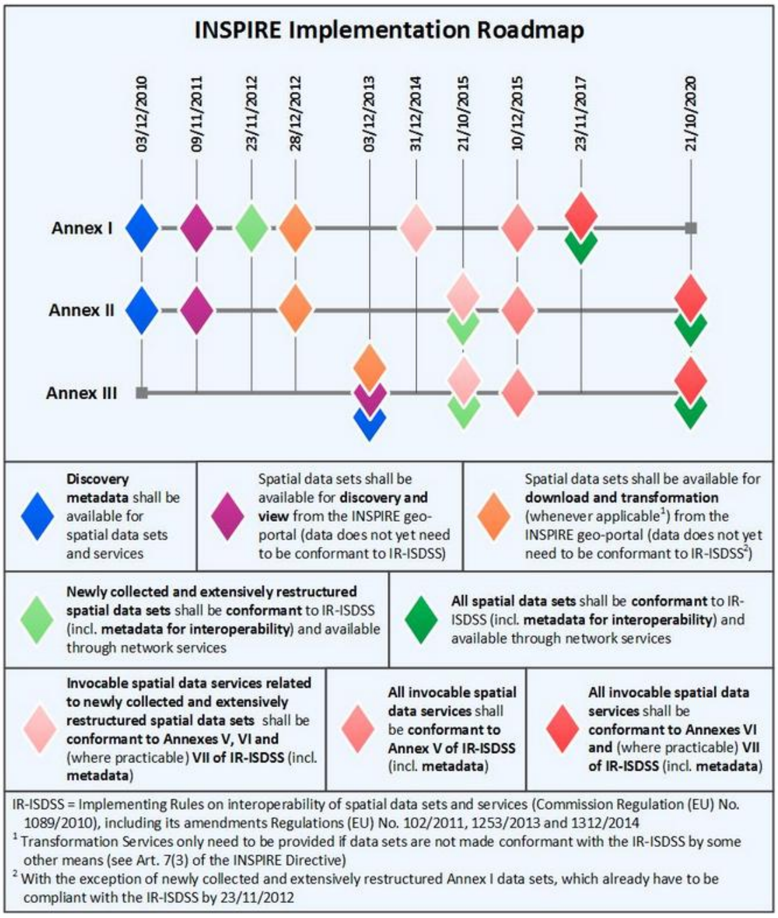

Figure 3-15 INSPIRE Implementation roadmap (INSPIRE, 2014)

The INSPIRE Thematic Working Group, Transport Networks, developed a technical guideline for data specification, which includes implementing rules for transport networks in general and an application schema for selected transport networks (TWG-TN, 2014):

•

Figure 3-16 illustrates the main objects of the Rail Transport Network.

In 2011, the INSPIRE geoportal website was launched: http://inspire-geoportal.ec.europa.eu. It provides the means to search for spatial data sets and spatial data services, and subject to access restrictions, to view spatial data sets from the MS within the framework of the INSPIRE directive.

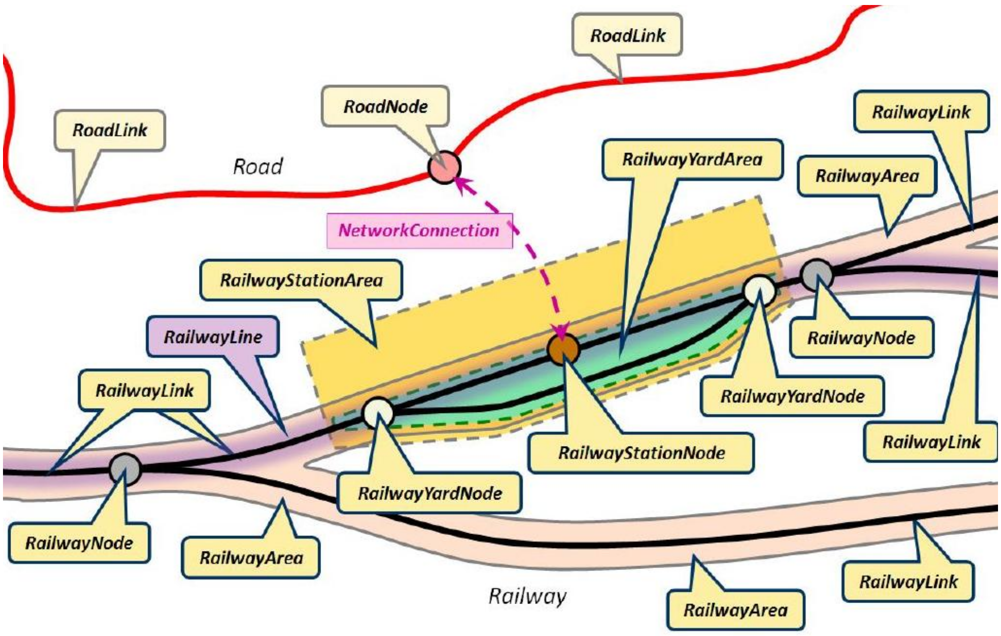

Figure 3-16 INSPIRE Overview of the main Rail Transport Networks objects (TWG-TN, 2014)

Some transport networks may be viewed without any privileges, e.g. rail transport network of ÖBB in Austria. Some data sets are restricted to defined user groups, e.g. railway network of Trafikverket in Sweden. Other IMs have not transferred their data sets yet due to legal aspects, e.g. rail network of Deutsche Bahn in Germany.

INSPIRE rail transport network will be further analysed focussing on the storyboard requirements in the following chapters:

INSPIRE		
Storyboard 1 **Infrastructure data for operation and simulation**	Storyboard 2 **Effective usage of cross mode capacity**	Storyboard 3 **Real-time data for cross-organisation operations**
2020 → chapter 4.2 2030 → chapter 4.3		

3.6 IDM^VU

IDM^VU may support storyboard 1 providing detailed infrastructure facilities relevant for maintenance as shown in Figure 3-17.

IDM^{VU} is a data model and data exchange format describing “infrastructure data management for transportation companies“, in particular regarding planning, construction, operation, maintenance and disposal of rail infrastructure facilities. (IDMVU.org, 2010)

IDM^{VU} is designed with the help of UML class diagrams for clear understanding of relationships and attributes. Based on the IDM^{VU} model the data exchange format IDM-GML has been developed, which forms a GML 3.2.1 application schema. The IDM^{VU} model and IDM-GML have been published as VDV-Recommendation 456 “*Interface Standard Infrastructure Data Management (IDM^{VU})*”.

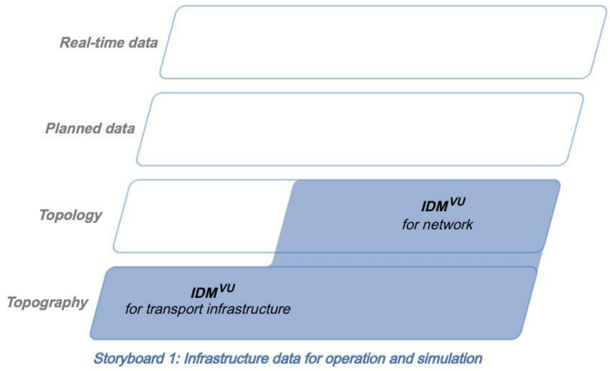

Figure 3-17 IDM^{VU} coverage for data classes and storyboards

The IDM^{VU} model describes network topology using a node-edge graph (see Figure 3-18).

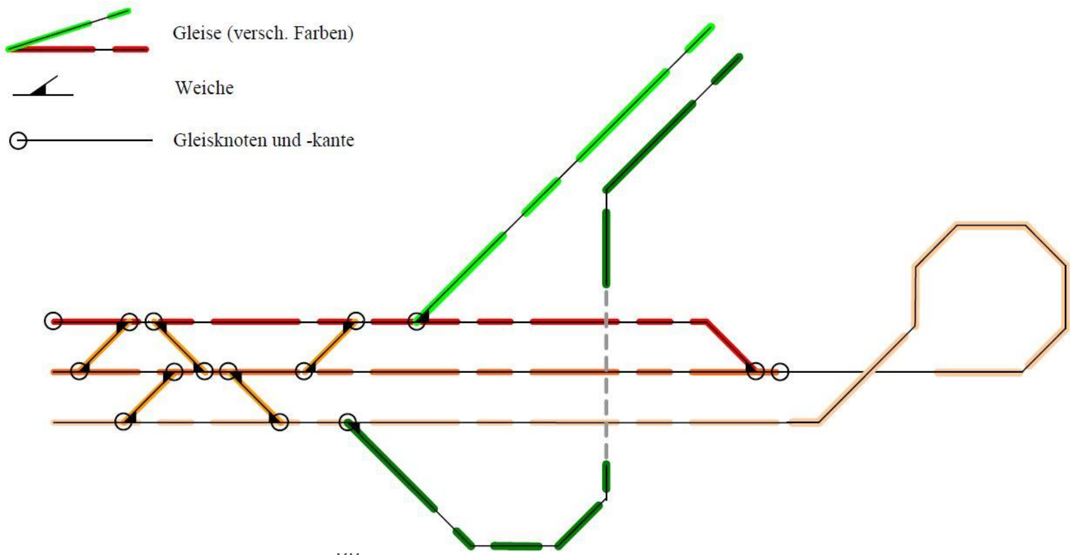

Figure 3-18 IDM^{VU} Drawing network topology (VDV456, 2014-05)

The top level [of the hierarchical structure] consists of the following object fields:

- Network model;
- Track formation;
- Power supply;
- Control/command and signalling systems;
- Stopping places;
- Structures;
- Cables and pipes;
- Telecommunications;
- Real estate;
- Depots and workshops;
- Emergency facilities;
- General objects;
- State data;
- Operational data;
- Commercial data.

The object fields of the IDMVU data model mentioned above have not been modelled to the same depth in all parts. The modelling was based on the requirement that it has to be possible for all public transport infrastructure managers to use the IDMVU data model. (VDV456, 2014-05)

IDM^{VU} will be further analysed focussing on the storyboard requirements in the following chapters:

IDM^{VU}		
Storyboard 1 **Infrastructure data for operation and simulation**	Storyboard 2 **Effective usage of cross mode capacity**	Storyboard 3 **Real-time data for cross-organisation operations**
2020 → chapter 4.2		

3.7 NeTEx - Network and Timetable Exchange

NeTEx may support storyboard 2 providing all relevant data structures and baseline models for storyboard 3 as shown in Figure 3-19.

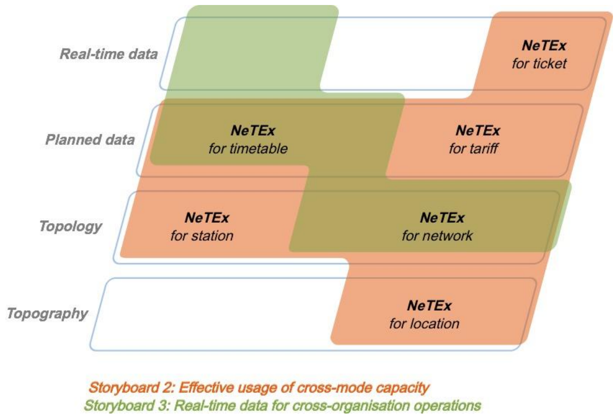

Figure 3-19 NeTEx coverage for data classes and storyboards

In 2014, the European Committee for Standardization (CEN) published the specification "Public transport - Network and Timetable Exchange (NeTEx)" for the harmonization of various standards that support public transport data exchange between interested parties. NeTEx is based on "Transmodel V5.1: EN 12986: 2006 Road transport and traffic telematics – Public transport – Reference data model" and "IFOPT EN 28701: 2012 Intelligent transport systems – Public transport – Identification of Fixed Objects in Public Transport". It complements and therefore may be used in combination with "SIRI CEN/TS 15531-1 to 5: 2011 Public transport – Service interface for real-time information relating to public transport operations" (see section 3.8).

NeTEx aims on providing an efficient European standard for exchanging Public Transport (PT) schedules and related data in XML files fulfilling the developed XSD schemas. Therefore it shall be capable for any mass PT mode, e.g. train, bus, metro, tramway, ferry, their sub modes; specific demands for air traffic are not considered so far.

NeTEx will facilitate interoperability between IT systems of involved PT parties by:

- Introducing common architectures for message exchange;
- Introducing a modular set of compatible services for real-time vehicle information;
- Using common data models and schemas for the messages exchanged for each service;
- Introducing a consistent approach to data management.

NeTEx covers the topics framework, network topology, timetables and tariffs (Figure 3-20).

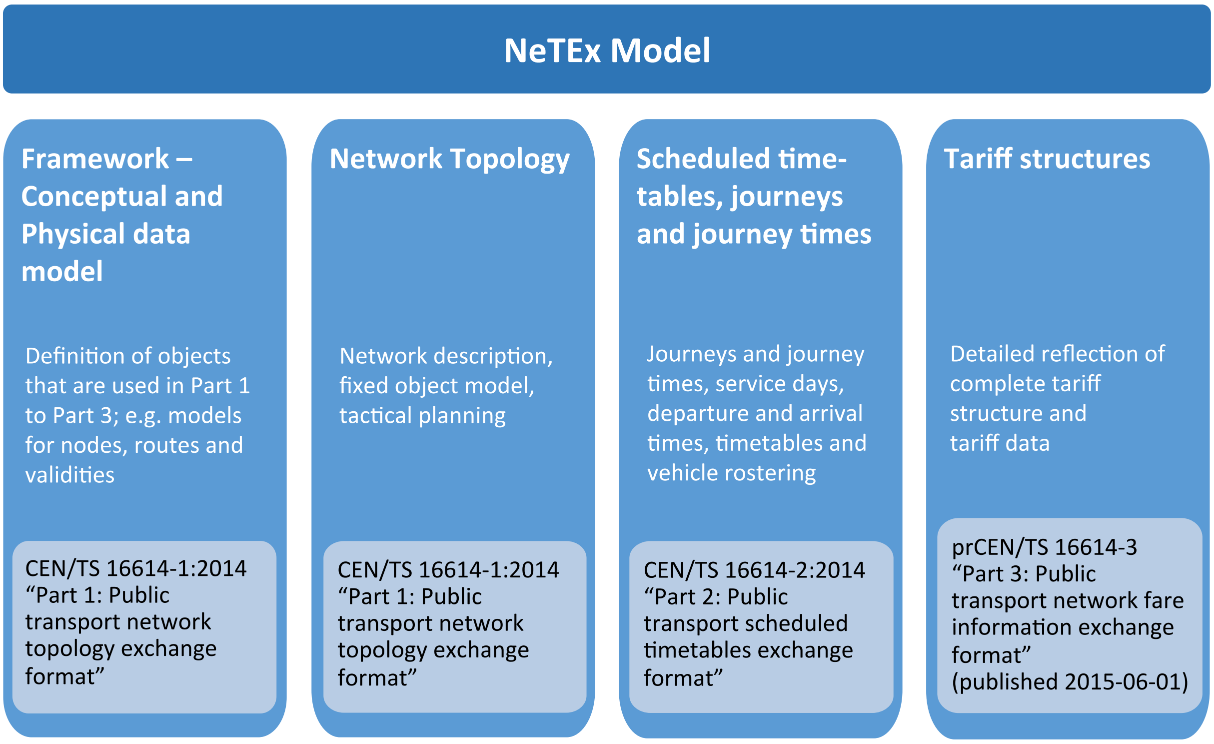

Figure 3-20 NeTEx Model

The development of NeTEx was pushed by United Kingdom (Department of Transport), Germany (VDV), France (CERTU, STIF) and Finland (VTT). Other countries were also involved, e.g. Norway, Sweden, Netherlands, Italy, Hungary and Switzerland. The concepts of already established country-specific standards were fully integrated into the NeTEx model.

NeTEx can be used to exchange (NeTEx, 2014):

- Public Transport schedules including stops, routes, departures times / frequencies, operational notes, and map coordinates;
- Routes with complex topologies such as circular routes, cloverleaf and lollipops, and complex workings such as short working expressed in patterns;
- Connections with other services;
- The days on which the services run, including availability on public holidays and other exceptions;
- Composite journeys such as train journeys that merge or split trains;
- Information about the Operators providing the service;
- Additional operational information, including positioning runs, garages, layovers, duty crews, useful for Automated Vehicle Location (AVL) and on-board ticketing systems;
- Data about the accessibility of services to passengers with restricted mobility;
- Management metadata for versioning data sets allowing updates across distributed systems;
- Fare structures, incorporating flat fares, point to point fares, zonal fares;
- Fare products, incorporating single tickets, return tickets, day and season passes;
- Fares, only applying at specific dates.

The modelling approach for NeTEx incorporating existing models is shown in Figure 3-21.

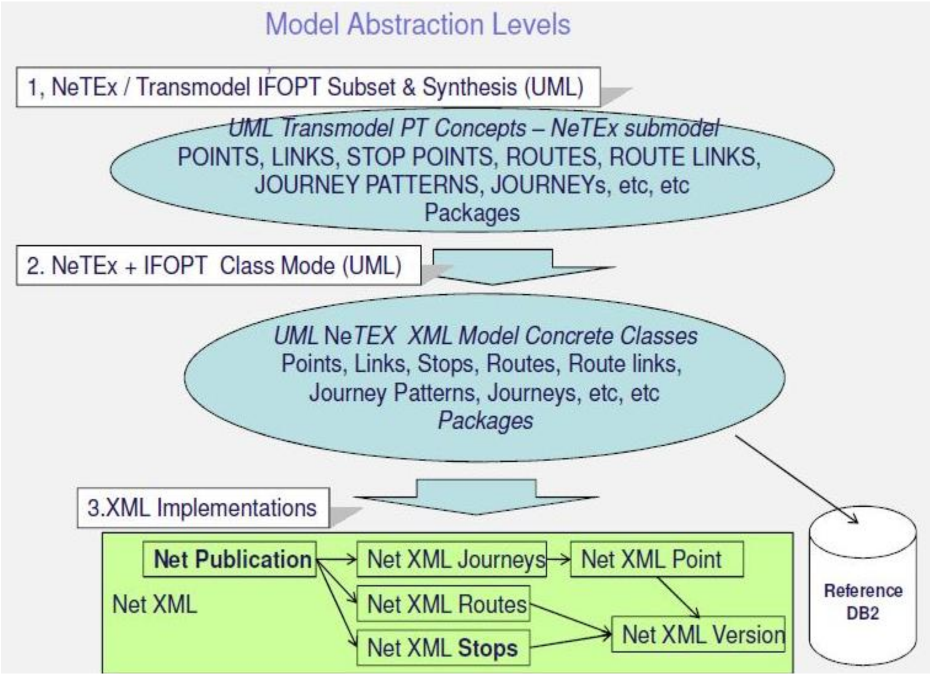

Figure 3-21 NeTEx Modelling Approach (CEN/TS16614-1, 2014)

Reusing a common communication layer shared with SIRI for various technical services makes the NeTEx standard readily extensible in future. For instance, Figure 3-22 illustrates parts of an XML file that define temporal validities.

```xml
<ServiceCalendarFrame version="any" id="SCF01">
    <ServiceCalendar version="any" id="SC01">
        <FromDate>2014-12-14</FromDate>
        <ToDate>2015-03-28</ToDate>
        <dayTypes>
            <DayType version="any" id="DT01">
                <Name>Montag bis Freitag</Name>
                <properties>
                    <PropertyOfDay>
                        <DaysOfWeek>Monday Tuesday Wednesday Thursday Friday</DaysOfWeek>
                    </PropertyOfDay>
                </properties>
            </DayType>
            <DayType version="any" id="DT02">
                <Name>täglich</Name>
                <properties>
                    <PropertyOfDay>
                        <DaysOfWeek>Everyday</DaysOfWeek>
                    </PropertyOfDay>
                </properties>
            </DayType>
            <DayType version="any" id="DT03"> [8 lines]
            <DayType version="any" id="DT04"> [8 lines]
        </dayTypes>
    </ServiceCalendar>
</ServiceCalendarFrame>
<TimetableFrame version="any" id="TTF01">
    <validityConditions>
        <AvailabilityCondition version="any" id="AC01">
            <Description>Fahrplan gültig vom 14. Dezember 2014 bis 28. März 2015</Description>
            <FromDate>2014-12-14T00:00:00Z</FromDate>
            <ToDate>2015-03-28T24:00:00Z</ToDate>
            <IsAvailable>true</IsAvailable>
        </AvailabilityCondition>
    </validityConditions>
    <contentValidityConditions>
        <AvailabilityCondition version="any" id="AC02"> [23 lines]
        <AvailabilityCondition version="any" id="AC03">
            <Description>Montag bis Freitag</Description>
            <FromDate>2014-12-14T00:00:00Z</FromDate>
            <ToDate>2015-03-28T24:00:00Z</ToDate>
            <dayTypes>
                <DayTypeRef ref="DT01"/>
            </dayTypes>
```


Figure 3-22 NeTEx XML file – example "Availability condition" (Excerpt)

NeTEx will be further analysed focussing on the storyboard requirements in the following chapters:

NeTEx		
Storyboard 1 **Infrastructure data for operation and simulation**	Storyboard 2 **Effective usage of cross mode capacity**	Storyboard 3 **Real-time data for cross-organisation operations**
	2020 → chapter 5.2 2030 → chapter 0 2050 → chapter 5.4	2020 → chapter 6.2 2030 → chapter 6.3 2050 → chapter 6.4

3.8 SIRI – Service Interface for Real-time Information

SIRI may support storyboard 3 providing relevant data structures as shown in Figure 3-23 relying on NeTEx base data.

The Technical Specification "Public transport – Service Interface for Real-time Information relating to public transport operation (CEN/TS 15531)" was first published in 2006 in 3 parts and became a CEN Technical Standard in 2007. Two further parts extended it in 2011. Since 2013, part 1 to 3 are published as draft European Standard (prEN 15531 rev) that are intended to supersede the CEN/TS 15531 documents of 2007.

SIRI provides a framework for specifying communications and data exchange protocols in order to exchange real-time information related to public transport operations. SIRI is designed as a modular and expansible standard that uses XSD schemas.

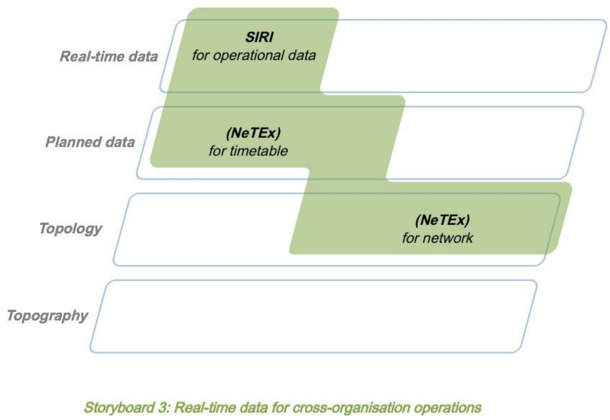

Figure 3-23 SIRI coverage for data classes and storyboards

As NeTEx is developed as a complementary format to SIRI, the revised version of SIRI includes references to NeTEx (see section 3.7). SIRI refers to "Transmodel V5.1: EN 12986: 2006 Road transport and traffic telematics – Public transport – Reference data model", wherever possible, which is also a basis for NeTEx. (prEN15531-1, 2013).

SIRI covers the topics framework, communications and functional service interfaces (Figure 3-24).

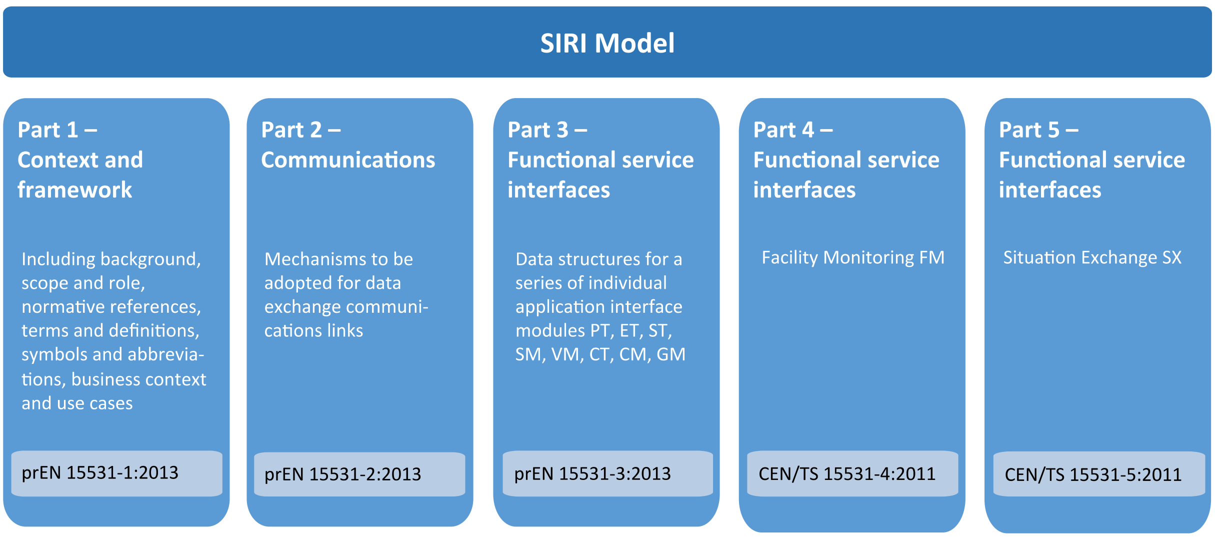

Figure 3-24 SIRI-Model

SIRI specifies several functional modules, based on 'use cases' identified in Annex B of Part 1 (Table 3-4), which are taken in part 3 to 5 presenting functional service interfaces in order to exchange data. For each of these functional services, SIRI defines purposes, capability and permission matrices, service requests, their related elements and provides examples.

Table 3-4 SIRI Functional Services (prEN15531-1, 2013)

	Service	Description
PT	Production Timetable	To send daily information on the operational timetable and associated vehicle running information
ET	Estimated Timetable	To send real-time information on timetable, including changes based on the production service and on actual running conditions
ST	Stop Timetable	To provide a stop-centric view of timetabled vehicle arrivals and departures at a designated stop
SM	Stop Monitoring	To send real-time arrival and departure information relating to a specific stop
VM	Vehicle Monitoring	To send real-time information on the movement and predicted movement of vehicles
CT	Connection Timetable	To send an operational timetable for a service feeding an interchange, in order to inform departing services of the possible need to wait for connecting passengers
CM	Connection Monitoring	To send real-time information on the running of a service inbound to an interchange, in order to advise departing services of the need to wait for connecting passengers; this can also be used to send real-time information to assist passengers in planning their onward journey following a connection
GM	General Message	To exchange informative messages between participants
FM	Facilities Management	To exchange information on the current status of facilities such as lifts, escalators or ticketing machines (additional compared to CEN/TS 15531- 1: 2007)
SX	Situation Exchange	To exchange information messages between identified participants in a standardised format suitable for travel information services (additional compared to CEN/TS 15531- 1: 2007)

Within the TAF TSI documents (see section 0), SIRI is explicitly named as standard to provide and exchange real-time timetable information with other modes of transport (European Commision, 2011).

SIRI will be further analysed focussing on the storyboard requirements in the following chapters:

SIRI

Storyboard 1 **Infrastructure data for operation and simulation**	Storyboard 2 **Effective usage of cross mode capacity**	Storyboard 3 **Real-time data for cross-organisation operations**
		2020 → chapter 6.2 2030 → chapter 6.3 2050 → chapter 6.4

3.9 TAF TSI

TAF TSI may support storyboard 3 providing train related data structures as shown in Figure 3-25.

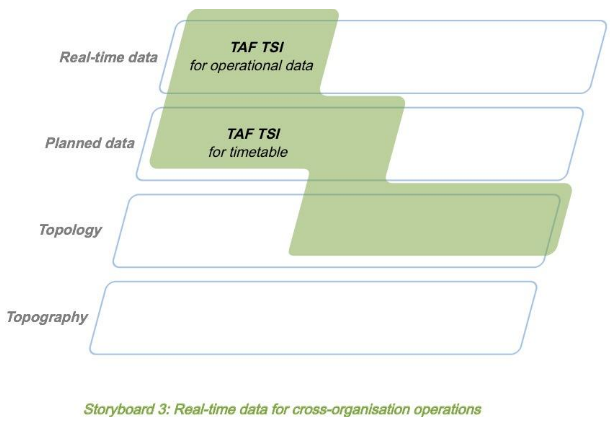

Figure 3-25 TAF TSI coverage for data classes and storyboards

The purpose of the „Technical Specifications for Interoperability for Telematics Applications for Freight" TAF TSI is to ensure the efficient interchange of information by specifying the underlying technical framework. It covers applications for freight services, and management of freight connections with other modes of transport.

The geographic scope of the TSIs is on:

- The trans-European conventional rail system network;
- The trans-European high-speed rail system network;
- Other parts of the network of the rail system in the European Union.

The TAF TSI covers (European Commision, 2014):

- Applications for freight services, including information systems, such as real-time monitoring of freight and trains;
- Marshalling and allocation systems, whereby train composition is meant;

- Reservation systems, whereby train path reservation is meant;
- Management of connections with other modes of transport and production of electronic accompanying documents.

More specifiable the benefits of TAF TSI are:

- **Simplifying the train/cargo handover processes.** Simplified, seamless data exchange, i.e. simplified data handover, implies also an easier train or cargo handover. Through its unique representation and interpretation, the receiving side instantly recognizes the data, and therefore also the train or cargo may be accepted without unnecessary delay.
- **Opening competition among Information and Communications Technology (ICT) vendors** means also pushing down the prices of ICT. As the data exchange is not driven by proprietary, closed standards anymore (which are often hard to get documented from the original vendor), the new ICT systems and components may be purchased from an arbitrary vendor, only on the condition of the standards implementation.
- **Shortening the handover and/or dwell times** of the trains means less idle times and lower losses on dwell for the RU or operator (more efficient usage of train paths, vehicles, staff etc., and therefore increasing capacity and reducing costs).
- That way a **better utilisation of the vehicles and other resources** is achieved, which leads again to increased capacity and higher cost efficiency and pushing the prices for the end customer down.

For the TEN-T Corridors, the focus is on transparent and accessible information provision around path allocation and traffic management, thus achieving higher flexibility and increased capacity. This is realised by XML-messages that covers data concerning the following functions (Table 3-5):

Table 3-5 Categories of TAF TSI functions (ERA Telematics Team, 2013)

RU only functions	Joint RU and IM functions
Consignment Note Data	Common Interface
Wagon & Intermodal Operating Unit Data (WIMO)	Reference Files
Wagon Movement	Train Running Information and Train Delay Cause
Shipment ETA	Train Forecast
	Service Disruption
	Train Enquiries
	Train Preparation
	Infrastructure Restriction Notice
	Ad hoc Path Request
	Train Transport Identifier

TAF TSI will be further analysed focussing on the storyboard requirements in the following chapters:

TAF TSI		
Storyboard 1 **Infrastructure data for operation and simulation**	Storyboard 2 **Effective usage of cross mode capacity**	Storyboard 3 **Real-time data for cross-organisation operations**
		2020 → chapter 6.2 2030 → chapter 6.3 2050 → chapter 6.4

3.10 ON-TIME RTTP - REAL-TIME TRAFFIC PLAN

The ON-TIME RTTP may support storyboard 3 providing operational train and infrastructure related data structures as shown in Figure 3-26 relying on railML base data.

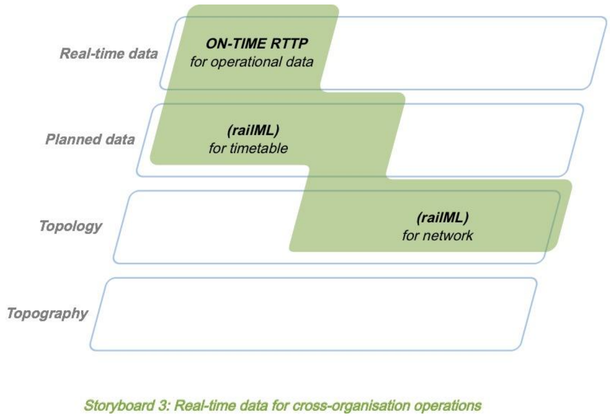

Figure 3-26 ON-TIME RTTP coverage for data classes and storyboards

The ON-TIME Real-Time Traffic Plan (RTTP) contains a microscopic conflict-free train path plan optimized for a short-term horizon. It was developed within the WP4 "Methods for real-time traffic management" of the ON-TIME project (ON-TIME_WP04, 2014).

As shown in Table 3-6 the RTTP comprises two perspectives containing sequences of events, the infrastructure and the timetable perspective. Both views are represented as event lists evolving along the time, i.e. the events associated with a given train will be executed one after the other, and different trains will follow each other over a certain track section. (ON-TIME_WP04, 2014).

Table 3-6 RTTP Functions (ON-TIME_WP04, 2014)

Perspective	Function	Description
Timetable	Rescheduled timetable for trains	For each train an ordered list of sections to drive on: The routing / timing part describes the sequence of track vacancy detection sections a train passes along its run together with its planned occupation and release times. The stopping part of the RTTP describes, on which of these sections, where exactly (along the section) and when the train will stop.
Infrastructure	Route setting request to signalling system	For each section an ordered list of trains to occupy it: It can be generated from the train view of the RTTP, i.e. the sequence of trains which will pass a certain track section and when this is expected to happen. The Traffic Control System can determine the sequence of routes setting.

The ON-TIME RTTP provides information about predicted train sequences on a selected network part. Train routes are provided as well as expected block occupancies. The ON-TIME RTTP data resulted from processing dynamic and static traffic data within the Perturbation Management Module (PMM), as shown in Figure 3-27.

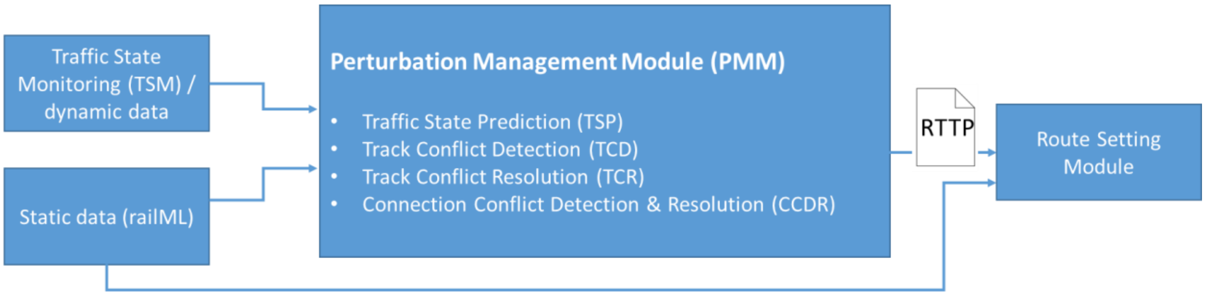

Figure 3-27 Perturbation Management Module and resulting Real-Time Traffic Plan (RTTP)

The purpose of the ON-TIME RTTP was to issue route-setting commands just in time. That means as early as necessary, such that a train should not see a restrictive signal aspect as a result of late route setting, while also not setting the route so early that it reduces flexibility within the network. A traffic management system may optimize between both contrary requirements and thus facilitates increased capacity.

The SBB currently utilises a system, called ADL (Adaptive Zuglenkung), which delivers speed recommendations to the locomotive driver for energy-efficient, smooth and comfortable driving. (SBB, 2014) SBB developed a proprietary data format, respective protocol specification, for sending the lowered speed aspect to the train display (Human-Machine-Interface HMI). For giving that information, the background system has also to provide Traffic State Prediction (TSP) and Track Conflict Detection (TCD). SBB currently implements lowering speed recommendations as one method of Track Conflict Resolution (TCR).

Further TCR and Connection Conflict Detection & Resolution (CCDR) methods would incorporate re-scheduling including estimating impacts on neighbouring trains.


```
      <tDSectionOccupation trainSequenceID="28" occupationStart="2014-05-04 07:22:48.461 CET" trainID="TB51" tDSectionID="TDS9_2"/>
      <tDSectionOccupation trainSequenceID="29" occupationStart="2014-05-04 07:23:20.769 CET" trainID="TB51" tDSectionID="TDS9_1"/>
      <tDSectionOccupation trainSequenceID="30" occupationStart="2014-05-04 07:23:53.76 CET" trainID="TB51" tDSectionID="TDS4567"/>
      <tDSectionOccupation ocpRef="S1P1_1" trainSequenceID="31" occupationStart="2014-05-04 07:24:25.384 CET" trainID="TB51" tDSectionID="TDS0_3"/>
      <tDSectionOccupation trainSequenceID="32" trainID="TB51" tDSectionID="TDS0_2"/>
      <tDSectionOccupation trainSequenceID="33" trainID="TB51" tDSectionID="TDS0_1"/>
    </rTTPForSingleTrain>
  </rTTPTrainView>
- <rTTPInfrastructureView>
  + <rTTPForSingleTDSection tDSectionId="TDS10_20">
  - <rTTPForSingleTDSection tDSectionId="TDS17_2">
      <tDSectionOccupation trainSequenceID="13" occupationStart="2014-05-04 06:30:08.571 CET" trainID="TA51" tDSectionID="TDS17_2"/>
      <tDSectionOccupation trainSequenceID="13" occupationStart="2014-05-04 07:14:08.571 CET" trainID="TB51" tDSectionID="TDS17_2"/>
      <tDSectionOccupation trainSequenceID="13" occupationStart="2014-05-04 07:58:08.571 CET" trainID="TC51" tDSectionID="TDS17_2"/>
      <tDSectionOccupation trainSequenceID="13" occupationStart="2014-05-04 08:42:08.571 CET" trainID="TD51" tDSectionID="TDS17_2"/>
      <tDSectionOccupation trainSequenceID="13" occupationStart="2014-05-04 09:26:08.571 CET" trainID="TE51" tDSectionID="TDS17_2"/>
    </rTTPForSingleTDSection>
  - <rTTPForSingleTDSection tDSectionId="TDS17_3">
      <tDSectionOccupation trainSequenceID="12" occupationStart="2014-05-04 06:29:37.142 CET" trainID="TA51" tDSectionID="TDS17_3"/>
      <tDSectionOccupation trainSequenceID="12" occupationStart="2014-05-04 07:13:37.142 CET" trainID="TB51" tDSectionID="TDS17_3"/>
      <tDSectionOccupation trainSequenceID="12" occupationStart="2014-05-04 07:57:37.142 CET" trainID="TC51" tDSectionID="TDS17_3"/>
```


Figure 3-28 RTTP – Excerpt XML file: Train view and infrastructure view

ETCS does not provide any facilities to recommend a lower speed aspect than the allowed one. The Movement authority (MA) in ETCS Level 2 specifies the maximum allowed speed for certain track distances. Enabling energy saving actions on the train, an ETCS text message may be used to inform the train driver about recommended speed aspects or upcoming closed signals.

In much the same way as SIRI, the ON-TIME RTTP is linked to another data format that provides the basis for real-time data: parts of railML 2.2 subschemas 'infrastructure', 'rollingstock', 'timetable' and draft 'interlocking' models are all involved. Figure 3-28 presents an excerpt of an RTTP XML file with train and infrastructure view.

An analysis of the ON-TIME RTTP presented in the context of the storyboard requirements will appear in chapters:

ON-TIME RTTP		
Storyboard 1 **Infrastructure data for operation and simulation**	Storyboard 2 **Effective usage of cross mode capacity**	Storyboard 3 **Real-time data for cross-organisation operations**
		2020 → chapter 6.2 2030 → chapter 6.3 2050 → chapter 6.4

3.11 OTHER DATA FORMATS AND OPEN DATA CONCEPTS

Although the XSD schema based XML standards are already widely spread in the public transport domain, there are alternative technologies, data formats and concepts, which fit the requirements of the storyboards:

- Non-XML data formats within the railway sector, e.g. TAP TSI;
- Data formats for other transport sectors, e.g. GTFS originally developed for urban public transport services;
- Generic key-value-pairs as data concept, e.g. OpenStreetMap data.

The following sections consider the practicalities of the underlying data formats alongside the ability to supply data. Added value is expressed according to Tim Berners-Lee's 5 star rating system for Linked Open Data (LOD) (Berners-Lee, 2010):

★ Available on the web (whatever format) with an open licence;

★★ Available as machine-readable structured data (e.g. excel instead of image scan of a table);

★★★ As (2) plus non-proprietary format (e.g. CSV instead of excel);

★★★★ Entire above, plus: Use open standards from W3C (RDF and SPARQL) to identify things, so that people can point at your stuff;

★★★★★ Entire above, plus: Link your data to other people's data to provide context.

The 5-star rating scheme is nicely illustrated and backed by examples, costs and benefits for each step on the 5 ★ Open Data website (http://5stardata.info/en/).

3.12 TAP TSI DATA FORMATS

TAP TSI may support storyboard 2 providing fare and ticket data structures as shown in Figure 3-29.

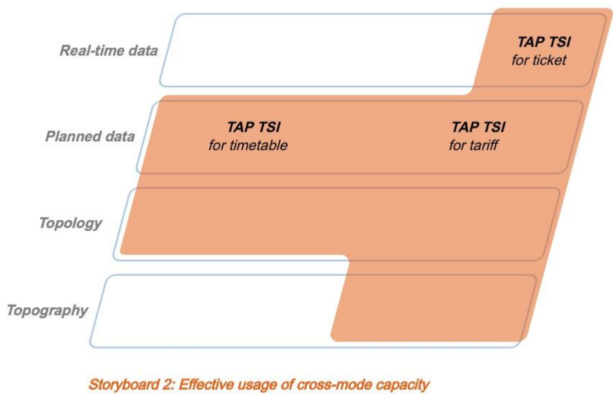

Figure 3-29 TAP TSI coverage for data classes and storyboards

The Technical Specification for Interoperability on “Telematics Applications for Passengers” (TAP TSI) defines protocols for the data exchange, which must be respected by the European rail sector (RUs, IMs, ticket vendors etc.) according to the European Rail Passengers’ Rights Regulations EC/1371/2007 and to the Interoperability Directive EC/2008/57.

TAP TSI deals with any information that is directly linked to a passenger and his (international) journey. Each technical document of this TSI in version 1.2.0 provides own data formats; a uniform approach was obviously not intended (Table 3-7). Contrary to that, TAF TSI defines one XML data schema.

Table 3-7 Data Formats used in TAP TSI (ERA Telematics Team, 2014)

Technical document	References	Data format
Tariff data	B.1; B.2; B.3	Fixed length text files
Timetabling	B.4	EDIFACT
Reservation messages (RTC2 standard)	B.5; B.6	Binary messages
Home printed tickets, PRM reservation	B.7	XML messages

Fixed length text files for tariff data

The “Computer generation and exchange of tariff data meant for international and foreign sales” is described within TAP TSI technical documents B.1 for non-reservation tickets (NRT) and B.2 for integrated reservation tickets (IRT). Following this specification, RUs may provide tariff data to commercial entities in a uniform manner, through subject specific text files complying with fixed numbers of characters (Table 3-8).

Table 3-8 TAP TSI file structure (ERA, 2015)

TAP TSI B.2	File	File name	Number of characters
Annex 1	Tariffs	PCTA-xxxx-xxx	304
Annex 2	Range	PCGA-xxxx-xxx	169
Annex 3	Cards / Memo	PCCA-xxxx-xxx	17
Annex 4	Exclusions	PCEX-xxxx-xxx	47
Annex 5	Sales conditions	PCCV-xxxx-xxx	21
Annex 6	After sales conditions	PCAV-xxxx-xxx	47
Annex 7	Prices	PCPR-xxxx-xxx	98
Annex 8	Zones	PCZO-xxxx-xxx	88
Annex 8b	Grouped OD	PCGO-xxxx-xxx	132
Annex 9	Name Cards, Memo	PCNC-xxxx-xxx	610
Annex 10	Distribution	PCDI-xxxx-xxx	169
Annex 11	Tariff Combinations, Dynamic Prices from/to	PCCD-xxxx-xxx	14
Annex 13	Header	PCET-xxxx-xxx	17 /15

Figure 3-30 illustrates the structure of these text files, namely the OD record: Origin/Destination.

Serial N° in record	Field	Characters	N° of charact.	M or O	Reference	Pos. Of char.	Version	Comments
1	Railway code	numerical	4	M	TAP TSI T. D. B.8	1-4	3	ex. 1080 = DB
2	Code entity	alpha N	3	M	*Code List B.2.1*	5-7	2	Unique Code per entity ex. TAL; ARN;
3	Code Origin / Destination Group	numerical	5	M		8-12	2	Unique Id by OD group and by RU/entity
4	Name OD Group	alpha N	32	M		13-44	2	
5	Origin Station Code	numerical	9	M	TAP TSI T. D. B.9	45-53	3	zero padded Country Code/Infrastructure holder Code (4) + location main code (5) example: 008765432
6	Name of station	alpha N	35	M		54-88	5	IT script without accents but with upper and lower case.
7	Destination Station Code	numerical	9	M	TAP TSI T. D. B.9	89-97	5	zero padded Country Code/Infrastructure holder Code (4) + location main code (5) example: 008765432
8	Name of station	alpha N	35	M		98-132	5	IT script without accents but with upper and lower case.

Figure 3-30 TAP TSI – file Structure of 'Grouped OD' (ERA, 2015)

A centralized and unique locations database, called Common Repository Domain (CRD) is being created in the form specified by TAF Regulation EC/62/2006. For linking these location codes with tariff data, the following subsidiary codes are needed (Dell'Arciprete, 2012):

- Tariff border;
- Tariff zone;
- Station name.

EDIFACT FOR TIMETABLING

Within TAP TSI B.4 UN/EDIFACT – in the full name of United Nations Electronic Data Interchange For Administration, Commerce and Transport – is used for the exchange of timetabling data. EDIFACT provides standard messages and associated service interfaces. Codes may be referenced from either universal lists, or from pre-defined extensions.

The conversion of some EDIFACT aspects to XML formats, like NeTEx and railML, is worth for further investigation (Figure 3-31, Figure 3-32).

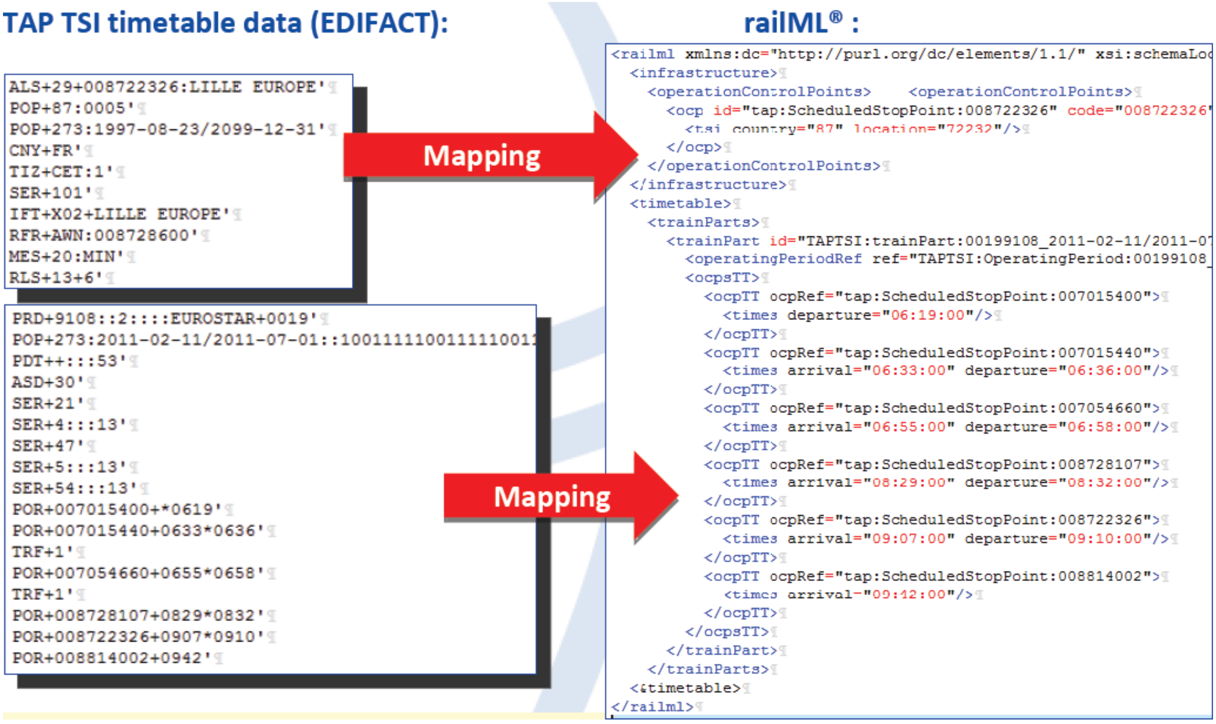

Figure 3-31 TAP TSI timetable – Mapping EDIFACT and railML (ERA Telematics Team, 2013)

BINARY MESSAGES FOR RESERVATION - "RCT2 COMPRESSED"

The Rail combined Standard 2 (RCT2), applicable for all travel documents in international and foreign sales, must be produced and issued electronically, according to the Commission Regulation (EU) No 454/2011. RCT2 – defined within TAP TSI B.6 "Electronic seat/berth reservation and electronic production of transport documents – Transport documents (RCT2 Standard)" – enhances TAP TSI B.5 "Electronic reservation of seats/berths and electronic production of travel documents – exchange of messages".

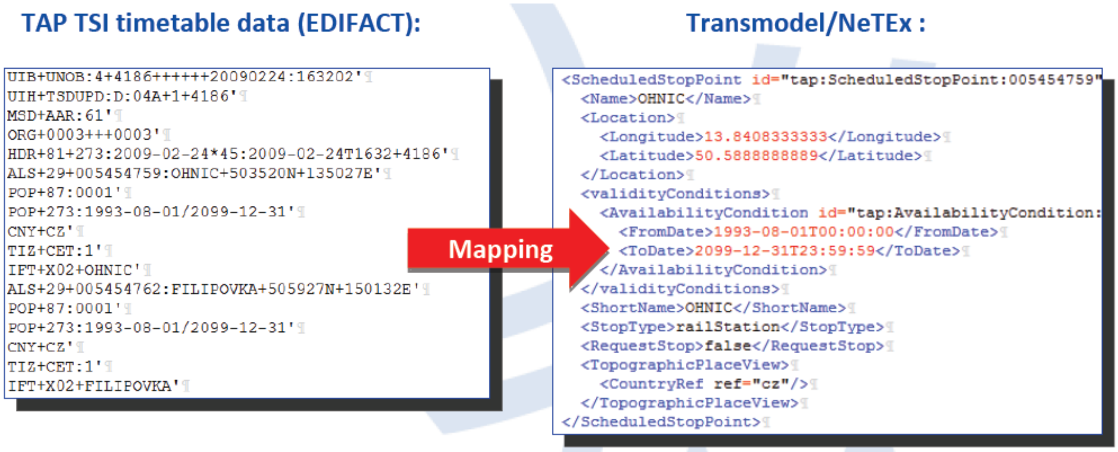

Figure 3-32 TAP TSI timetable – Mapping EDIFACT and NeTEx (ERA Telematics Team, 2013)

The RCT2 standard is designed for issuing combined "Ticket + Reservation" travel documents and thereby enables multi-service products, integrating transport tickets, reservations and any ancillary services (e.g. provision of support for disabled access). Besides this classic RCT2 Standard, the "RCT2 compressed" spreads out, coding the main information elements of the electronically issued ticket into 2D barcode(s). They are preferred for check-in and after sales activities, also allowing additional encoded elements for security and verification process.

XML MESSAGES FOR HOME PRINTED TICKETS

As a consequence of organizing journeys via Internet, there has been an increase in the number of customers choosing to print their tickets at home. The layout of "print at home" tickets is described in TAP TSI B.7 "International rail ticket for home printing". In order to unify layouts and respect bilateral standards, the RCT2 layout of tickets is adapted.

TAP TSI has been designed to provide data sets to a certain interest group, meaning no Open Data. It will be further analysed focussing on the storyboard requirements in the following chapters:

TAP TSI		
Storyboard 1 **Infrastructure data for operation and simulation**	Storyboard 2 **Effective usage of cross mode capacity**	Storyboard 3 **Real-time data for cross-organisation operations**
	2020 → chapter 5.2	

3.13 GTFS - General Transit Feed Specification

GTFS may support storyboard 2 providing multimodal timetable data as shown in Figure 3-33.

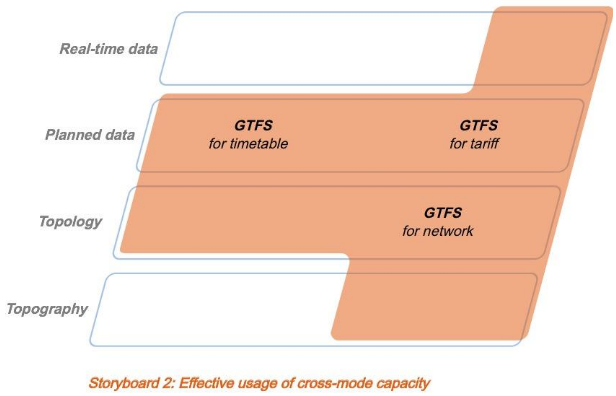

Figure 3-33 GTFS coverage for data classes and storyboards

The General Transit Feed Specification (GTFS) defines an open data format for public transport schedules and associated geographic information applicable to a range of different transport modes. GTFS data are exchanged in form of a collection of CSV files, also called GTFS feed (Figure 3-34).

Filename	Required	Defines
agency.txt	**Required**	One or more transit agencies that provide the data in this feed.
stops.txt	**Required**	Individual locations where vehicles pick up or drop off passengers.
routes.txt	**Required**	Transit routes. A route is a group of trips that are displayed to riders as a single service.
trips.txt	**Required**	Trips for each route. A trip is a sequence of two or more stops that occurs at specific time.
stop_times.txt	**Required**	Times that a vehicle arrives at and departs from individual stops for each trip.
calendar.txt	**Required**	Dates for service IDs using a weekly schedule. Specify when service starts and ends, as well as days of the week where service is available.
calendar_dates.txt	Optional	Exceptions for the service IDs defined in the calendar.txt file. If calendar_dates.txt includes ALL dates of service, this file may be specified instead of calendar.txt.
fare_attributes.txt	Optional	Fare information for a transit organization's routes.
fare_rules.txt	Optional	Rules for applying fare information for a transit organization's routes.
shapes.txt	Optional	Rules for drawing lines on a map to represent a transit organization's routes.
frequencies.txt	Optional	Headway (time between trips) for routes with variable frequency of service.
transfers.txt	Optional	Rules for making connections at transfer points between routes.
feed_info.txt	Optional	Additional information about the feed itself, including publisher, version, and expiration information.

Figure 3-34 GTFS feed files (Google Developers, 2015), CC-BY-3.0

Although the CSV formatting of GTFS feeds makes them harder to cross check than XML based formats, the simplicity of using the data, particularly for relatively simple transport networks, means that they are widely used. Thus, Knowles & Miller proposed a Transmodel based XML schema for exchanging transit stop and timetable data, which would be fully compatible and interoperable with both Google's GTFS and Transmodel based data sets. In particular as GTFS ought be enhanced to cover further more complex capabilities, Transmodel can be used to guide and validate the design, drawing on many years of industry experience. (Knowles & Miller, Transmodel for google, 2008)

Trip planning tools like Google Transit (http://maps.google.com/transit) or OpenTripPlanner (http://www.opentripplanner.org/) rely on up-to-date GTFS data integrating transit stops, routes, schedules and fare information into their routing algorithms. Although the GTFS feeds were originally designed only for the use in Google Transit, the specification is widely used as import format in other applications. (AWT Consortium, 2014)

Providing GTFS as Linked Open Data (LOD)

GTFS feeds are increasingly provided to the public as ★★★ LOD (Google TransitDataFeed - PublicFeeds, 2015) (GTFS Data Exchange, 2015):

- By railway agencies, e.g. LEO Express/Czech, SNCB/Belgium;
- By other transport agencies, e.g. ACTV/Italy, AirCoach/Ireland, FlixBus/Germany;
- By transport authorities for their responsible region, e.g. Berlin-Brandenburg, Budapest, Helsinki, Madrid, Oslo, Paris, Roma, Vilnius;
- By legally mandated authorities for rail transport, e.g. ATOC and DfT for Great Britain National Rail train services;
- By legally mandated authorities regarding entire public transport, e.g. in Estonia, in the Netherlands, in Sweden.

  GTFS is not the only format of this type, the same data are often provided as ★★★ LOD in other data formats (Google TransitDataFeed - PublicFeedsNonGTFS, 2015):

- By railway agencies, e.g. NS/Netherlands;
- By transport authorities for their responsible region, e.g. TfL/UK;
- By legally mandated authorities regarding entire public transport, e.g. in Switzerland.

In contrast to the LOD movement, some railway agencies provide GTFS data with a proprietary license to dedicated partners, e.g. Deutsche Bahn to Google Inc. (Meyer, 2012)

GTFS will be further analysed focussing on the storyboard requirements in the following chapters:

GTFS		
Storyboard 1 **Infrastructure data for operation and simulation**	Storyboard 2 **Effective usage of cross mode capacity**	Storyboard 3 **Real-time data for cross-organisation operations**
	2020 → chapter 5.2 2030 → chapter 0 2050 → chapter 5.4	

3.14 GTFS-REALTIME

GTFS-realtime may support storyboard 3 providing multimodal real-time data as shown in Figure 3-35 relying on GTFS base data.

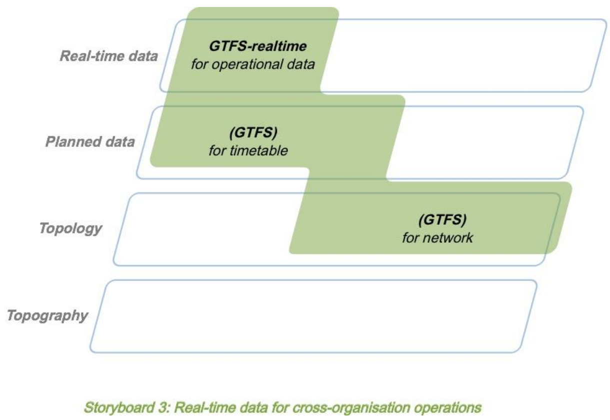

Figure 3-35 GTFS-realtime coverage for data classes and storyboards

GTFS-realtime (rarely GTFS-RT) is a feed specification that allows public transportation agencies to provide real-time updates about their fleet to application developers. It is an extension to GTFS and was introduced and released in 2011. The specification currently supports the following types of information (Google Developers, 2015):

- Trip updates – delays, cancellations, changed routes;
- Service alerts – stop moved, unforeseen events affecting a station, route or the entire network;
- Vehicle positions – information about the vehicles including location and congestion level.

Unlike GTFS, the GTFS-realtime feeds are provided as binary files, where updates of each type are provided in a separate feed. GTFS-realtime needs to be combined with GTFS schedule data to be meaningful. This combination enables to deliver data about all routes and vehicles at once using a minimum of bits. (MBTA, 2015)

In Europe only OVapi provides Open Access to GTFS-realtime data [for buses from different transit agencies in the Netherlands as ★★★ LOD] so far, whereas it is already provided by several US transit agencies, such as BART (Oakland, CA), TriMet (Oregon, WA), or MARTA BUS (Atlanta, GA). (Steiner, Hochmair, & Paulus, 2015)

GTFS-realtime will be further analysed focussing on the storyboard requirements:

GTFS-realtime		
Storyboard 1 **Infrastructure data for operation and simulation**	Storyboard 2 **Effective usage of cross mode capacity**	Storyboard 3 **Real-time data for cross-organisation operations**
		2020 → chapter 6.2

3.15 OpenStreetMap (OSM)/ OpenRailwayMap

OpenStreetMap may support storyboard 1 providing open infrastructure data where gaps exist in official provision as shown in Figure 3-36.

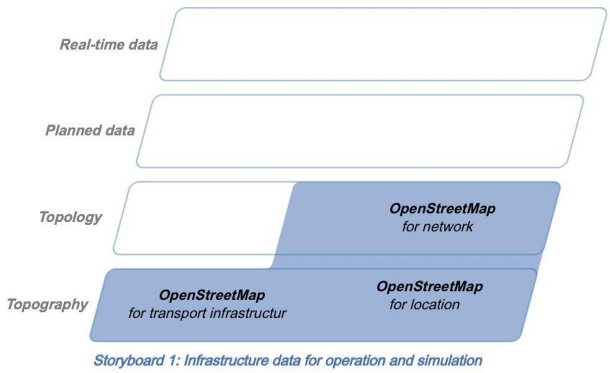

Figure 3-36 OpenStreetMap coverage for data classes and storyboards

The OpenStreetMap (OSM) project intends to create and distribute worldwide free geographic data without technical restrictions in order to encourage people to use them in creative, productive or unexpected ways. The project started in 2004 and publishes its data under the Open Data Commons Open Database License (ODbL) v1.0. OSM aims to profit from crowdsourcing. Lots of people around the world complete and update OSM data in order to generate or just use thoroughly created up-to-date maps. OSM data is the basis for many publically available maps in a number of domains, including sailing, cycling, hiking, motorcars, public transport, skiing, power networks.

OSM database objects are transferred in different data formats, depending on the consuming application, common variants include (Wiki OpenStreetMap.org, 2015):

- OSM XML – xml-format provided by the API;
- PBF Format – highly compressed, optimized binary format similar to the API;
- o5m – for high-speed processing, uses PBF coding, has same structure as XML format;
- OSM JSON – json variant of OSM XML;
- Level0L – plain text file, compatible with OSM XML.

  Data structures used within OpenStreetMap are developed in an organic way according to the data collection ('mapping') needs without any central coordination. The generic key concept comprises:

- 'node' with latitude and longitude according to WGS84;
- 'way', also known as edge or link, as a poly-line between consecutively referred 'nodes';
- 'relation' as a collection of certain 'nodes', 'ways' and other 'relations';

- 'tag-value', also known as key-value-pair, enabling semantics for 'nodes', 'ways' and 'relations'.

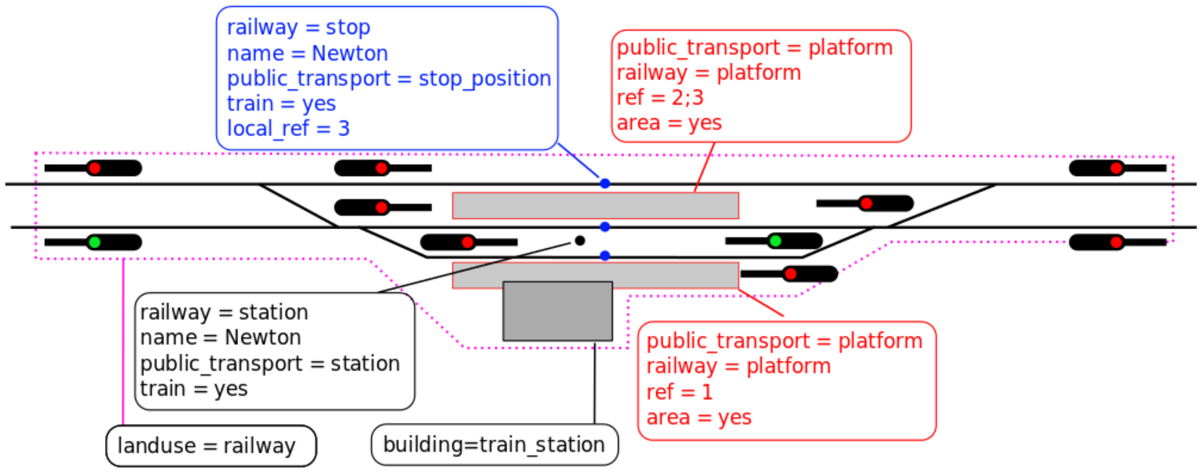

Figure 3-37 OSM Tag "railway=station" (Wiki OpenStreetMap.org Tag:railway=station, 2015)

As illustrated in Figure 3-37, a station may be modelled based on the vocabulary introduced above:

- Tracks are defined by connected 'ways';
- Signals are defined by 'nodes' nearby the tracks;
- Platforms are defined by enclosed 'ways', see red notices;
- A building is defined by an enclosed 'way', see black notice;
- Stop positions are defined by 'nodes' touching 'ways', see blue notice;
- A station is defined by a 'node' nearby the stop positions, see black notice;
- A station area is defined by a 'relation', see magenta notice;
- Any semantics and relationships are defined by 'tag-value'-pairs.

In order to create a map from these data, the rendering machine needs sophisticated algorithms for deducing the favoured symbols, points, lines or areas at correct positions at a certain map scale.

OpenRailwayMap

OpenRailwayMap (previously called 'Bahnkarte') is a detailed online map of the world's railway infrastructure, built using OSM data: http://openrailwaymap.org/. It provides three overlays (Figure 3-38), which may be applied to various base maps (Wiki OpenRailwayMap.org, 2014):

- Infrastructure, including milestones, switches, track numbers, signal boxes;
- Maximum track speeds;
- Signals.

The search features of OpenRailwayMap rely on the underlying OSM data and therefore result in data restricted to general railway concepts available at this level, such as stations, milestones and level crossings.

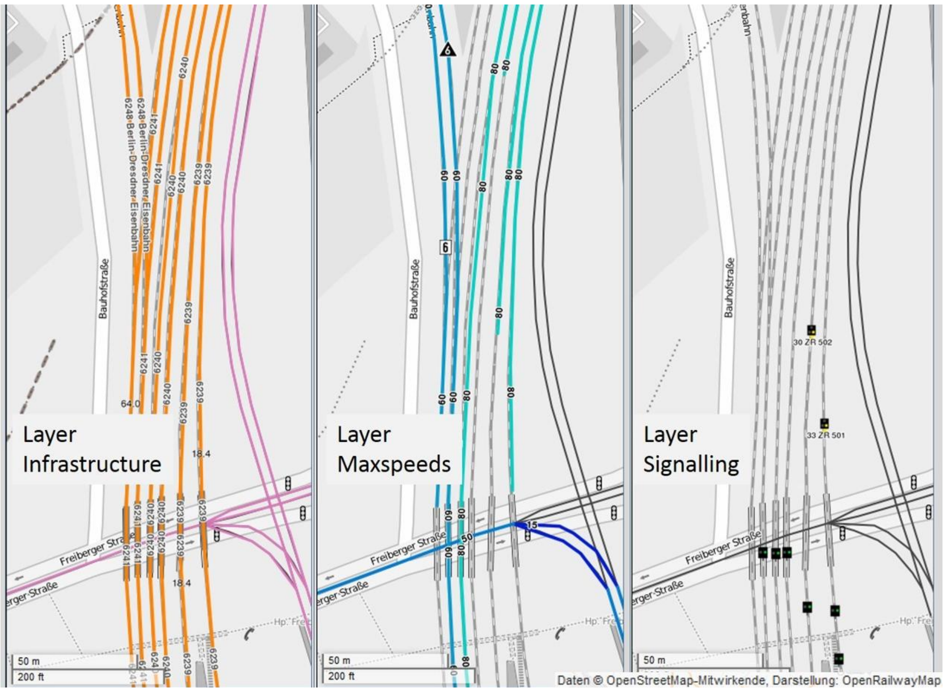

Figure 3-38 OSM / OpenRailwayMap Microscopic View, Different Layer

OpenRailwayMap highlights several use cases demonstrating the need for both raw railway-related mapping data, and pre-rendered maps (Wiki OpenRailwayMap.org, 2014):

- **Simulation:** Source for simulating trains and signal boxes close to reality;
- **Research:** Analyse the railroad network, simulate changes and use for educational purposes;
- **Base map for real-time traffic:** Track the position of trains, provide a real-time view of the trains' current position, show construction sites, detours, blockings and traffic density;
- **Routing:** Develop a routing application with detailed mapped railway infrastructure;
- **Public transport:** Data source for public transport applications like timetables, routing etc.

Besides the usage of OSM as base map in several railway software packages, it has been successfully launched for data generation in commercial tools (Wiki OpenStreetMap.org Eisenbahn, 2015):

- Importer for basic network data in simulation software VISUM (PTV Group, 2015);
- Importer into CAD layers for alignment in CARD/1 (Braun, 2012).

Railway-related OSM data has also been used for other, highly specific tasks, including:

- Generation of a routable railway network (Czioska, Thiemann, Giese, & Vogt, 2014);
- Generation of railway simulation landscape based on OSM data (Rahmig & Richter, 2014).

PROVIDING OSM AS LINKED OPEN DATA

Railway-related OSM data is available to the public in different quality at the OSM database at minimum as ★★★ LOD and may be used in any project honouring the terms and conditions. Combined with the OSM Semantic Network, which is encoded in SKOS vocabulary, OSM data may reach ★★★★★ LOD rating. (Wiki OpenStreetMap.org, 2015)

OSM will be further analysed focussing on the storyboard requirements in the following chapters:

OSM		
Storyboard 1 **Infrastructure Data for Operation and simulation**	Storyboard 2 **Effective usage of cross mode capacity**	Storyboard 3 **Real-time data for cross-organisation operations**
2020 → chapter 4.2 2030 → chapter 4.3 2050 → chapter 4.4		

3.16 EMERGING TECHNOLOGIES FOR SEAMLESS DATA EXCHANGE

Although XML allows the flexible transfer of information between applications, as a technology it is still reliant on human software engineers correctly interpreting the schemas they wish to use when building their software. Recent years have seen an increasing amount of interest in the Semantic Web, an extension of current Internet technologies in which machines will be able to explore online resources and infer new information autonomously. The Semantic Web will require that data is described online in a way that conveys its semantic (the meaning of the data) unambiguously to computers as well as humans, and this has led to the creation of a new, alternative family of data models, ontologies.

Ontologies are “content theories about the sorts of objects, properties of objects, and relations between objects that are possible in a specified domain of knowledge. They provide potential terms for describing our knowledge about the domain” (Chandrasekaran, Josephson, & Benjamins, 1999). Put simply, ontologies describe data by referencing published models of a domain; because of this, as long as the published models are both reasonably complete and share common root concepts, ontologies allow computers to automatically relate pieces of information drawn from different sources, understand the relationships between them, and infer new facts about the world.

Recent models have attracted a certain amount of interest in the transport domain. This is particularly relevant to multimodal journey planning, where mashup applications on the web must bring together similar data (timetables, ticket prices, live departure information) from a number of providers, who describe it using different sets of terminology (normally based on the terms used in their own transport modes).

The majority of the interest in semantic web technologies to date from within the industry has been focused on Remote Condition Monitoring (RCM). This is addressed by SP4 of CAPACITY4RAIL, but is not the primary focus of this document.

Ontologies generally provide knowledge representations with decent vocabularies, which may be used for enhancing open data sets to Linked Data according ★★★★ LOD through introducing URIs, that may be referred from other data sets. On top of thus enriched data sets, links into foreign data sets may be integrated, which lifts the open data set to ★★★★★ LOD, enabling Semantic Web technologies.

The benefits of Linked Data may be also capitalized in internal processes without providing the data to the public. Whereas the Open Data movement enables crowd sourcing and software development for free.

3.17 Ontology-based data verification and debugging

Semantic Railway Infrastructure (RI*) may support storyboard 1 enabling verification and debugging as shown in Figure 3-39.

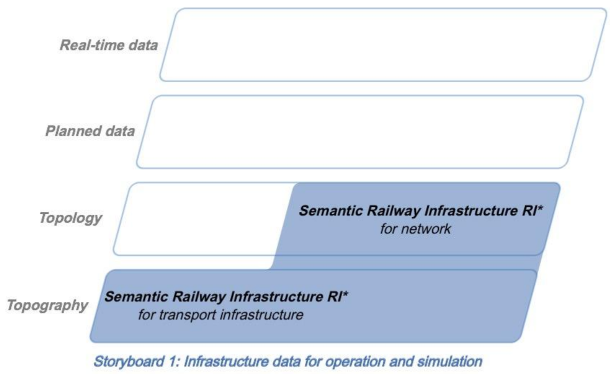

Figure 3-39 Semantic Railway infrastructure RI coverage for data classes and storyboards*

In (Lodemann, Luttenberger, & Schulz, 2013) a railway infrastructure ontology set RI* (Figure 3-40) is developed, which aims at command and control verification and debugging of railway infrastructure according to certain planning instructions.

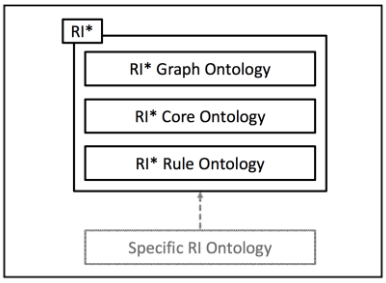

Figure 3-40 Railway infrastructure ontology set RI (Lodemann, Luttenberger, & Schulz, 2013)*

Table 3-9 gives an overview about the purpose of each ontology within the hierarchical set. The 'separation-of-concerns' approach, which was used as guide for the ontology's architecture, reflects different kinds of input to be provided from different parties.

Table 3-9 RI ontologies (Lodemann, Luttenberger, & Schulz, 2013)*

Name	Description	Providing party
RI* Graph Ontology	Physical and logical railway networks through vertices, edges, and ports	(Authors of the paper)
RI* Core Ontology	Hierarchy of railway specific concepts: approx. 50 railway elements equipped with approx. 80 object and 80 data type properties	Senior railway engineers
RI* Rule Ontology	SWRL rules according to RI planning instructions for 'terminological correctness' and 'value correctness'	Senior railway engineers
RI* Individual Ontology	Specific RI planning data, which are to be verified against the above defined knowledge base	Actual railway engineer

The RI planning data originates from XML, which underlies the Closed World Assumption (CWA) whereas the OWL ontology environment underlies the Open World Assumption (OWA). In order to perform data verification it is necessary to keep the open world as close as possible. (Lodemann, Luttenberger, & Schulz, 2013)

The framework does not only enable railway infrastructure verification but also debugging with the help of special Semantic Web Rule Language (SWRL) rules, called 'Semantic

Constraint (SC)', and additional heuristic methods for increasing performance. The authors describe positive experiences, when testing the developed ontologies with real industry data.

Linked Infrastructure Data regarding verification and debugging will be further analysed focussing on the storyboard requirements in the following chapters:

Linked Infrastructure Data / RI*		
Storyboard 1 **Infrastructure data for operation and simulation**	Storyboard 2 **Effective usage of cross mode capacity**	Storyboard 3 **Real-time data for cross-organisation operations**
2030 → chapter 4.3		

3.18 RAILWAY CORE ONTOLOGY

Railway Core Ontology (RaCoOn) may support storyboard 1 linking different data models as shown in (Tutcher, Easton, & Roberts, 2015).

A number of research projects and industrial initiatives concerning knowledge management and data modelling for railway data have been undertaken over the last decade, aiming to allow better integration of data between systems. For the development of RaCoOn the following concepts have been taken into consideration:

- railML, see also section 3.2;
- UIC RailTopoModel, see also section 3.3;
- Railway Domain Ontology (RDO) as constructed in the frame of the EU FP6 InteGRail project;
- National Public Transport Access Node (NaPTAN), see also sections 3.20, 3.21 and 3.22;
- ArcGIS Esri model.

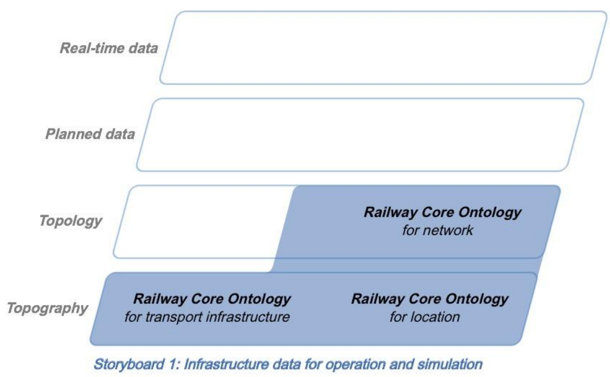

Figure 3-41 Railway Core Ontology coverage for data classes and storyboards

Although initially developed with the representation of signalling and rail infrastructure in mind, the model rapidly developed into a general model for the railways, including a “core” of generic railway concepts with extensions capturing particular subdomains (infrastructure, timetabling, rolling stock etc.) and an upper level model to define concepts used more broadly than rail (e.g. transport). The layered design philosophy behind the model is shown in fig 3-42.

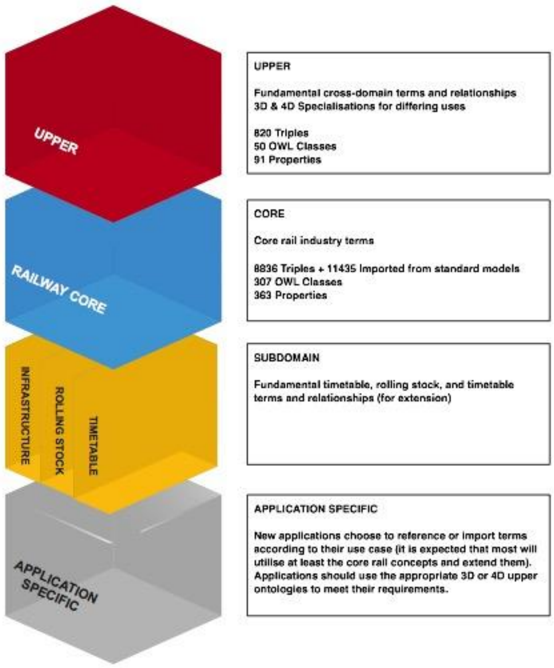

Figure 3-42 Layered design philosophy underpinning the RaCoOn model (Tutcher, Easton, & Roberts, 2015)

The RaCoOn upper level ontology contains knowledge of generic upper level concepts that transcend the railway domain. Such concepts include space and time, and are mostly reused from existing “gold standard” vocabularies, shown in Table 3-10.

Table 3-10 “Gold standard” vocabularies, included in RaCoOn (Tutcher, Easton, & Roberts, 2015)

Name	Description
W3C Time Ontology / Allan time relations	Representing instants, intervals; Entities are labelled with start and end times where required.
W3C Geo / Ordnance Survey Spatial Relations	Location positioning.

Name	Description
NASA Quantities, Units, Dimensions and Types ontology	Exhaustive list of quantities, units, dimensions and data types used in the upper ontology in conjunction with an appropriate design pattern to represent measurements and data types.
ISO 15926:2	Meta-model for entity types; classifies objects into independent and dependent, which is useful in defining acceptable ranges and domains for properties.

The rail core vocabulary ontology is a result of work carried out manually constructing and curating knowledge from other domain models and from UK industry experts. The vocabulary and its sub-modules predominantly draw upon corresponding elements in railML 2.2, relying on both its XML syntax and human-readable documentation in building an equivalent semantic data model.

The feasibility of RaCoOn was demonstrated at two use cases. The first demonstrator, which was presented at the 2014 IEEE Conference on Big Data (Tutcher, 2014), showed how the use of a linked data approach to the handling of asset information could add value as part of a scalable asset management platform. The second demonstrator, presented in the ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems (Tutcher, Easton, & Roberts, Enabling Data Integration in the Rail Industry Using RDF and OWL - the RaCoOn Ontology, 2015), aimed to show how the use of ontology and linked data can help the industry maximise on investment in existing information systems despite changes elsewhere in an increasingly technology-driven railway system.

In particular, the demonstrators set out to show how the use of ontology can provide a bridge between legacy systems and newer replacement services without sacrificing functionality, and how interfaces between such legacy systems and more contemporary linked data-based systems can be set up. As the volumes and variety of data gathered in new information systems on the railway continue to increase, the second demonstrator seeks to illustrate the practical uses of semantic data models in simplifying interfaces and applications, and enriching content.

Linked Data regarding the RaCoOn will be further analysed focussing on the storyboard requirements in the following chapters:

Linked Data / Railway Core Ontology		
Storyboard 1 **Infrastructure data for operation and simulation**	Storyboard 2 **Effective usage of cross mode capacity**	Storyboard 3 **Real-time data for cross-organisation operations**
2050 → chapter 4.4		

3.19 Enriched open GTFS data using Transit Ontology

Transit Ontology may support storyboard 2 enabling linked planned schedule data as shown in Figure 3-43.

In (Mishevska, Najdenov, Jovanovik, & Trajanov, 2014) GTFS data from the transport agency JSP Skopje are transformed into RDF and therefore enhanced to ★★★★ LOD. This process is based on:

- The Transit Ontology;
- W3C Geospatial Ontology;
- Own GTFS-ext ontology.

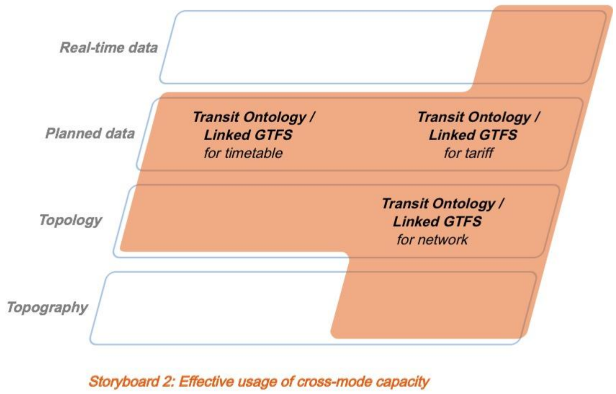

Figure 3-43 Transit Ontology coverage for data classes and storyboards

The Transit Ontology (Davis, 2012) is based on the General Transit Feed Specification and provides a vocabulary for describing transit systems, routes, stops and schedules. Figure 3-44 shows the core classes, which already reveals the basic GTFS concept (see also section 3.13).

As stated on the GitHub website of 'Transit Vocabulary', this schema has been superseded by 'Linked GTFS' (Colpaert & Byrd, 2015). Mishevska et al. demonstrated, that the transformation of GTFS CSV files into semantically enriched RDF data works quite well by presenting three use cases, which retrieve reasonable results by applying SPARQL queries.

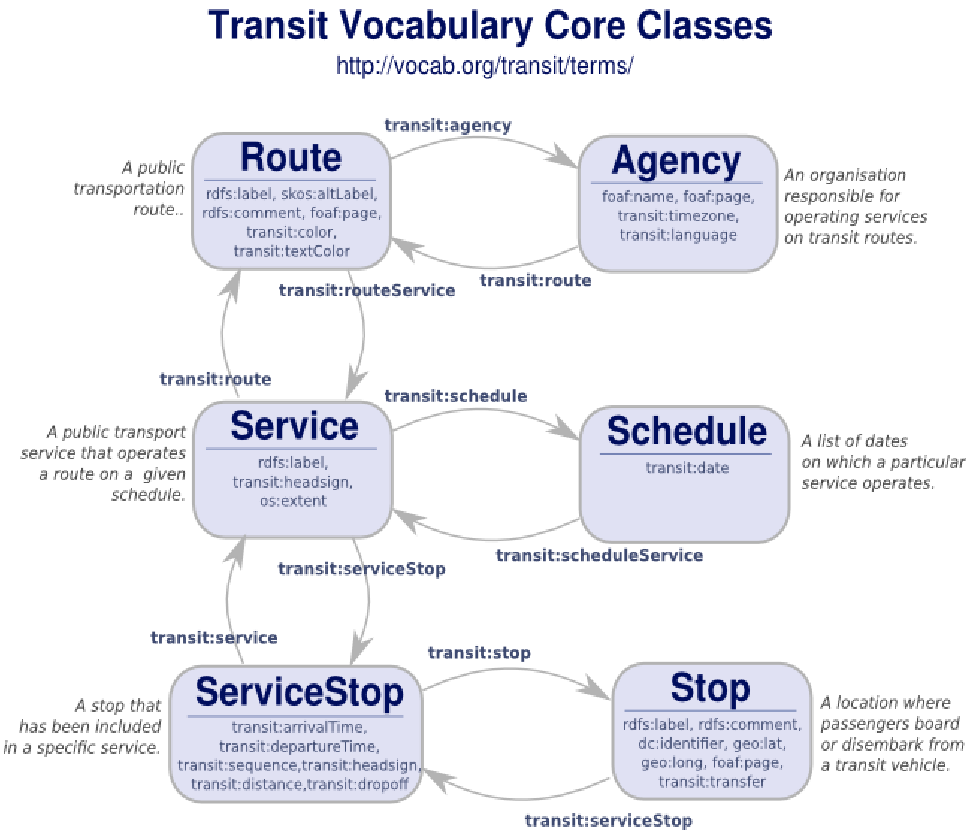

Figure 3-44 Transit Ontology core classes (Davis, 2012)

Linked Data regarding Transit Ontology or Linked GTFS will be further analysed focussing on the storyboard requirements in the following chapters:

Transit Ontology / Linked GTFS		
Storyboard 1 **Infrastructure data for operation and simulation**	Storyboard 2 **Effective usage of cross mode capacity**	Storyboard 3 **Real-time data for cross-organisation operations**
	2050 → chapter 5.4	

3.20 LOD BASED ON NEPTUNE

NEPTUNE Ontology may support storyboard 2 enabling interlinked planned schedule data as shown in Figure 3-45.

As (Plu & Scharffe, 2012) elaborated, publishing and interlinking transport data on the Web is feasible, targeting at ★★★★★ LOD, and paves the way for applications from other domains incorporating public transport services. They developed a new NEPTUNE Ontology (Plu, 2012) based on the already available ‘Norme d’Échange Profil Transport collectif utilisant la Normalisation Européenne’ (NEPTUNE) conceptual model and XML schema.

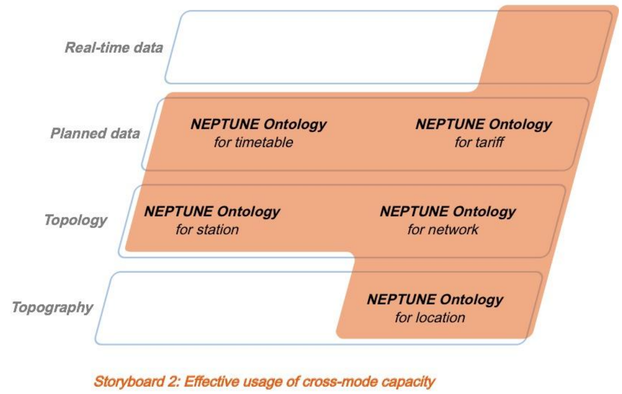

Figure 3-45 NEPTUNE Ontology coverage for data classes and storyboards

NEPTUNE, the French Standard (NFP-99506) specifying the reference format for data exchange of theoretical transport offers, particularly useful for the development of multimodal information systems, will evolve in the context of works within the European standardization of the NeTEx project and is based on a former version of Transmodel. NEPTUNE data are already offered as ★★★ LOD to the public.

Two potentially reusable ontologies, UK-centric ‘National Public Transport Access Node’ (NaPTAN) and Transit, could not be used. The NEPTUNE vocabulary (Plu, 2012) seemed to diverge somewhat from existing ontologies despite a few similarities.

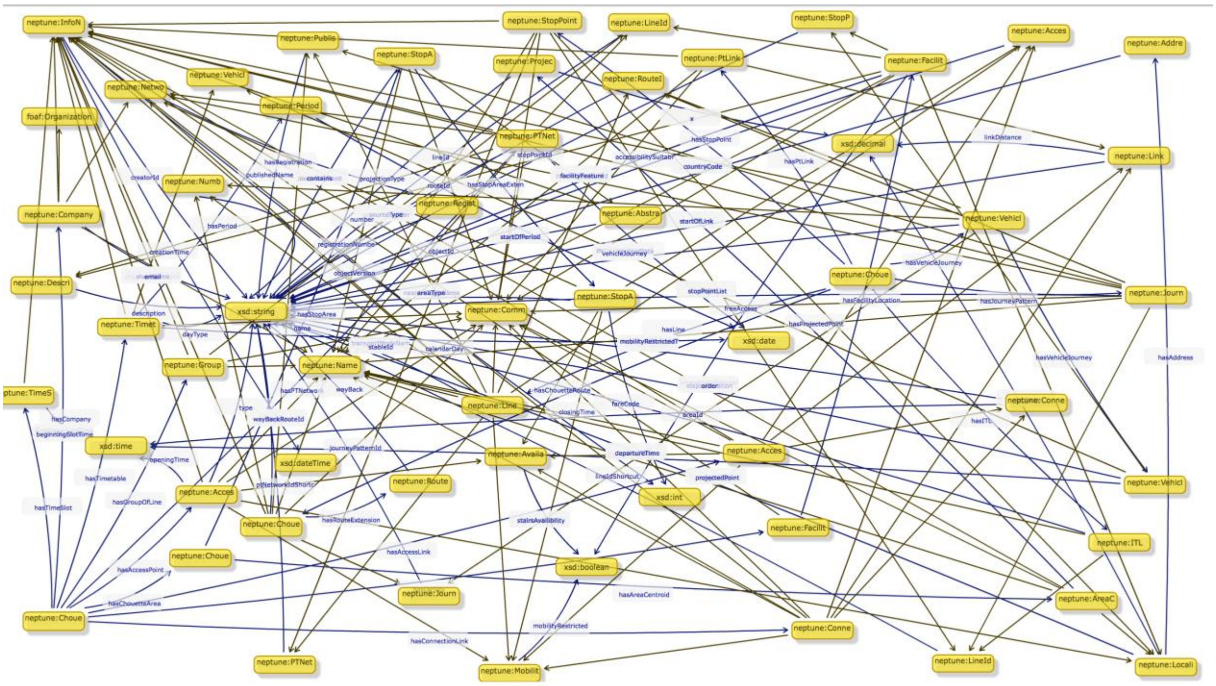

Figure 3-46 NEPTUNE Ontology (Plu, 2012)

The authors state, that the European standard, NeTEx should unify the vocabulary and formats for data exchange in Europe for public transport data. See also section 3.7 for more information about NeTEx. NEPTUNE concepts are now incorporated in NeTEx, as their authors were involved in the NeTEx development.

Thus published and interlinked data enable development of applications using multiple datasets simultaneously deploying generic tools and well-formulated queries.

Linked Data regarding the NEPTUNE ontology will be further analysed focussing on the storyboard requirements in the following chapters:

Linked Data / NEPTUNE Ontology		
Storyboard 1 **Infrastructure data for operation and simulation**	Storyboard 2 **Effective usage of cross mode capacity**	Storyboard 3 **Real-time data for cross-organisation operations**
	2050 → chapter 5.4	

3.21 LOD FOR ADVANCED PASSENGER INFORMATION SYSTEMS

Ontology for advanced passenger information system may support storyboard 2 enabling linked planned schedule data with other sources from the Web of Data as shown in Figure 3-48.

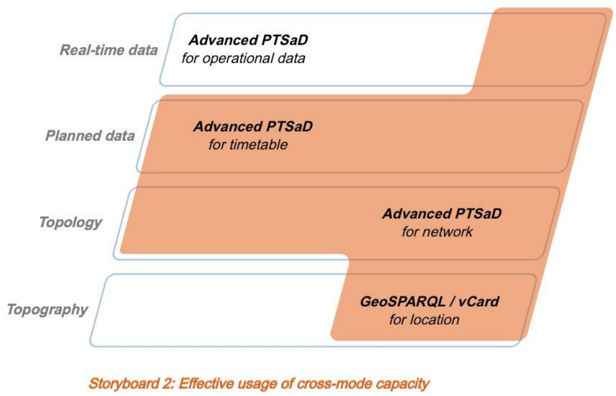

Figure 3-47 Ontology for advanced public transport services coverage for data classes and storyboards

Compared to (Plu & Scharffe, 2012) the researchers (Keller, Brunk, & Schlegel, 2014) show, how to enrich Public Transport data to ★★★★★ LOD and demonstrate the benefits with a tourism-themed prototype in form of a smart mobile application.

Therefore they developed their own ontology for Public Transport Services and Data (PTSaD) partially borrowing from existing models, like Transmodel, Transport Protocol Expert Group

(TPEG), NaPTAN, VDV 452, IFOPT, SIRI as well as NeTEx for terminology and common concepts.

The geographical entities are classified with the help of already existing ontologies GeoSPARQL and vCard Ontology as shown in Figure 3-48. Means of transport comprise as public as individual transport (same figure), the ontology thus supports multimodal trip suggestions by design. On the other hand, 'situation' and 'event' classes are defined in order to handle delays or disruptions.

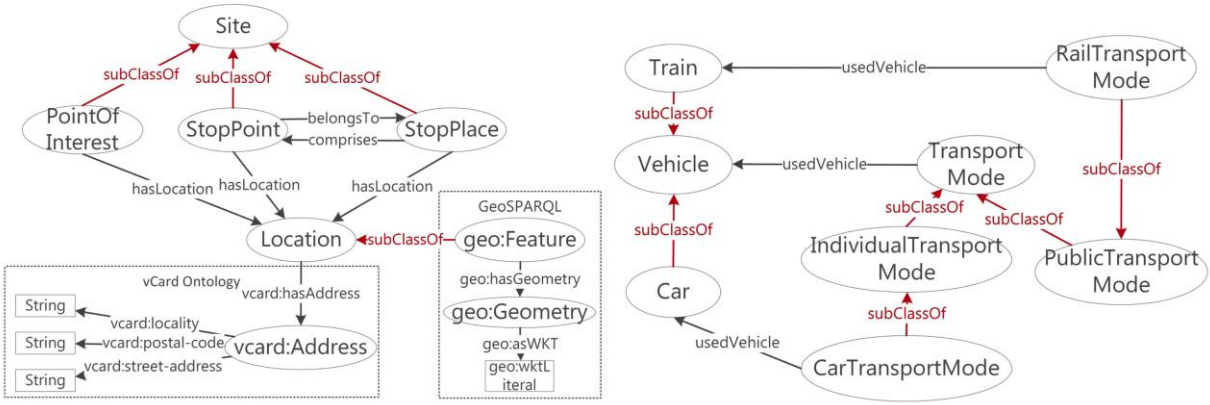

Figure 3-48 Public transport ontology (Excerpt) (Keller, Brunk, & Schlegel, 2014)

The concept for context-adaptation and filtering of data relies on the exploitation of semantic data, describing different facets of the user's situation and information needs: the public transport situation, weather conditions and points of interests but also user preferences. The application integrates existing passenger information and adds data sources from the Web of Data to exploit common links and structures. (Keller, Brunk, & Schlegel, 2014)

Linked Data regarding advanced passenger information systems will be further analysed focussing on the storyboard requirements in the following chapters:

Linked Data / Advanced passenger information ontology		
Storyboard 1 **Infrastructure data for operation and simulation**	Storyboard 2 **Effective usage of cross mode capacity**	Storyboard 3 **Real-time data for cross-organisation operations**
	2030 → chapter 0	

3.22 REAL-TIME DATA AQUIRED THROUGH SEMANTIC SENSOR DATA

Ontology for informed rural passenger may support storyboards 2 and 3 enabling data acquisition through users as shown in Figure 3-49.

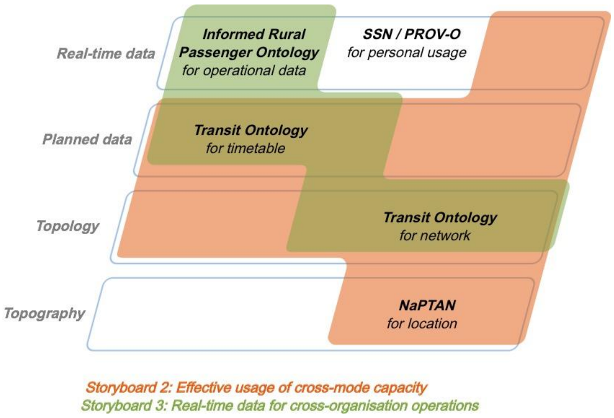

Figure 3-49 Ontology for real-time PT data acquisition coverage for data classes and storyboards

Cosar et al. describe in (Corsar, Edwards, Baillie, Markovic, Papangelis, & Nelson, 2013) the use of ontologies in GetThere, a real-time passenger information system (RTPI) for rural areas, to represent and integrate citizen sensors with data required to provide RTPI (e.g. timetable and route descriptions). Missing sensor data (e.g. vehicle locations from GPS) are collected by citizens, acting as data sensors. As recompense the users get a mobile PIS based on real-time data.

For achieving this complex service, a complex ontology framework is established, see Figure 3-50, i.e. comprising of:

- Linked NaPTAN (see also sections 3.20 and 3.21);
- Transit Ontology (see also section 3.19);
- FOAF (Friend of a Friend Ontology);
- SSN (W3C Semantic Sensor Network);
- PROV-O (W3C Provenance Ontology).

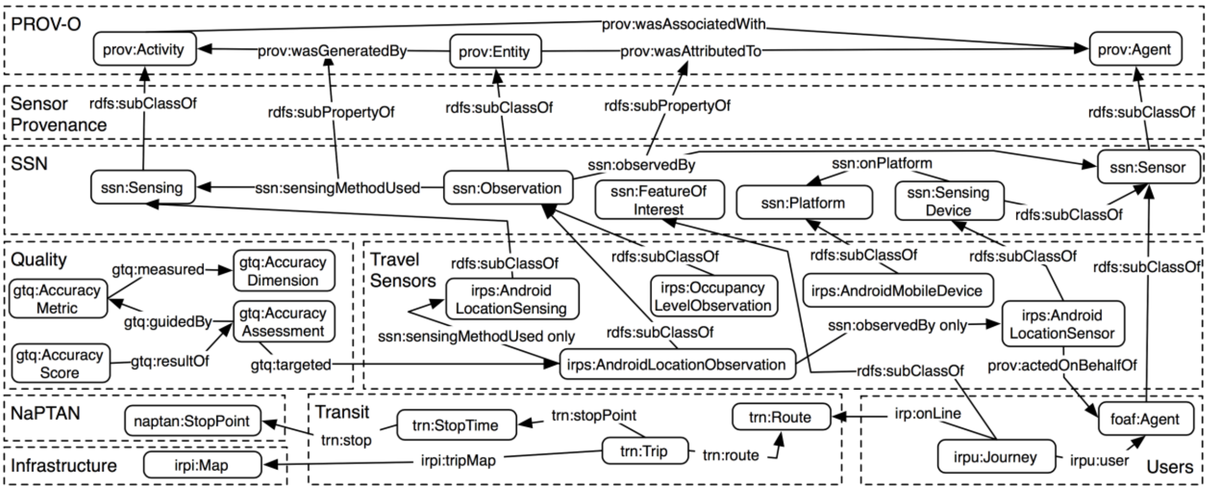

Figure 3-50 Ontologies integrating sensor data with other data to support RTPI provision (Corsar, Edwards, Baillie, Markovic, Papangelis, & Nelson, 2013)

This project aims at a robust service provision based on semantically enriched data sets, but not on publishing them, therefore no LOD rating is applicable. The paper states that the data could be published, than they would reach ★★★★★ LOD.

Linked Data regarding ontology based data acquisition will be further analysed focussing on the storyboard requirements in the following chapters:

Linked Data / Ontology based acquisition		
Storyboard 1 **Infrastructure data for operation and simulation**	Storyboard 2 **Effective usage of cross mode capacity**	Storyboard 3 **Real-time data for cross-organisation operations**
	2050 → chapter 5.4	2050 → chapter 6.4

4. STORY 1 CONSISTENT CROSS INDUSTRY INFRASTRUCTURE DATA

This chapter refers to the first storyboard presented in section 2.1 ”Consistent cross industry infrastructure data in support of planning, simulation and operations”. Promising data exchange formats, data resources and data concepts are considered and evaluated in order to determine their suitability / compatibility with the data related requirements of transport planning, simulation and operations in both today’s railway and the railway of 2050:

- railML / UIC RailTopoModel, see sections 3.2 and 3.3;
- RINF, see section 0;
- INSPIRE, see section 3.5;
- IDMVU, see section 3.6;
- OpenStreetMap / OpenRailwayMap, see section 3.15;
- RaCoOn, see section 3.18.

4.1 DATA FOR RAILWAY INFRASTRUCTURE DATA

The data highlighted in Figure 4-1 is needed to support cross industry applications for planning, simulation and operations across borders. A key element of these tasks is the need for detailed topographic and topological data.

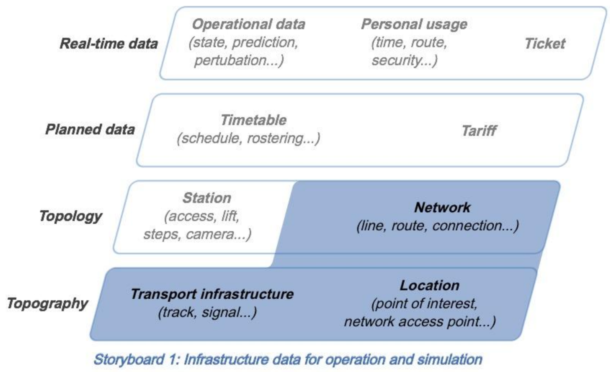

Figure 4-1 Data classes for storyboard 1

Current railML and IDM[VU] are the foremost data formats that are publicly documented and therefore ready for the implementation within railway business processes. They are solely developed for the data exchange between different software applications and different business partners.

UIC RailTopoModel is still work in progress showing some first successful internal implementations.

RINF and INSPIRE are initiatives mainly for publishing data to special interest groups. Therefore they also developed publicly documented data formats that are ready for the data exchange to the central repositories.

OSM is mainly a data collection initiative offering some maps for visualization of the database contents. The data format is subject to permanent development without any releases and with a more or less complete documentation.

Figure 4-2 sorts the proposed subjects out into pure data formats, and more or less Open Data initiatives in order to settle a clear understanding of the feasible range of use.

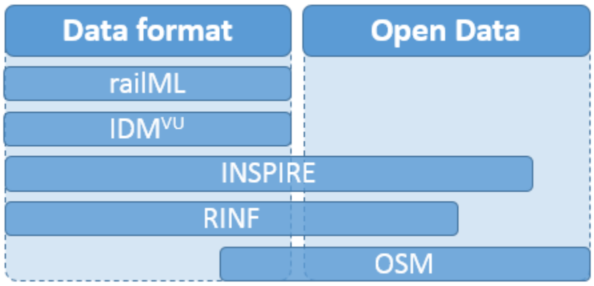

Figure 4-2 Data formats and Open Data for railway infrastructure

Generally, linking data sets in the data models given above would enable consistency validation. Furthermore cross-mode recovery strategies could be developed in case of serious or extended disruptions.

4.2 OVERVIEW 2020 – CONSOLIDATED DATA RESOURCES

Multiple data resources for the same infrastructure contain potentially conflicting views of the same physical assets. Data is presented using a range of models, limiting the easy integration of data from multiple IMs; as a result manual alignment and integration of the data is required before cross-border simulation or planning work can be performed.

For the near future, the consolidation of different data sources by individual stakeholders builds the focus of work in order to establish a uniform interface to single data sources that are immediately updated in case of changes. All data resources are versioned in order to unambiguously trace each single change in data sets. Conflicting data sources are detected by automatic cross checks and manually corrected in a short term. These newly created interfaces build the basis for all business processes.

Appropriate topological granularities give a guideline for sorting different views of the same physical assets (Figure 4-3). See section 2.1 for detailed description of each granularity level.

At the first step, no aggregation or disaggregation is envisaged. The correct assigning of available data sets to its contained granularity levels is much more important and a necessary basis for further enhancements. A sound network topology is a core component of all data formats and therefore serves as guidance for the data consolidation process.

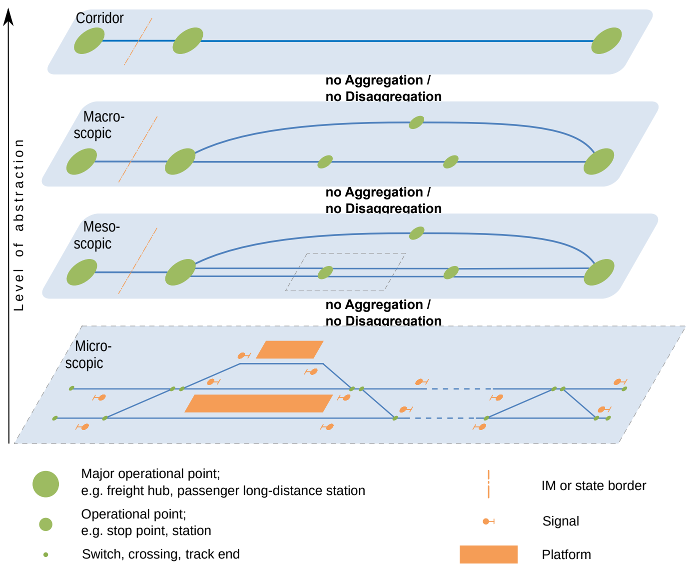

Figure 4-3 Vision 2020 – no aggregation of data between granularities

The data formats and open data movements were established with different intentions and therefore handle different topological granularities (Table 4-1).

Table 4-1 Capability of data concepts to handle topological granularities in 2020

Data format/ Data Granularity	railML 3	IDMVU	INSPIRE	RINF	OSM
Corridor	Possible	Out of scope	Not available	Out of scope	Out of scope
Macroscopic	Possible	Possible	Possible	Possible	Possible
Mesoscopic	Possible	Out of scope	Possible	Out of scope	Possible
Microscopic	Possible	Possible	Out of scope	Possible	Possible

Obviously, some data formats are more and other less suitable for providing an adequate syntax and structure regarding available railway related data sets from IMs. The introduction of a unique and consistent interface that is suitable for different business processes requiring different topological granularities builds the base for the next steps focussing on interoperability of railway infrastructure data.

4.3 Perspective 2030 – Consistent data in a uniform format

Rail infrastructure data is provided in a common, open format by IMs at multiple, predefined levels of granularity appropriate to operational and planning tasks.

Automated consistency checks ensure old or conflicted data is flagged for review by expert staff.

The mid-term vision for 2030 assumes, that open data formats are in use, enabling data exchange between operation planning tools without the risk of conflicting definitions / usage patterns or contextual ambiguity.

As next step data aggregation from microscopic to mesoscopic level shall be provided (Figure 4-4), enabling disaggregation from mesoscopic to microscopic level along the way. In most cases, this data development step can be performed within the responsibility area of a single IM facilitating consistency checks.

No new data has to be collected. Furthermore, available data sets have to be enriched and smoothly linked within the same data model. UIC RailTopoModel may be again used as guideline for this data development process.

Nevertheless it must be ensured, that the definition of each granularity is unambiguous and strictly complied by each IM in order to be able to exchange consistent data. Prospectively the ability to aggregate from detailed to more abstract granularities is required for a seamless and manageable operation planning. Data aggregation will be crucial for future use.

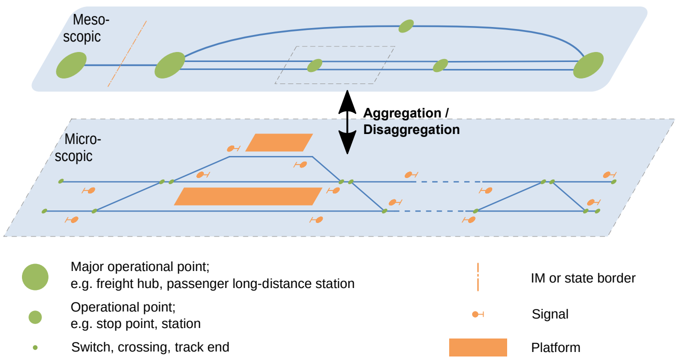

Figure 4-4 Vision 2030 – aggregation / disaggregation between microscopic and mesoscopic level

Simulation tools require different granularities for short-term and significant results. Basing a simulation on macroscopic instead of microscopic data may result in not enough resilient outputs. On the other hand, feeding microscopic data into the same simulation tool may produce better decision support but in a disproportional amount of processing time. In those cases the mesoscopic level would be the best compromise.

If a data format does not provide the ability to aggregate from the microscopic to the mesoscopic level, it is probably not sustainable considering further granularities. Table 4-2 gives an overview over the aggregation capabilities of the considered data concepts.

railML 3 will provide flexible aggregations from microscopic through mesoscopic granularity level. IDM^{VU} does not consider aggregation between both levels. INSPIRE instead does not provide data on a microscopic level. Mesoscopic level is out of scope for RINF data. Finally, OSM can provide data in any granularity (as far as they are available) and is able to aggregate from one level to another.

Table 4-2 Capability to aggregate from Microscopic to Mesoscopic level

Perspective 2030	railML 3	IDM^{VU}	INSPIRE	RINF	OSM
Aggregation microscopic → mesoscopic	Possible	Mesoscopic level out of scope	Microscopic level out of scope	Mesoscopic level out of scope	Possible

IDM^{VU} as a national initiative seems not to be sustainable for a European-wide approach, because of the missing aggregation model.

RINF AND INSPIRE

RINF uses partly the same information that are already collected for INSPIRE. Object types have been aligned between RINF and INSPIRE concerning 'rail' within the data theme transport networks. Details about harmonisation of attributes and definitions of INSPIRE 'rail object types' are listed in Annex H of "IU-Recommendation of RINF-final Report" (ERA, 2010). It is clarified, where a mapping between RINF and INSPIRE is directly possible and where a RINF-specific extension of the INSPIRE transport network model (i.e. a RINF-specific INSPIRE application schema) might be necessary. The goal is to make RINF data available in compliance with INSPIRE.

Possible future cooperation from RINF with the INSPIRE project were identified. Table 4-3 shows benefits for cooperation using geographical information for RINF:

Table 4-3 Possible future cooperations INSPIRE – RINF (ERA, 2010)

Cooperation	Implementation	Benefits
Join INSPIRE and RINF data (short term)	Use common identifiers for operational points and tracks	Data can be assessed which is not provided by the other source (e.g. geometries, area objects)
Overlapping data available as INSPIRE RINF application schema (short term)	RINF data available as GML compliant with RINF GML application schema	Same data fulfils (some of the) requirements from two legal obligations
Complete RINF data available as INSPIRE RINF application schema (long term)	RINF GML application schema adds also additional properties/parameters	INSPIRE data becomes accessible to RINF users and vice-versa, RINF can benefit from developments and implementations already available in INSPIRE

So far none of these cooperation tasks started. The INSPIRE development and utilization will be thoroughly observed.

RINF AND RAILML 3

RINF is identified as one of several use cases for the development of railML 3 and is being used as evaluation of railML 3 based on the UIC RailTopoModel.

A proof-of-concept tool, which shows that railML® is able to handle RINF requirements, exists in form of the software *railML4RINF*. This tool is capable of converting railML data into the current RINF format (Figure 4-5). To cope with incomplete railML data a configuration file for default values is provided. Fields that are not compatible with the conversion process are identified and automatically reported to the user (Fraunhofer IVI, 2014).

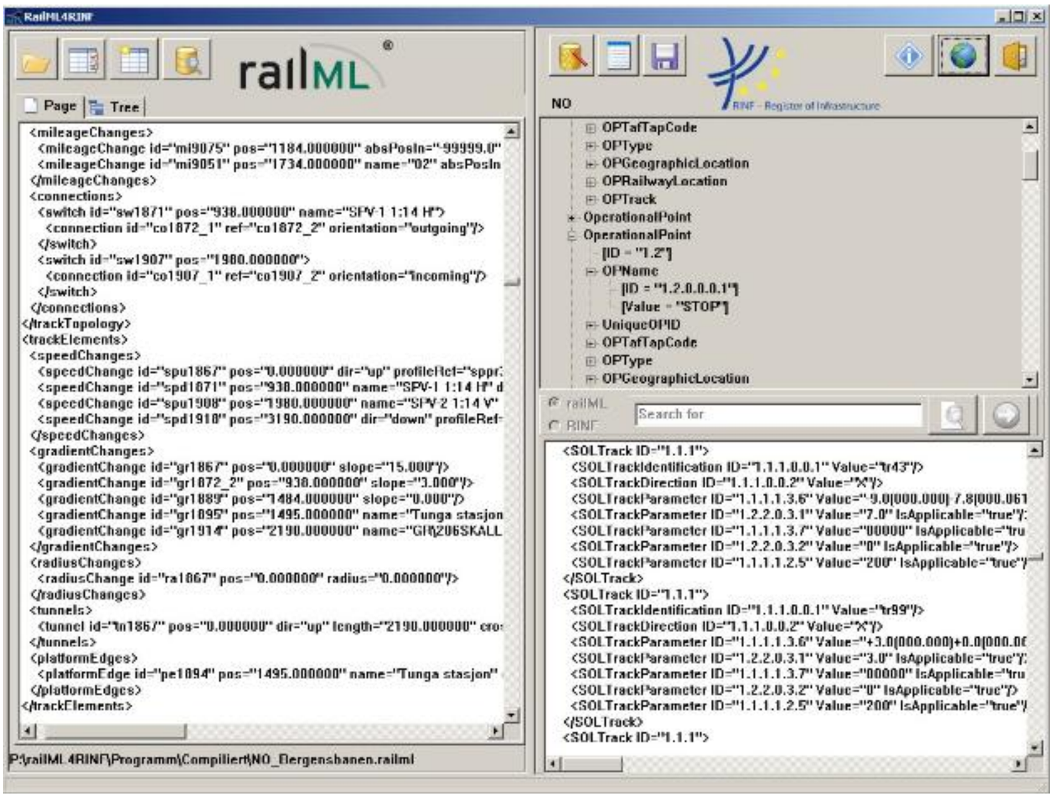

Figure 4-5 railML4RINF converter (Kolmorgen, 2014)

OSM – OPENRAILWAYMAP

The development of OpenStreetMap / OpenRailwayMap is independent from EC or IM driven initiatives. As the OSM contributors work voluntarily without time constraints and based on their personal interest no development assumptions are possible, neither regarding the topics nor the schedule. However work in the railway sector will proceed, alongside efforts related to other transport modes.

The independence at OSM on IM / industry data enables IMs to use OSM as a “public backup” for infrastructure data (respecting legal aspects) and to benefit from the “mappers“, who complete and update the IM’s data with own intentions (Figure 4-6).

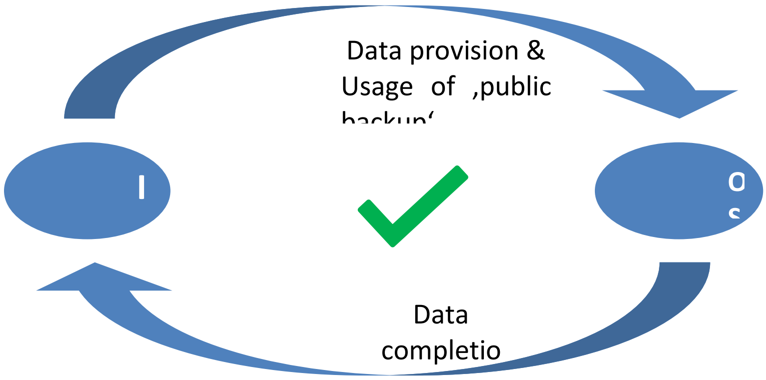

Figure 4-6 Mutual benefit potentials of IMs and OSM

DATA CONSISTENCY CHECKS

The aspect of closing gaps is already described in the previous paragraphs, but the aspect of detecting inconsistent data sets and their repair has to be addressed, too.

Wunsch outlined a way for checking railML data sets staying in the XML domain in (Wunsch, 2010). This means not only looking for well-formed XML files or meeting the XML Schema definitions, but also verifying the intrinsic railway-specific data relations. Additionally data sets may be checked against use case specific constraints, which would be even stronger than general rules and thus guarantee the applicability of the data sets for certain scenarios.

This way of data checks require the development of Schematron rules:

- general ones regarding the railway system;
- supplementary specific ones for the data format and model (railML 3/UIC RailTopoModel);
- supplementary use case specific rules (see also page 33);
- supplementary project specific rules are conceivable.

4.4 PERSPECTIVE 2050 – LINKED STANDARDIZED DATA EXCHANGE

Distributed, single sources of truth for rail infrastructure data are available at the IM level, thanks to common concept-based treatment of models.

Detailed data can be auto-aggregated to support planning, operations and simulation of the infrastructure at any granularity (micro, macro, etc.). Data models relate to over-arching representations of the railway system, allowing interoperability of data across domains.

Moving beyond the provision of an aligned data exchange format, the longer-term vision for 2050 considers that the topological granularities are established. It is assumed that switching between the granularity levels results in minimised time and computing effort, e.g. for long-distance operation. Sections with high coordination efforts (e.g. passing points, stations) can be considered on a microscopic level whereas sections without relevant alternatives can be handled more superficial on a mesoscopic or macroscopic level (Figure 4-7).

No data has to be provided at each level, but it may be aggregated 'upstairs' and disaggregated 'downstairs'. Available data are sorted into its equivalent level of granularity with no need for further data acquisition from the model's point of view. The model allows for fluid switching between the levels of granularity without loosing the track for the desired result.

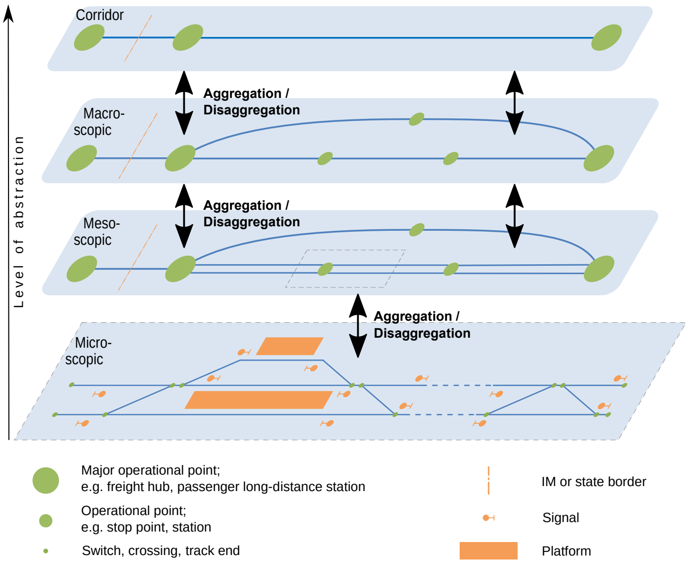

Figure 4-7 Vision 2050 – Data aggregation between any granularity level

In terms of data formats, it is assumed that the UIC-RailTopoModel in combination with railML 3 will fulfil this requirement. In 2050 the ERA-driven RINF format will be fully integrated into the model. Furthermore, INSPIRE would benefit from incorporating the UIC-RailTopoModel as generic topology approach for different modes of transport.

As OSM will be developed independently, railML data can be evaluated with the help of OSM data.

Furthermore the best-structured data sets are enriched through semantic annotation and interlinking with other data sets. The RaCoOn may serve as basis, but has to be further aligned to the topology model and enhanced by extended railML 3 data concepts resulting from the use case approach. Thus enriched data sets may bridge between different traditional software systems and newly emerging services, e.g. in multimode scenarios.

As introduced in section 3.4.1, Lodemann et al. demonstrated semantic railway infrastructure verification and debugging with the help of a hierarchical ontology set. Whereas 'debugging' means 'pointing to the source of error'. This concept may be adapted to semantically enriched railML 3 / UIC RailTopoModel data sets regarding different levels of granularity and different use case specific constraints.

New services are developed, which steadily check the data sets for their consistency using other sources of the Web of Data. These high-quality data sets are made available to other interested parties, in order to ensure best methods for planning, operations and simulation.

4.5 Gap Analysis and conclusion for story 1

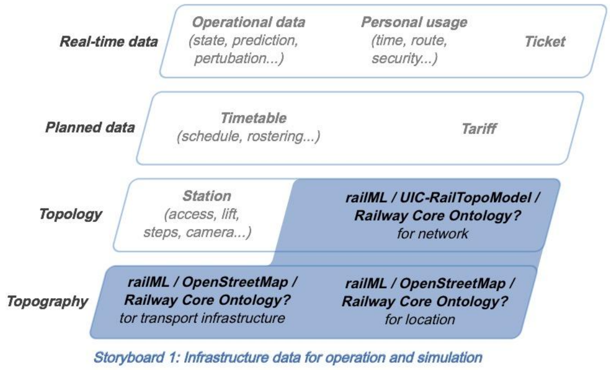

Figure 4-8 Data formats, models and concepts for storyboard 1

Summarizing the discussion in the perspectives for 2020, 2030 and 2050, most future-proof data formats, models and concepts will support the story 1 for achieving higher capacity in optimized railway operations (see Figure 4-8).

Increased capacity will be accomplished by the application of planning, dispatching and simulation across IM and state borders. These processes heavily rely on sound infrastructure data, which will be provided in a uniform data format (railML 3) for different, but clearly assigned, levels of granularity regarding a uniform topology model (UIC-RailTopoModel).

Subject to further research shall be a comparison of the Schematron-based and the ontology-based approaches for semantic railway data verification regarding rule creation effort, ability of real-time processing and flexibility for future adjustments.

Reporting obligations, such as RINF or INSPIRE, will be complied with standard processes across the EU, reducing implementation efforts. Ongoing integration of RINF into railML 3 is subject to the ERA.

railML 3 as ★★★★★ LOD would enable linking to other sources of the Web of Data. Further development of RaCoOn may be aligned with the upcoming raiML 3 in order to create semantical data sets. Interdependent development of RaCoOn and railML may be supported by standard tools that are capable to translate UML class models into ontology syntax, and vice versa. Thus enriched data sets enable automatic data validation and correction suggestions applying Semantic Web technologies.

Ontology-based processes may drive detection of data gaps in the IM's asset databases.

While OSM data is already available in various qualities, robust services may be developed, which fill the data gaps with converted crowd-sourced OSM data honouring different mapping styles and levels of detail.

The OSM community may gain data donations, which will be kept up-to-date by their members, enabling an early warning system for changed assets at the IM’s databases. For such a service, on the one hand fix identifications may ensure re-integration of changed IM’s data, but on the other hand, the changed asset may be semantically identified not relying on identification marks.

Applications will visualize the more detailed railway-related OSM data as base maps, attracting data correction processes both at IMs and open community.

Based on these findings, the overall question, how to link the data sets in applications will be answered in deliverable 3.4.2.

5. STORY 2 – EFFECTIVE USAGE OF CROSS-MODE CAPACITY

This section refers to the second storyboard in section 2.2. The "Effective usage of cross-mode capacity" primary aims at motivating people to use public transport systems, especially railway systems, as the project CAPACITY4Rail focuses on this transport mode. A well-informed passenger plays an important role for increasing the share of public transport and thereby increasing the effective capacity of railway systems.

Besides environmental and economic motives, following aspects play a role in transport mode decisions:

- Individual effort to inform oneself about possible alternative journeys;
- Reliability of information;
- Availability of transport modes for the specific journey;
- Comfort of inevitable transfer.

Customer satisfaction may be increased with up-to-date and easy accessible passenger information. However, there is a risk that the best-informed passengers could be headed to unusual journey variants, which represent the best compromise of time, costs and comfort, whereas the non-informed passengers take their typical routes with already known delays.

Online Passenger Information Systems (PIS), such as journey planners, support people with needed information in order to make reasonable decisions for getting to the destination at the best compromise of time, costs and comfort. Therefore the offered transport modes and services play an important role in this process. PIS are supplied with data from transport operators concerning routes, timetables, and fares, in various data formats. Across Europe, several guidelines and standards were developed to manage the information flow between transport service providers in order to offer adequate public transport facilities. In contrast to infrastructure data formats where the XML-format is well established the data format landscape for PIS is fragmented.

For the aspect of effective usage of multimodal transport systems, the following data formats, models and concepts are of interest:

- GTFS, see section 3.13;
- NeTEx, see section 3.7;
- TAP TSI, see section 3.12;
- railML / UIC RailTopoModel, see sections 3.2 and 3.3;
- Transit Ontology, see section 3.19;
- NEPTUNE Ontology, see section 3.20;
- Ontology for advanced passenger services and data, see section 3.21;
- Real-time data acquired through semantic sensor data, see section 3.22.

5.1 DATA FOR PASSENGER INFORMATION

Figure 5-1 illustrates the range of interlinked data resources needed to support passengers in the effective use of available multimodal capacity.

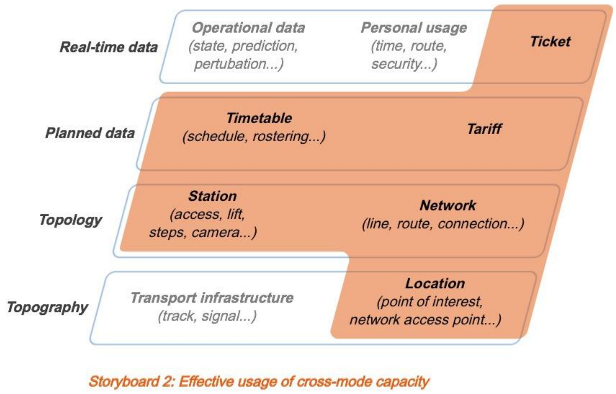

Figure 5-1 Data classes for storyboard 2

Multimodal journey planning requires access to static information, such as infrastructure, routes, planned timetables, and dynamic information, such as real-time vehicle positions or ad-hoc service changes, regarding all involved transport modes. The information may contain raw data that is updated periodically. Most transport operators use their own online information and ticketing systems. For a multimodal PIS the data formats used by the various operators have to be harmonised (AWT Consortium, 2014).

Current projects, such as the Full Service Model Initiative (FSM) driven by the Community of European Railways (CER) and ticket vendors associations, aim to set up industry standards for the transport sector, paving the way for unlimited multimodal trip planning. The Initiative is driven by a consortium of railways and ticket vendors, represented by the ticket vendors associations ETTSA and ECTAA and by the railway association CER. (FSM Initative, 2014), (CER; ETTSA; ECTAA, 2013)

Leaving aside the necessity of passenger information systems to deal with multimodal transport systems, each public transport party needs to receive any needed data in order to enable intermodal transport information. Required information shall be available to stationary and mobile staff as well as to the passenger itself. The high demand on alternative routes integrating alternative transport modes occurs essentially in case of delays or disruptions.

Dynamic passenger information screens facilitate access to actual information or alternative connections / routes via widely spread mobile devices. Both information concepts shall consume the same data in the same data format providing location and situation based services.

TIMETABLE

The most obvious inputs for passenger information systems are timetables of available transport modes. Knowledge of actual departure times is essential for considering a (spontaneous) change of modal transport system. Nevertheless, further information are needed for decision making:

- Validities of timetable;
- Travel duration;
- Number and duration of vehicle changes;
- Vehicle equipment (for PRM, toilets, bistro);
- Vehicle category (long distance, regional, urban transport vehicles);
- Facilities at start, transfer and end stations (e.g. lounge, WLAN);
- Compliance with further personal demands.

STATION (INCLUDING TRANSFER AND NAVIGATION PATH INFORMATION)

In order to include the aspects of multimodality and passenger needs, stations shall be described in detail as they serve as system access and provide transfer facilities. Information about available interchanges such as bus stops, metro entrances or taxi stands, are elementary to guide passengers between modes.

Accessibility to the railway system within the station, e.g. entrances with opening hours, lifts, stairs, is of similar value as are attributes concerning PRMs, positions of ATMs, etc. Furthermore, the number of platforms, the links between them and their sections should be described.

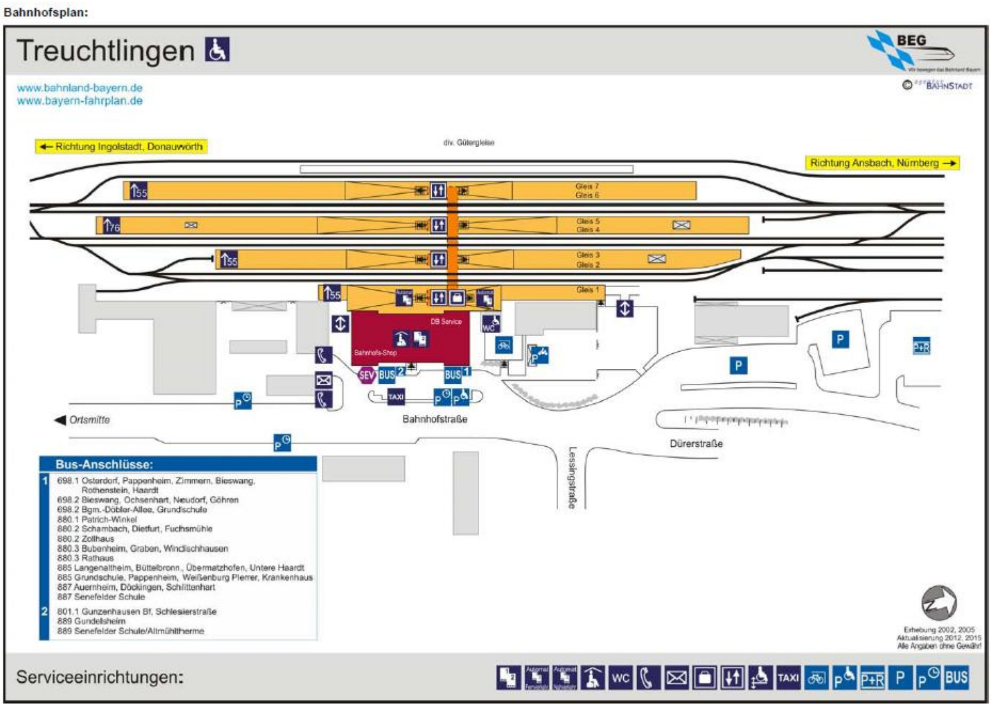

Figure 5-2 Example of a railway station map (http://bahnland-bayern.de/stationsdatenbank)

Generally spoken, the desired data exchange format shall provide data structures for describing station elements as shown in Figure 5-2.

Concerning transfers in general and in between different transport modes, the following transfer conditions are representative:

- Transfer between two specific journeys;
- Frequent transfer between two regular services;
- General possibility for transfer at a certain location.

Navigation Path information plays an important role in multimodal routing applications also considering access restrictions for persons with reduced mobility, e.g. wheel chair or baby carriage for estimating reasonable and individual transfer times.

A navigation path is composed of several parts, e.g. plain sections, stairs, lifts (see Figure 3-3). Combining navigation path information with points of interest (ATM, ticket sales) discloses an emerging business case.

Optimized navigation paths may significantly improve the acceptance of provided systems, taking not only paths within a PT building but also between nearby PT entry points into account, e.g. footpath from Gare du Nord to Gare de l'Est in Paris or navigation path from the Suisse railway station (SBB) to the French railway station (SNCF) in Basle.

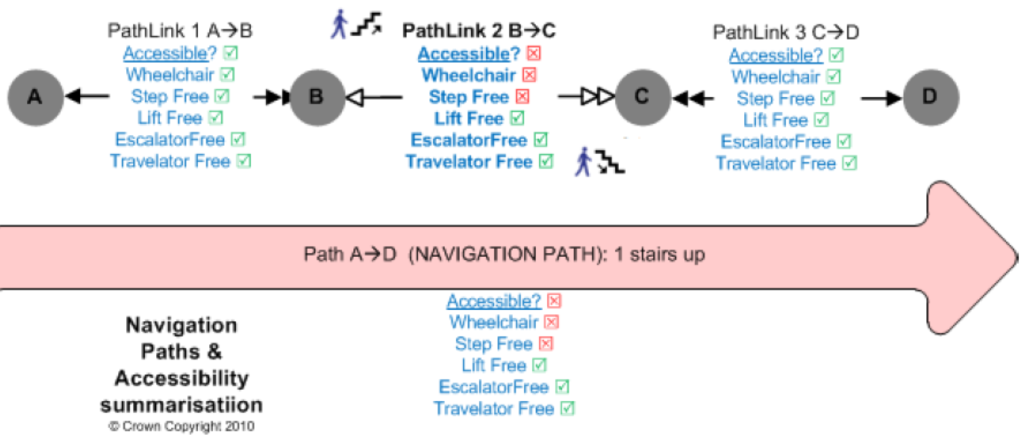

Figure 5-3 NeTEx Navigation Path Accessibility Example from CEN TS 16614-1

FARES

Ticket fares highly influence the passengers' motivation for using PT in general or for moving a journey from busy peak to less crowded off-peak hours. As a result, various fare models are in common use in a range of modes (Figure 5-4).

While particular fare models target specific usage scenarios, the combination of different fare models during a journey often comes difficulties.

One classic ticketing model is based on distance or zone based tickets. Fares sometimes vary regarding the selling point, e.g. ticket for mobile devices, pre-sale paper ticket or on demand paper ticket. Furthermore flexible price policies, e.g. the approach of Yield Management (YM), as successfully applied by some airlines, invade the PT market.

Roughly speaking, Yield Management means "selling the *right seat* to the *right customer* at the *right price*" in order to spread the volume of passenger evenly. However, YM is not considered for ground-based PT analogous to air traffic, as the willingness to commit oneself to a certain trip by reservation is only given for long-distance travel and not for urban or regional trips where the frequency of trip is higher (Li, van Heck, Vervest, & Rooijmans, 2006).

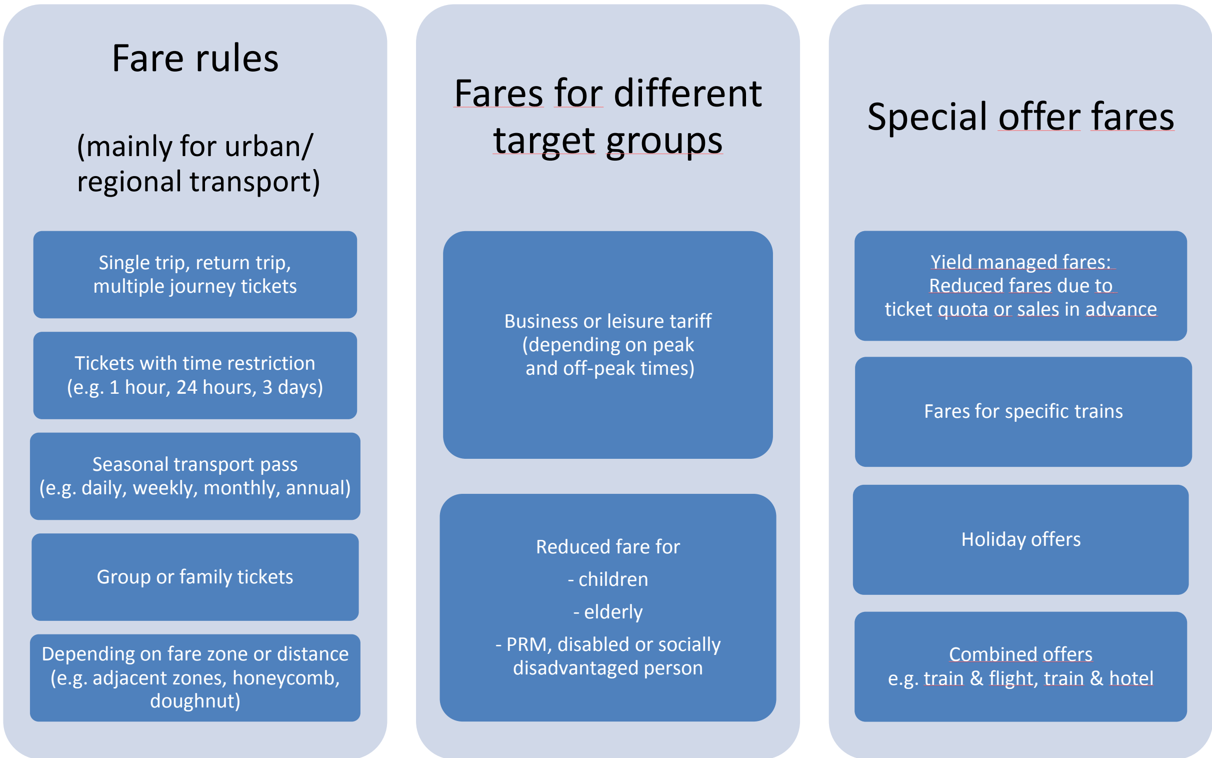

Figure 5-4 Facets of fare structures, influencing factors

In some cases urban or regional fare models are already linked with national long-distance journey tariffs. The exchange of fare information between operators and handling of ticket sales is hard to harmonize, especially in the case of multimodal transport systems. From the perspective of operations, passenger loading in individual modal segments must also be calculated, adding a further aspect to the multimodal ticketing question.

5.2 Overview 2020 – Spreading information

Disruption results in lengthy delays for passengers as they wait for resolution of issues / later services. Passenger and staff may be left out-of-position for follow-on travel.

Transport modes that offer well-known alternative connections may be heavily congested while less obvious, but not less valid route choices have spare capacity available.

Impacts of disruptions vary according to the affected transport mode as well as the efforts to overcome. A road closure due to an accident can in most cases be easily bypassed, while trains affected by a track closure due to train collisions are more complex to reroute.

Besides the handling of these operational aspects, the information flow requires more or less the same data to enable passengers who have already started their journey to choose alternative connections. Relevant data for passengers' decision include information about

- Timetabling;
- Station and Navigation paths;
- Intermodal transport options;
- Fares.

The following paragraphs illustrate whether the available data formats presented in chapter 3 cover these aspects.

GTFS

Timetabling and fare data are distributed in several text files. Geographic locations of stations or stopping points are given inside stops.txt, as geographic coordinates are fixed to the Coordinate Reference System World Geodetic System 1984 (WGS 84), which is also used by the Global Positioning System (GPS).

Station aggregation is also foreseen in this file through "parent station" bindings. Supplementary information, such as wheelchair boarding or carriage of bicycles is offered, too. The GTFS feeds treat any transport mode the same way. Within the transfers.txt details for minimum transfer times may be specified.

Fare information can be given per transport operator according to simple rules or combinations: depending either on origin to destination station or on zones the itinerary passes through or on which route the itinerary uses.

In summary, the GTFS feeds may deal with most of the requirements except navigation paths. Although the coordinates of different stops (e.g. platforms) are implemented, the paths between them (e.g. including stairs, lifts) cannot be figured out, as a station plan is missing supporting the concept of an available routable map.

Compared to Transmodel (section 3.7) GTFS lacks the following aspects (Knowles & Miller, Transmodel for google, 2008):

- Rail Services with splitting trains;
- Multi-operator services;
- Stop Areas, Connection Links Interchanges;
- Station and Transit Interchange Navigation and Accessibility;
- Connection Protection or Guaranteed Connections;
- Services across Midnight;
- Services across Time-zones;
- More general Day Types;
- Different Stop Labelling;
- Transport Modes, Vehicle Equipment & Accessibility.

NeTEx, based on the latest version of Transmodel, has the potential to fill these gaps.

NeTEx

NeTEx covers all aspects that are needed to inform the passenger. Timetabling and stations can be described in detail. Validity triggers (external events known to result in disruption to normal service, e.g. flooding, bad weather, track/road closure for works) can be activated, which are linked to special timetables.

Besides these elements incorporated from Transmodel, the description of navigation paths is developed. To demonstrate the concept, the NeTEx standard includes comprehensive examples for navigation paths, e.g. through the Olympics Main Site in London 2012 or Wimbledon Station (CEN/TS16614-1, 2014). Transport modes are treated consistently, so intermodality is assured.

NeTEx dedicates the third part of the standard to fare information and tariff structures. The handling of fares is highly flexible. Among others, it will be able to meet the requirements of the TAP TSI B.1 to B.3 documents as well as national, regional and urban particularities (Knowles, 2014). Furthermore, it covers electronic payment cards, such as Oyster (London region), Navigo (Paris region) or Sube T (Madrid) and other complex fare structures.

Zone based fare systems of any topology (adjacent zones, honeycomb, doughnut, etc.) can be described as well as mixed zonal and stage systems. Furthermore, NeTEx is able to deal with search parameters to find the best fare for a user (e.g. age, possession of rail cards) including Yield Management. (CEN/TC278/WG3 SG9, 2015)

TAP TSI

TAP TSI components, such as timetabling and tariff data, could be easily adapted for other transport modes. Nevertheless new fare concepts are hardly adaptable and would require new data formats. So far, the focus of TAP TSI is on interoperability and international traffic, specific features in urban or regional traffic are not considered.

Description of stations and thus Navigation Paths are out of scope and not feasible. Hence, intermodality cannot be treated extensively.

railML / UIC-RailTopoModel

RailML is intended to support railway operation but may be also used to for passenger information, which is relevant to facilitate the effective usage of cross-mode capacity. Timetabling is one sub-schema of railML.

Stations are implemented in the ‘IS’ schema and considered as operational control points (‘ocp’). The operational point of view reveals for instance in the fact, that locations of stop signs and signals are provided, whereas different platform sections for allocating certain coaches are currently missing. Despite, railML concentrates on the infrastructure, station buildings or access to the platforms are not offered. Neither stations nor navigation paths can be satisfactory defined. The railML structure is railway specific and not suited for adoption in other transport modes. Fare information is also out of scope.

Linked (open) data based on ontologies

As successfully demonstrated with Transit Ontology (see section 3.19), NEPTUNE ontology (see section 3.20) and an application specific ontology for advanced passenger services and data (see section 3.21) enriching data sets with URIs and interlinking them will support a range of new, interacting applications in the field of public transport services.

Comparison

TAP TSI is identified as a railway specific concept that includes several data formats, with focus on interoperability of railways, thus it doesn't seem to be a sustainable concept for cross-mode capacity and passenger information.

GTFS and NeTEx are intended to cover multimodal transport modes, as illustrated by an increasing number of Public Transport companies provide their GTFS data to Google or others in order to attract more passengers (AWT Consortium, 2014).

The development of NeTEx is intended to harmonize previous EU standards (Transmodel, IFOPT, VDV 452, TransXChange, Trident, Bison) and to complement the SIRI standard. Furthermore, NeTEx could circumvent formerly disclosed problems and therefore provides the best state-of-the-art standard for its domain.

railML uses an XML format, which facilitates data integration with other XML-based data concepts. As railML is designed for data exchange in railway systems and does, at least in short term, not consider other transport modes, the combination of this along with railML's focus on operational scenarios causes model insufficiencies for detailed passenger information.

Supplementing NeTEx with railML 3 and the UIC-RailTopoModel respectively would link the multimodal passenger centric view with the railway-specific operational view, thus enabling new fields of application. The thorough combination of NeTEx with railML through linked data technologies may provide a strong backend for flexible and reasonable multimodal passenger information concerning adaptability, extensibility, use of standardised interfaces and common architecture.

Although NeTEx is the only data format that provided navigation paths, it is hard to imagine that format to be used for the near future of 2020. So far, the NeTEx standard is not widely adopted, and conversion tools are needed for the first step getting from proprietary formats to NeTEx until tools will export NeTEx natively.

5.3 PERSPECTIVE 2030 – ENHANCED, WIDE-SPREAD TRAVEL INFORMATION

Passengers have easy access to data on alternative connections, including those that run via different modes.
Open ticketing information allows informed decision-making, allowing the correct trade-off between cost, speed of transit, and available facilities to be made.

The effective usage of multimodal Transport Systems is already applied in various urban and regional environments: Public transport services work together in transport associations in order to coordinate their offers to gain customers' satisfaction by improving connections, minimising transfer times and adopting their services on passengers' demand. Public transport users raise a claim to get their journey organized without consulting different public transport undertaker for planning and transfer modalities.

By this point, it is likely that public transport operators (PTO) will inform their passengers not only about their own connections, but will also offer data from other PTOs as it is already done for (regional) transport associations. This development is due to the fact that the market for open source as well as commercial multimodal travelling information systems is emerging.

Several movements are uprising to improve the exchange of multimodal data and to inform passengers, e.g.

- National Authorities: The French Ministry of Ecology, Sustainable Development, Transports and Housing (MEDDTL) develops the open source software "CHOUETTE" (http://www.chouette.mobi) in order to facilitate public transport theoretical information exchange by specifying a XML exchange profile. The tool addresses to Transport management Authorities, PTOs and their contractors, others State services, software developers, information service providers as well as researchers;
- European Approach: the "All Ways Travelling (AWT)" consortium works on a model for a multimodal pan-European passenger transport information and booking system (AWT Consortium, 2013-2015);
- Global player: Google offers with "Google Transit" (http://www.google.com/transit) an instrument for passengers to inform themselves about different transport modes available – as far as data are provided by the PTOs;
- Driven by Operator: DB Vertrieb GmbH (Distribution) offers a "personal mobility advisor" called QIXXIT (https://www.qixxit.de) that claims to include any transport mode available irrespectively who prides the transport service (e.g. competitors, coaches, car or bike rentals).

It cannot yet be foreseen which approach will prevail, but the need for harmonized data format to easily exchange multimodal data is undisputed. Although multimodal information systems may deal with different data formats and convert them, this is not sustainable and rather seen as a step towards uniform data formats.

As far as shown in the advanced public transport service (see section 3.21), linked open data offer quite awesome mobile applications, where the public transport trip assistance may also play a minor role giving the person in charge the feeling of freely catching the desired targets at low cost. The transport mode decision is not central, but satisfying the mobility needs at best compromise of time, costs and comfort.

5.4 Perspective 2050 – Capacity increase through well-established guidance

Good utilisation of capacity allows for frequent services. Flows of passengers can be controlled at peak times via dynamic modal advice to individuals, managing their arrival / departure at long distance terminals and spreading load across feeder networks. Disruptions are handled through seamless interaction between modes, electronic tickets allow free-flow of passengers between available transport modes.

The perspective for 2050 considers an active routing of passengers based on well-established passenger information systems consuming sound data. So far, passenger navigation is only possible from platform to platform, and in case of reservations from one boarding position to another. Assumed that the willingness to commit oneself to a specific journey, especially for regional and commuter trains, the need for other indicators than reservation is obvious to enable navigation from alighting to boarding position.

Active passenger routing from alighting to boarding position

In case that no reservation is done, passengers should be routed to a platform section where free seats in the train are available. This would lead to minimised time for passenger exchange, which is preferable not only in case of previous delays or train failures.

Therefore, train data are needed, e.g. the occupancy rate in each coach including number of available seats per vehicle, considering reservations as well as actual occupied seats. Furthermore, the knowledge about habits in commuter transport should be considered: Do commuters take their seat near their preferred exit/entrance or do they choose less crowded sections?

The potential exists for citizen based data acquisition (see section 3.22) to be used on older vehicles to supplement the limited data produced by the train itself by applying ontology-based reasoning at the data integration layer. Such methods may deliver both train positions, and maybe even occupancy indicators at mobile and stationary equipment, meaning train wagons and platforms, or interchange pathways.

Adequate reliable data on operations will increase the acceptance in micro-routing applications. Changed platforms or even wagon orders shall be correctly displayed or announced. Therefore the railway-specific data format and model railML / UIC-RailTopoModel enters the scene, playing its consolidated role based on more than 10 years experiences in productive systems updated to serve new use cases.

In case of system or connection failures, the re-routing of passengers to alternative transport mode should also consider the free capacity of the supplement means of transport. If the actual demand cannot be satisfied immediately, alternative routes shall be made attractive to passengers.

As already stated, NeTEx is able to deal with navigation paths in stations. Supplementary, data about occupancy rates and equipment of vehicles are required.

As demonstrated with GTFS (see section 3.19) public transport data may be enriched to ★★★★ LOD with the help of Transit Ontology (developed into: Linked GTFS) enabling new services for passenger information. While GTFS lacks some provisions for more complex scenarios and railway related concepts, NeTEx closes these gaps.

Deploying the NEPTUNE Ontology into the NEPTUNE data sets together with interlinking even reaches ★★★★★ LOD (see section 3.20) enabling versatile applications. But as NEPTUNE is a French settled public transport data standard, NeTEx comprises its concepts as a European initiative. Therefore it seems to safe future investments by going the NeTEx way enriched to ★★★★★ LOD.

ELECTRONIC TICKETING

Electronic ticketing enables passengers to spontaneously use different transport systems without explicitly dealing with various fare models. Electronic ticket cards are able to adjust fares according to the consumption, e.g. change fare of regular trips to a daily rate if this is cheaper. This approach requires the storage and evaluation of route profiles and the handling of complex fare models.

Flexible electronic ticketing with mobile devices across multiple operators opens the door for detailed multimodal passenger routing data deploying ontology-based reasoning. Furthermore, the interoperable extension of fare zones shall be possible.

5.5 GAP ANALYSIS AND CONCLUSION FOR STORY 2

Summarizing the discussion of the perspectives for 2020, 2030 and 2050, most future-proof data formats, models and concepts will support the story 2 for achieving higher capacity by increased passenger count (see Figure 5-5). In other words a set of existing models be used to meet the needs at this scenario, given the required industrial “hardening”.

While data exchange within a single mode may be established with low effort, there may be some value in considering multimodal links, both for freight and passenger services. Even though there is scope for modelling work in this space, the far more obvious work would be in the data management space, and particularly in the open data arena, where by making railway data available to the public at large, the CAPACITY4RAIL team could encourage the development of “mashup” transport apps and services by independent developers.

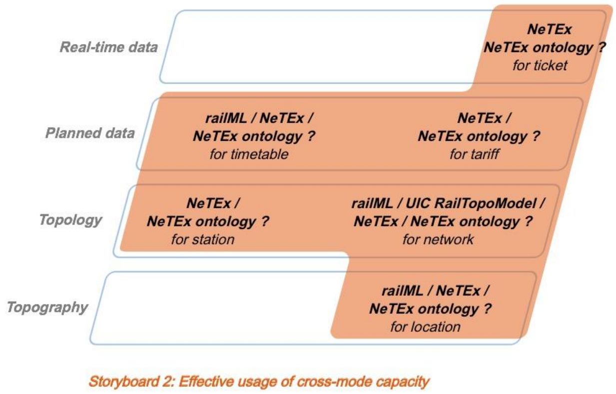

Figure 5-5 Data formats, models and concepts for storyboard 2

In terms of advising passengers of their best compromise of time, costs and comfort, during a journey a close interaction of operations from multiple operators and independent passenger information systems is crucial to a successful outcome.

Thus, railML 3 combined with UIC-RailTopoModel may serve the railway-specific operator's side. Operations of other modes, e.g. regional bus or local tram services, may be modelled with the UIC-RailTopoModel concepts justified by their specific needs. This approach does not aim at collecting new data, but on providing the available data in a uniform way enabling easy reasoning within the operator specific data sets.

Sophisticated passenger information systems, as well stationary as mobile, will guide the citizen to his desired destination. The passenger has trust in the technology when it advises him to reroute around a disruption, and this is key to the system remaining fluid. Consistent, accurate data is essential to achieving this.

NeTEx may serve PIS with multimodal transport data. Enriching NeTEx to ★★★★★ LOD will enable suites of powerful new applications to be developed. Therefore a NeTEx ontology has to be developed, taken the previously done research on GTFS and NEPTUNE in mind. As GTFS and NEPTUNE concepts are integrated into NeTEx, which is advanced for more complex scenarios, NeTEx may replace their current field of application.

Interlinking enriched NeTEx with semantic sensor data may be used for data acquisition by the users themselves.

Data responsibility is a critical subject regarding operational as well as legal aspects. Update intervals for data strongly influence the reliability of passenger information, which can only be provided on a well-defined responsibility basis. Publishing timetables as Open Data will be mostly done by the operators themselves instead of a general data provider, such as Google.

Based on these findings, the overall question, how to link the data sets in applications will be answered in deliverable 3.4.2.

6. STORY 3 REAL-TIME OPERATIONAL DATA ACROSS ORGANISATIONAL AND MEMBER STATE BOUDARIES

This section refers to the third storyboard (section 2.3) dealing with real-time exchange of operational data. In general, the coordination of railway traffic is planned in detail, and well in advance of the date of operations. Nevertheless, unforeseen events require immediate operator intervention in order to minimise the consequences of incidents.

Promising data exchange formats, data resources respectively data concepts are considered and evaluated in order to determine their suitability for and compatibility with the data related requirements of real-time operation decision support and real-time passenger information in both today's railway and the railway of 2050:

- SIRI / NeTEx, see sections 3.8 and 3.7;
- GTFS / GTFS-realtime, see sections 3.13 and 0;
- TAF TSI, see section 0;
- ON-TIME RTTP, see section 3.10;
- railML / UIC RailTopoModel, see sections 3.2 and 3.3;
- Ontology based reasoning for acquiring real-time public transport data, see section 3.22.

6.1 DATA FOR REAL-TIME OPERATION

According to the relevant data classes in Figure 6-1, real-time operational data across organisational and member state boundaries enabling optimized operations need sound timetable and operational state data smoothly linked to a solid network topology.

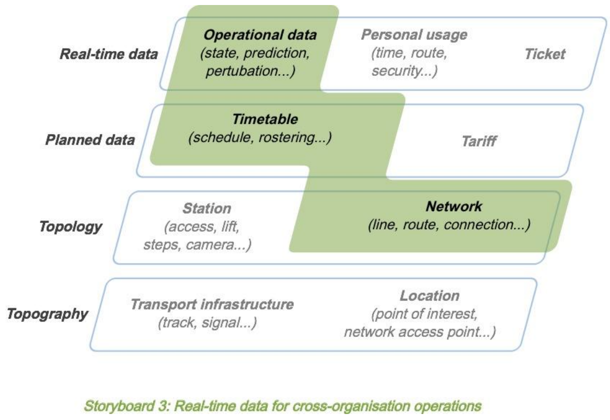

Figure 6-1 Data classes for storyboard 3

Data exchange formats for real-time data shall include any relevant data that gets influenced by irregularities within the network. The data format has to be compatible with infrastructure and operational data in order to minimise process times and fault interpretations.

For managing traffic in real-time, knowledge about the state of the infrastructure and vehicles is elementary. The challenge is not only to react on any deviation but also to absorb resulting chain reactions. Thus, real-time data are required not only to point out and react on traffic incidents, but also to enable operators to forecast train positions and speeds in order to optimise the traffic management decisions within the whole network.

This storyboard relates to traffic management decisions, as they are subject of WP3.2 “Simulations and Models” within the CAPACITY4RAIL project. Real-time operational strategies are often realised in local traffic management systems (TMS). However, in order to enable joined-up management of the network, particularly at the interfaces between the areas of responsibility of different Railway Operations Centres (ROCs), the individual route TMS must be able to communicate with each other – exchanging information on service timings, the position of vehicles, and any ancillary information associated with the service – and with external systems, such as timetable planning, vehicle and crew rostering, and track access planning tools.

This context is highly complex and requires sophisticated data architectures (e.g. UK-based LINX-TM), which will be further subject of deliverable D3.4.2.

The data formats and respective protocols of SIRI and GTFS-realtime are dedicated to multimodal public transport:

- GTFS-realtime is complementary to GTFS and aims on real-time information for passengers. GTFS and thus GTFS-realtime lacks railway-specific services, like coupling and sharing of trains.
- The implementation of SIRI is more extensible and comprises railway applications. SIRI works as supplement to NeTEx, also referring to Transmodel. Nevertheless, the recent developments on the SIRI standard direct towards the implementation of real-time data on station equipment and other multimodal passenger information services.

By comparison, TAF TSI covers only freight railway traffic from a European perspective.

For managing traffic in real-time, the technical state of the infrastructure assets and vehicles must be monitored to detect faults or failures, which might impact the control options. Furthermore, the traffic state (train positions and speeds) as captured by trackside and onboard sensors must be detected and continuously compared with the timetable in order to identify deviations (ON-TIME_WP04, 2014).

The RTTP data format developed in the ON-TIME project raises the claim not only to deliver real-time data, but also transmit operational state predictions for railway operations. This format is embedded in the Perturbation Management Module (PMM) of ON-TIME WP4, and the ON-TIME Architecture. The potential of the RTTP-format depends on the result of the PMM, which in turn needs accurate and feasible input data, especially monitoring data (see SP4 of CAPACITY4RAIL project).

The RTTP-format settles on top of railML 2.2 infrastructure, timetable and rolling stock subschemas, thus it can't be used without any modifications regarding railML 3. Since it is also based on the interlocking draft of 2013, this part has to be adapted to the newly developed railML interlocking subschema, which will be integrated into official railML 3.

6.2 Overview 2020 – Early information about on-route change

Plans for cross-border services available to undertakings involved, but data is at varying levels of granularity and not necessarily up to date if on-route changes have been made or disruptions are involved.

By 2020 it is anticipated that a variety of data formats and granularities of data will still be in use. This hinders IMs in data exchange, requiring specific routines for each IM-RU-relation. RUs that operate on several network areas with different IMs also need to deal with several formats and granularities.

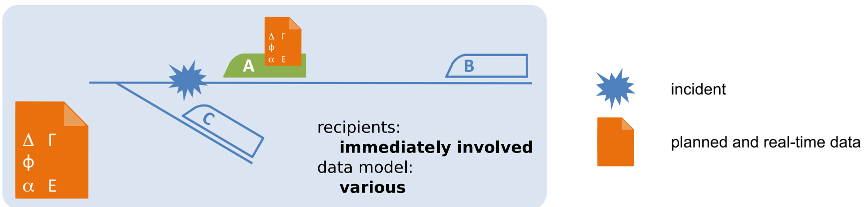

Figure 6-2 Vision 2020 – disruption data for the immediatly involved train

The data provision process gives a guideline for sorting different recipients and data models (Figure 6-2). See section 2.3 for detailed description of exploitation step. Immediately involved trains shall automatically get data about any kind of deviation or disruption.

So far, the SIRI standard provides the possibility to exchange real-time data. Thus, although any aspect is covered, the granularity for railway data may not be sufficient in any case.

GTFS-realtime instead provides only a subset of SIRI's facilities with the focus on passenger information. Operational views on disturbances cannot be reflected, and the format is not in use for operational procedures of railway systems anyway. GTFS-realtime offers no additional value for railway operations. Thus, the data format will not considered in the further analysis.

For a long time, the aspect of real-time data within the TAF TSI was limited on real-time monitoring of freight and trains. Nevertheless, the TAF TSI is a living concept and will be further developed. In 2014, the Commission Regulation (EU) No 1305/2014 repealing the Regulation (EX) No 62/2006 (European Commision, 2014) was published in line with the TAF TSI Master Plan (ERA Telematics Team, 2013).

The newly implemented TAF TSI functions "Train Running Information Message" (Target Implementation Milestone 2017) and "Train Delay Cause Message" (Target Implementation Milestone 2018) play a central role for real-time applications. Their functionalities are:

- Long-term planning;
- Path request on short notice;
- Train preparation;
- Train running forecast / Information;
- Service Disruption Information.

Figure 6-3 illustrates the communication links between IM and RU running the train. The "Train Running Information Message" includes information concerning departure from departure point and arrival at destination as well as arrival and departure at handover points, interchange points and at agreed reporting points based on contract (e.g. handling points). In case of delay, its cause (first assumption) must be sent in a separate "Train Delay Cause Message" as soon as it is identified (ERA, 2014).

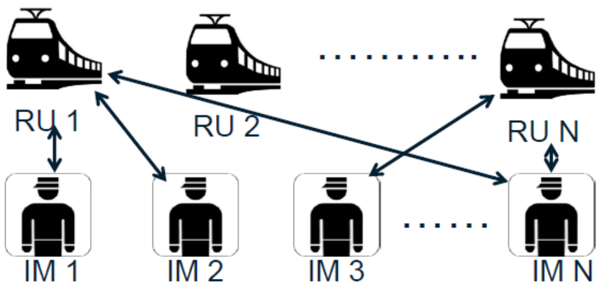

Figure 6-3 TAF TSI – Train Running Information Message (ERA, 2014)

Similar to feasible real-time data handling within TAF TSI, the ON-TIME RTTP is not yet in productive operation. So far, it was successfully demonstrated within the ON-TIME project prototype. The reliance on static railML data and dynamic Traffic State Monitoring data are deemed as promising approach. But RTTP has to be upgraded in order to benefit from railML 3's granularity concept.

6.3 Perspective 2030 – Immediate updated traffic plans

Cross-border service data available at appropriate levels of granularity and in consistent models for all operators / undertaking along the planned route. Changes to schedule of services pushed in real-time to undertakings as updated traffic plans.

As Figure 6-4 illustrated, not only trains in the responsibility area of the IM, where the disruption occurred, are informed but also other trains along the route in the same direction. Therefore a kind of subscription service is required, in order to acquire the perturbation feed already from a neighbouring IM area.

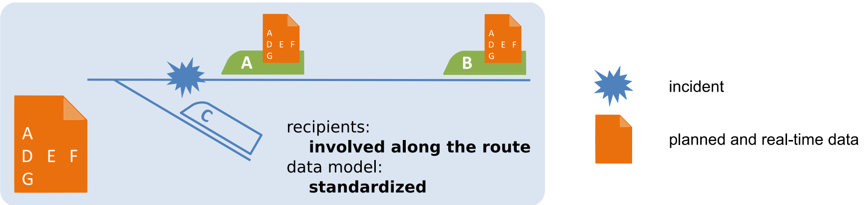

Figure 6-4 Vision 2030 – disruption data for trains involved along the route

The challenge for 2030 is to harmonise real-time data exchange between IM and RUs respecting several aspects such as kind of information, level of granularities and consistency. Furthermore, the harmonisation shall involve service data even beyond IM borders, which de facto results in a (European-wide) unified approach on real-time data exchange.

The conditions described are an ideal use case for the SIRI standard, but although the standard will be further developed, the focus will be on multimodal issues and on further alignment with NeTEx, e.g. the revision of part 4 and 5 (Station and Passenger Information Systems).

The situation of TAF TSI in 2030 is resilient. According to the TAF TSI Master Plan (ERA Telematics Team, 2013), relevant functions for real-time information (Short term Path Request, Train Run) shall be implemented at least in 2022, see Figure 6-5.

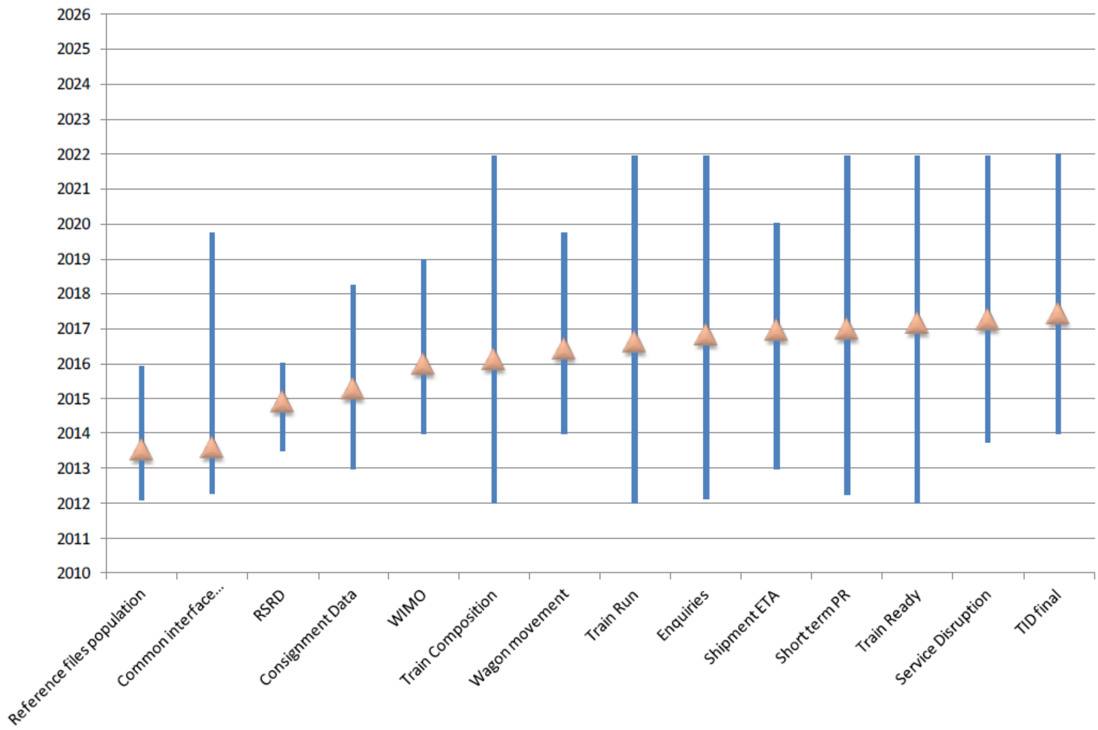

Figure 6-5 TAF TSI Functions – Minimum and Maximum Implementation Dates (ERA Telematics Team, 2013)

The upgraded ON-TIME RTTP has the capability to supplement the railML 3 standard in the domain of real-time information respecting the levels of granularity of UIC RailTopoModel. The upgrade process may be aligned with the TAF TSI real-time information developments in order to achieve at minimum a harmonized terminology.

6.4 PERSPECTIVE 2050 – REAL-TIME NETWORK STATE TO ALL RUS

Live views of network and asset state available at high resolution to operational staff. Real-time service and routing information available to all involved railway undertakings through concept-based models for infrastructure, operational practices, and asset condition. Real-time network information allows more accurate prediction of service arrivals and optimisation of live traffic plans.

The aim for 2050 is to provide real-time data for any interested party in a fully equipped level of granularity considering all available information. This includes the consideration of infrastructure and train data with support of available sensor data.

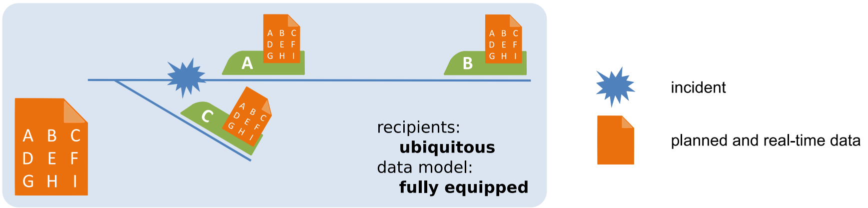

Figure 6-6 Vision 2050 – disruption data for all interested trains

This objective requires the availability of prepared sensor data, e.g. wayside data feeds from Hot Axle Box Detectors (HABDs) or trainside in case of propulsion failures. SP4 of CAPACITY4RAIL will focus on monitoring data for the railways, however it is unclear, what kind of monitoring data would be evaluated, when they'll be available, where they'll be collected, nor in what data format they will be presented or if sensor data will be clustered.

In the case of installed interlinked monitoring systems, consolidated train- or even axle-specific raw data are already reported to the train control centre (TCC), which in turn supports the maintenance team and dispatcher with scaled suggestions. In case of a hot axle box, the TCC commands a stop for the train outside tunnels. In case of reduced train propulsion, the TCC adjusts its track occupancy predictions and therefore may change the routes and/or times, furthermore it serves up-to-date passenger information.

SIRI would perfectly fit into the passenger information task, but this topic would be more subject to storyboard 2, while this storyboard focuses on data exchange for railway operations, which is at present not aimed by SIRI.

If perturbation data is already made available with a subscription process in 2030, it may be easily extended to all interested parties in the railway domain. The level of detail regarding the transferred data may depend on the role of the subscriber.

So far, no efforts are taken to respect sensor data within the TAF TSI functions. It must be observed, how the TAF TSI will be extended after the full implementation of the TAF TSI Master Plan.

The concept of the ON-TIME RTTP allows implementing various data, e.g. extended infrastructure for evaluated sensor data or train performance data.

Based on the perceptions at real-time data acquisition through semantic citizen sensor data in Scotland (see section 3.22), it seems to be promising, to extend the ON-TIME RTTP by base real-time related data deploying ontology-based reasoning for up-to-date sensor integration aiming at fully interlinked data sources across various IMs and RUs. ON-TIME RTTP would therefore be enriched to ★★★★★ LOD, but missing the publishing ('Open Data') facet.

Really published Open Data may comprise train locations and predictions for following stations, applying SIRI, while 'closed' data may be provided to RUs containing additional track occupancy predictions regarding certain trains and train order predictions regarding certain locations at the infrastructure.

6.5 GAP ANALYSIS AND CONCLUSION FOR STORY 3

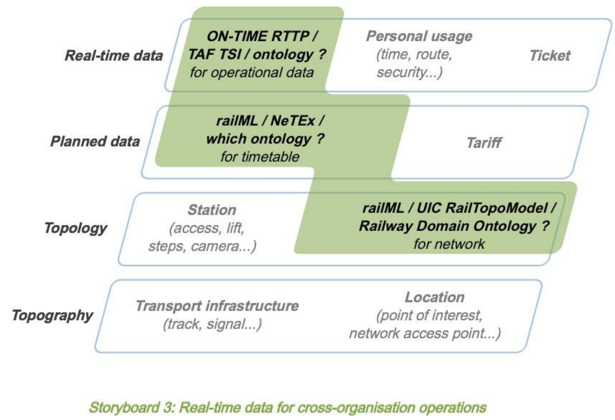

Figure 6-7 Data formats, models and concepts for storyboard 3

Summarizing the discussion in the perspectives for 2020, 2030 and 2050, most future-proof data formats, models and concepts will support the story 3 for achieving higher capacity by short-term optimized operations (see Figure 6-7).

Real-time data, such as infrastructure and rolling stock states shall be made instantly available to the IM. Track- and train-side monitoring systems collect sensor data and send standardized messages to the IM or RU in case of anomalies. These processes are developed in SP4 of CAPACITY4RAIL. Currently no concrete information about the data models, data formats and interfaces are available.

Based on the track- or train-side consolidated messages, TCCs re-calculate train paths, command stops in case of safety-related issues etc. Relocated train paths result in changed train stops, which have to be broadcasted to the passengers. RUs rely on more detailed data; they therefore get additional track occupancy predictions regarding that certain situation.

Both RUs and passengers information are defined in a uniform data format, which enables both perspectives and easy to integrate future enhancements.

All data have to be related to a solid topology network, which will be served by railML 3 with UIC-RailTopoModel. RUs timetable data needs may be served by railML, especially with its timetable subschema, which also comprises rostering constraints. While the passengers NeTEx may handle timetable needs.

Relocated paths or other path constraints may be exchanged with upgraded ON-TIME RTTP. Therefore the current ON-TIME RTTP protocol and data format has to be dissolved from the Perturbation Management Module (PMM) and demonstrated in a productive software environment. Furthermore it has to be upgraded from railML 2.2 related data to railML 3 incorporating different levels of granularity and the newly developed interlocking subschema. Afterwards, an extension for the monitoring requirements resulting from SP4 may be added. At the end, the newly born RTTP may be enriched to ★★★★★ LOD, with the 'Open' star in case of passenger information or without it in case publishing to interested RUs.

Thus enriched real-time data may be used for semantic reasoning, maybe also providing early warning for disruptions through citizen sensing.

Based on these findings, the overall question, how to link the data sets in applications will be answered in deliverable 3.4.2.

7. CONCLUSIONS

The three storyboards cover different areas of railway operation and data exchanges, all aiming at increased railway capacity. Some of the data formats currently in use already have planned development activities for the future, and of particular promise are EU-led efforts. Linking back to the data classification in public transport services shown in Figure 3-1 the following

Figure 7-1 illustrates which data formats, models and concepts may be used for the different data classes.

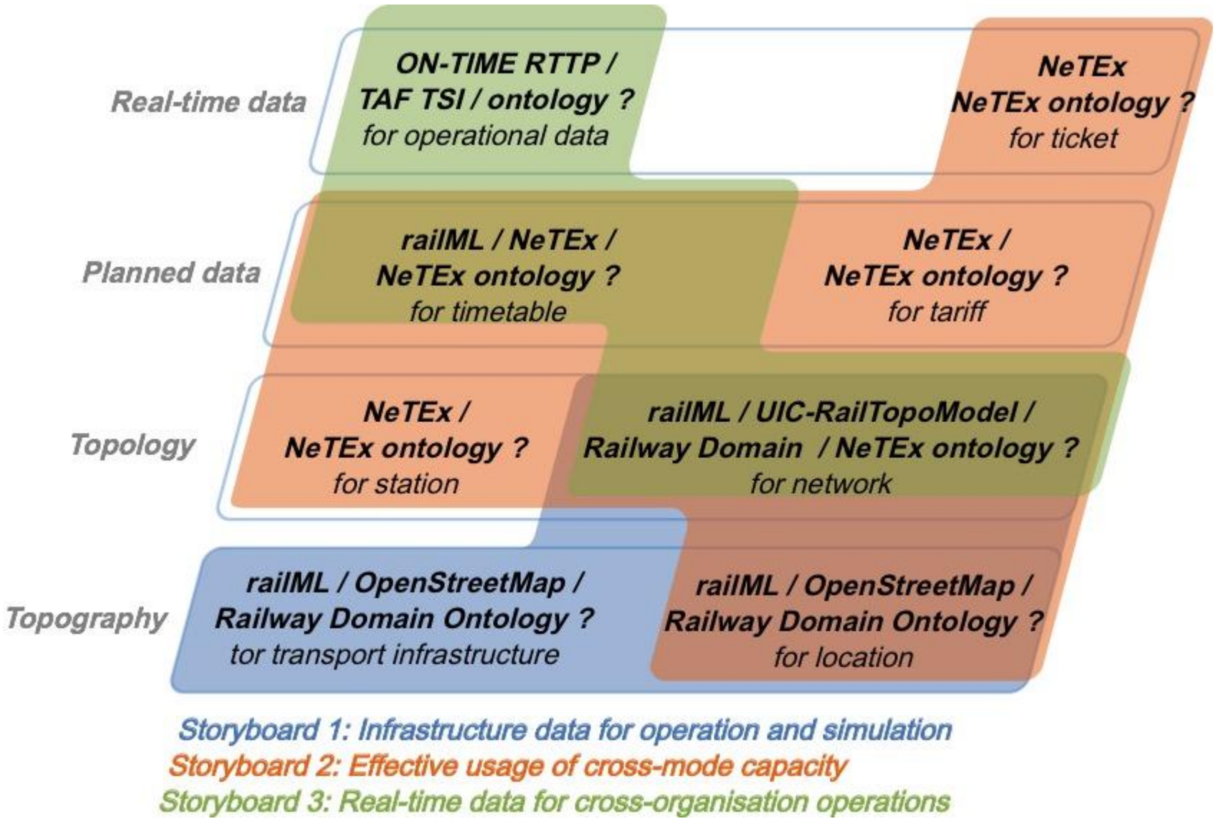

Figure 7-1 Data formats, models and concepts for public transport services (Jutand, 2015)

A key element in the success of the CAPACITY4RAIL project will lie in its handling of the relationship between rail and other transport modes, allowing timely connections and giving seamless access to alternative modes in the event of extreme disturbances.

All transport processes rely on a solid network topology, which correctly handles different levels of granularity. The UIC RailTopoModel will fulfil this task, bringing railML 3 as XML schema based data format as reference implementation.

Linked sensor networks, as researched in SP4 of CAPACITY4RAIL, serve as indicators for disturbed railway infrastructure and rolling stock, supporting both maintenance staff and those responsible for railway operations, who may react to events by adjusting the real-time traffic plan based on robust simulations taking the actual traffic into account.

Railway asset data may be handled by railML. Crowd-sourced OpenStreetMap data may feature in the systems landscape where there are data gaps or a degraded operating scenario. A symbiotic relationship may be established, where the OpenStreetMap community gathers data donations and the supplying IM gains steady data corrections.

Passenger information systems are best served by NeTEx, integrating navigation paths, various fare models, and provision for electronic ticketing.

Timetables may be provided in railML as well as in NeTEx, while railML focuses on the operation's point of view and NeTEx serves the passenger needs. Equal data may be converted between both XML formats, as both support different types of software tools.

Real-time data from the IM to RUs may be provided by an upgraded version of the RTTP-format, which resulted from the ON-TIME project. This XML format has to be aligned with railML 3 and the underlying UIC-RailTopoModel as well as with consolidated sensor data.

Regarding all single data aspects, equivalent ontologies provide facilities to enrich them to Open Linked Data sets, which in turn enable the technologies of the Semantic Web, e.g. ontology-based reasoning. While data formats and models are largely covered by EU specifications and UIC standards, the ontology topic is subject to individual studies for certain use cases.

Further research regarding data formats, models and concepts shall include:

- Interaction of IM asset data sets with OpenStreetMap data in a round-trip process;
- Upgrade of ON-TIME RTTP regarding railML 3 / UIC-RailTopoModel, and proposing it to the railML community;
- Incorporate the consolidated findings of SP4 on sensor data into the upgraded RTTP;
- Comparison of Schematron- and Ontology-based approaches for railway data verification;
- Development of ontologies supporting Linked open data from specific formats such as railML and NeTEx;
- Demonstrating the developed ontologies in typical use cases, oriented at the stories of this document.

The question, how the data sets in the proposed data formats and models shall interact in order to enable scenario-oriented software solutions, will be answered in the deliverable D3.4.2 in form of architecture recommendations.

REFERENCES

AWT Consortium. (2014). *All Ways Travelling - To develop and validate a European passenger transport information and booking system across transport modes (phase 1).* Final Report, Contract MOVE/C2/SER/2012 489/SI2.646722.

AWT Consortium. (2013-2015). *All Ways Travelling*. Retrieved May 19, 2015 from http://www.allwaystravelling.eu/

Berners-Lee, T. (2010). *Linked Data*. Abgerufen am 24. Aug 2015 von http://www.w3.org/DesignIssues/LinkedData.html

Braun, B. (2012). Detailreich, nützlich, for free – OpenStreetMap-Daten . *interAktiv* (1), S. 22.

CAPACITY4RAIL. (2013). *Annex I - "Description of Work", Part B1 Concept and objectives, progress beyond state-of-the-art, S/T methodology and work plan.* Grant Agreement no: 605650.

CEN/TC278/WG3 SG9. (2015). *NeTEx FAQ*. Retrieved June 09, 2015 from http://netex-cen.eu/?page_id=111

CEN/TS16614-1. (2014). *Public transport - Network and Timetable Exchange (NeTEx) - Part1: Public transport network topology exchange format (May 2014).* European Committee for Standardization.

CEN/TS16614-2. (2014). *Public transport - Network and Timetable Exchange (NeTEx) - Part 2: Public transport scheduled timetables exchange format (May 2014).* European Committee for Standardization.

CEN/TS16614-3. (2015). *Public transport – Network and Timetable Exchange (NeTEx) – Part 3: Public transport fares exchange format (Apr 2015).* European Committee for Standardization.

CER; ETTSA; ECTAA. (2013). *Taking rail ticket distribution to the next level: railways and ticket vendors launch the 'Full Service Model' initiative.*

Chandrasekaran, B., Josephson, J., & Benjamins, V. (1999). What are ontologies, and why do we need them? *IEEE Intelligent Systems* , pp. 14(1), 20-26.

Colpaert, P., & Byrd, A. (2015). *OpenTransport/linked-gtfs GitHub*. Abgerufen am 09. Sep 2015 von http://vocab.gtfs.org/

Commission Regulation (EU) 2015/302 of 25 February 2015 amending Regulation (EU) No 454/2011 on the technical specification for interoperability relating to the subsystem 'telematics applications for passenger services' of the trans-European rail system (February 25, 2015).

Corsar, D., Edwards, P., Baillie, C., Markovic, M., Papangelis, K., & Nelson, J. (2013). Short Paper: Citizen Sensing within a Real-Time Passenger Information System . *Proceedings of the 6th International Workshop on Semantic Sensor Networks , 1063*, S. 71-76.

Czioska, P., Thiemann, F., Giese, R., & Vogt, H. (2014). Ableitung eines routingfähigen Bahnnetzes aus nutzergenerierten Gleisdaten (OpenStreetMap) durch Generalisierung. *DGPF Tagungsband* (23).

Davis, I. (2012). *TRANSIT: A vocabulary for describing transit systems and routes*. Abgerufen am 24. Aug 2015 von http://vocab.org/transit/terms/.html

Dell'Arciprete, U. (2012). *Tariffs IT Specifications.*

EC, T. E. (2007). *DIRECTIVE 2007/2/EC establishing an Infrastructure for Spatial Information in the European Community (INSPIRE).*

EIM. (2014). *EU railway system conference 08/05/2014 EIM/ERA overview of RINF.* European Rail Infrastructure Managers.

ERA. (2010). *IU-Recommendation of RINF-final Report.* EUROPEAN RAILWAY AGENCY, Interoperability Unit.

ERA. (2015). *Register of Infrastructure (RINF)*. Retrieved Apr 30, 2015 from http://www.era.europa.eu/Core-Activities/Interoperability/Pages/RINF.aspx

ERA. (2014). TAF TSI - Where is my train? *Innotrans 2014.* Berlin.

ERA. (2015). *TAP TSI: ANNEX B.2.*

ERA Telematics Team. (2013). *TAF-TSI Master Plan.*

ERA Telematics Team. (2014). Telematics Application for Passenger TSI TAP. *railML conference Braunschweig March 26th, 2014.*

ERA Telematics Team. (2013). Telematics Applications for Passengers and Freight TAP/TAF TSI. *railML conference Paris, September 18th, 2013.*

ERIM Workgroup. (2014). *UIC RailTopoModel Railway Network Description - A conceptual model to describe a railway network (Version RC2).* Paris.

European Commision. (2011). *C(2011) 2962 final: COMMISSION REGULATION of 5.5.2011 on the technical specification for interoperability relating to the subsystem 'telematics applications for passenger services' of the trans-European rail system.*

European Commision. (2014). *COMMISSION REGULATION (EU) No 1305/2014 of 11 December 2014 on the technical specification for interoperability relating to the telematics applications for freight subsystem of the rail system in the European Union and repealing the Regulation (EC) No 62/.*

European Commission. (2011). *WHITE PAPER - Roadmap to a Single European Transport Area - Towards a competitive and resource efficient transport system Com(2011) 144 final.*

Fraunhofer IVI. (2014). *railML®-capable Applications, railML4RINF*. Retrieved Oct 16, 2014 from http://www.railml.org//index.php/applications.html?show=29

FSM Initative. (2014). *Full Service Model Initiative - Latest development.* Presentation.

Gely, L., Dessagne, G., Pesneau, P., & Vanderbeck, F. (2008). *A Multi Scalable Model Based On A Connexity Graph Representation.*

Google Developers. (2015, May 27). *Realtime Transit*. Retrieved June 17, 2015 from https://developers.google.com/transit/gtfs-realtime/

Google Developers. (2015). *Static Transit*. Retrieved Feb 2, 2015 from https://developers.google.com/transit/gtfs/

Google TransitDataFeed - PublicFeeds. (2015). Abgerufen am 24. Aug 2015 von https://code.google.com/p/googletransitdatafeed/wiki/PublicFeeds

Google TransitDataFeed - PublicFeedsNonGTFS. (2015). Abgerufen am 24. Aug 2015 von https://code.google.com/p/googletransitdatafeed/wiki/PublicFeedsNonGTFS

GTFS Data Exchange. (2015). Abgerufen am 24. Aug 2015 von http://www.gtfs-data-exchange.com/

IDMVU.org. (2010). Retrieved August 04, 2014 from http://www.idmvu.org

INSPIRE. (2014). *Infrastructure for Spatial Information in the European Community.* Retrieved July 30, 2014 from http://inspire.ec.europa.eu/

Joint CER/EIM working group on RINF. (2013, Dec 19). *Joint position paper by CER and EIM on the Register of Infrastructure - RINF - Request to adopt the principles of an industry-standardised data exchange format for RINF*. Retrieved Oct 06, 2014 from http://www.cer.be/index.php?eID=tx_nawsecuredl&u=0&file=uploads/media//CER_EIM_RINF_data_model_position_paper.pdf&t=1412675372&hash=ce6e4a26961b0452b821b2f78a92d1a1de1880e3

Jutand, F. (2015). *Ouverture des données de transport.* rapport remis au sécretaire d'État chargé des Transports, de la Mer et de la Pêche.

Keller, C., Brunk, S., & Schlegel, T. (2014). Introducing the Public Transport Domain to the Web of Data . *WISE 2014, Part II , LNCS 8787*, S. 521–530 .

Knowles, N. (2014). *Using the CEN NeTEx Standard to represent basic Rail fare data - Briefing Report for the UIC.* Abgerufen am 24. Aug 2015 von http://netex-cen.eu/wp-content/uploads/2014/12/20140610-UIC-NetEx-TAP-Report.pdf

Knowles, N., & Miller, P. (2008, Dec 29). *A Transmodel based XML schema for the Google Transit Feed Specification With a GTFS / Transmodel comparison.* Retrieved August 18, 2014 from http://www.dft.gov.uk/transmodel/schema/doc/GoogleTransit/TransmodelForGoogle-09.pdf

Kolmorgen, V. P. (2014). *2nd UIC RailTopoModel and railML® Conference - Session C : The usage of railML® for the RINF project of ERA.* Retrieved Oct 16, 2014 from http://documents.railml.org/events/slides/2014-04-08_uic-railtopomodel+railml-conference_railml4rinf.pdf

Langer, G., Lewis, R., & Roberts, C. (2008). Ontology driven railway RCM data integration. *4th IET International Conference on Railway Condition Monitoring* , S. 1-8.

Li, T., van Heck, E., Vervest, P., & Rooijmans, P. (2006). Improve Yield in Public Transport - a Focus on ICT Capability. *IEEE International Conference on Service Operations and Logistics, and Informatics, 2006 (SOLI'06).*, (pp. 516 - 521). Shanghai.

Lodemann, M., Luttenberger, N., & Schulz, E. (2013). Semantic Computing for Railway Infrastructure Verification. *8th IEEE International Conference on Semantic Computing* , S. 371-376.

MBTA. (2015). *Massachusetts Bay Transportation Authority - Real-Time Data in GTFS-realtime Format*. Retrieved June 17, 2015 from http://www.mbta.com/rider_tools/developers/default.asp?id=22393

Meyer, D. (2012). *Open transport data in Germany? Not if you're not Google*. Abgerufen am 24. Aug 2015 von https://gigaom.com/2012/09/28/open-transport-data-in-germany-not-if-youre-not-google/

Mishevska, E., Najdenov, B., Jovanovik, M., & Trajanov, D. (2014). Open Public Transport Data in Macedonia .

Mnif, S., Galoui, S., Elkosantini, S., Darmoul, S., & Ben Said, L. (2015). Ontology based performance evaluation of public transport systems. *4th IEEE International Conference on Advanced Logistics and Transport (ICALT)* , S. 205-210.

Nash, A., Huerlimann, D., Schütte, J., & Krauss, V. (2004). railML - a standard data interface for rail-road applications. *Computer in Railways IX (Comprail IX)* (pp. 233-240). WITpress.

NeTEx. (2014). *NeTEx - Network and Timetabling Exchange*. Retrieved June 5, 2015 from http://netex-cen.eu

Nissi, E., Jeanmaire, A., Seybold, B., & et.al. (2013). Rail TopoModel and railML® - The foundation for an universal Infrastructure Data Exchange Format. *24. railML.org conference (Paris, September 2013).*

OLTIS GROUP. (2014). Technical Specifications for Interoperability. *CAPACITY4RAIL MS 4* .

ON-TIME_WP04. (2014). *ONT-WP04-D-IFS-051-02 - Tools for reraltime perturbation management including human machine interface.*

Pease, A. (2011). *SIGMA Knowledge base Browser - SUMO*. From http://sigma.ontologyportal.org:4010/sigma/Browse.jsp?kb=SUMO

Plu, J. (2012). *NEPTUNE ontology*. Abgerufen am 24. Aug 2015 von http://data.lirmm.fr/ontologies/neptune

Plu, J., & Scharffe, F. (2012). Publishing and linking transport data on the Web. *WOD '12 Proceedings of the First International Workshop on Open Data* , S. 62-69.

prEN15531-1. (2013). *Public transport – Service interface for real-time information relating to public transport operations (SIRI)–Part 1: Context and framework; Oct 2013.* European Committee for Standardization.

PTV Group. (2015). *PTV Compass Blog, OSM*. Abgerufen am 30. Apr 2015 von http://compass.ptvgroup.com/tag/osm-en/?lang=en

Rahmig, C., & Richter, A. (2014). *Converting OpenStreetMap geo data into railML® for a Railway Simulation Environment.* railML.org conference, Paris.

railML.org. (2015). 4th UIC RailTopoModel and railML conference (Paris, April 2015).

railML.org. (2015). *railML TT:UseCase*. Retrieved July 29, 2015 from http://wiki.railml.org/index.php?title=TT:UseCases

railML.org. (2012). *railML.org- The Project*. Retrieved Apr 29, 2015 from http://www.railml.org//index.php/events2012.html

SBB. (2014). *From timetable to train operation.* Abgerufen am 17. Jun 2015 von https://www.sbb.ch/content/dam/sbb/en/pdf/en_mobile/Operations.pdf

Steiner, D., Hochmair, H., & Paulus, G. (2015). Quality Assessment of Open Realtime Data for Public Transportation in the Netherlands. *GI_Forum* , S. 579-588.

TAP TSI. (2012). *Telematics Applications for Passenger Services Technical Specifications for Interoperability (TAP TSI)*. Retrieved June 17, 2015 from http://tap-tsi.uic.org/

Tutcher, J. (2014). Ontology-driven Data Integration for Railway Asset Monitoring Applications. *2014 IEEE International Conference on Big Data* , S. 85-95.

Tutcher, J., Easton, J. M., & Roberts, C. (2015). Enabling Data Integration in the Rail Industry Using RDF and OWL - the RaCoOn Ontology. *ASCE-ASME Journal of Risk and Uncertainty in Engineering Systems, Part A: Civil Engineering* .

TWG-TN. (2014). *D2.8.I.7 Data Specification on Transport Networks – Technical (v3.2).* INSPIRE Thematic Working Group Transport Networks.

UIC. (2015). *TAF TSI Schema*. Abgerufen am 29. July 2015 von http://www.era.europa.eu/Document-Register/Documents/taf_cat_complete.xml

UIC. (2014, Nov 20). UIC RailTopoModel WORKGROUP MEETING. Paris.

VDV-Schrift 456 Standardschnittstelle Infrastruktur-Daten-Management (IDMVU) Version 3.0 (2014-05).

Wiki OpenRailwayMap.org. (2014). Retrieved August 04, 2014 from http://wiki.openstreetmap.org/wiki/OpenRailwayMap

Wiki OpenStreetMap.org. (2015). Retrieved Apr 30, 2015 from http://wiki.openstreetmap.org/wiki

Wiki OpenStreetMap.org Eisenbahn. (2015). Abgerufen am 30. Apr 2015 von http://wiki.openstreetmap.org/wiki/Eisenbahn

Wiki OpenStreetMap.org Tag:railway=station. (2015, April 30). From http://wiki.openstreetmap.org/wiki/Tag:railway%3Dstation

Wunsch, S. (2010). *Constraining railML with Schematron and Experiences with ETD.* Abgerufen am 24. Aug 2015 von http://www.railml.org/index.php/fruehere-veranstaltungen-2002.html?file=tl_files/railML.org/documents/events/slides/2010-03-17_wunsch-schematron.pdf