INTERNATIONAL UNION
OF RAILWAYS

FEASIBILITY STUDY

UIC RailTopoModel

and data exchange format

Zurich / Paris; September 2013

Table of Contents

Table of Figures

1 Foreword

In April 2013, the ERIM (*European Rail Infrastructure Masterplan*) Task Force of the UIC (International Union of Railways) launched a Request for Proposal to investigate whether "*an international infrastructure data model for railway topology and corresponding common data exchange format could be achieved*". The the feasibility study was entrusted to trafIT solutions, Zurich (Switzerland).

In this study, we analysed existing models, determined requirements for a standardised and universal data exchange scheme, identified work packages based on the existing railML exchange format and estimated the work load to establish a UIC RailTopoModel and corresponding exchange format based on railML.

The feasibility study provided a common focus point and objective.

The focus lay on working with the ERIM group and the railML consortium, both of which were very open for discussions. In four phases of one month each, we challenged both groups with ideas and proposals. At the end of each phase, we identified core points and important findings and strived to reach a common understanding for the next phase.

Our thanks go to the ERIM group, especially to Erika Nissi (UIC) and Alain Jean-maire (RFF), and equally to the railML consortium, especially to Vasco Paul Kolmorgen, Christian Rahmig and Susanne Wunsch for their open-mindedness.

While our work shows that it is technically feasible to establish a UIC RailTopoModel and a corresponding data exchange format supported by tools and an active user community, it remains to the ERIM group to find political consensus and financing for the necessary steps to realize this vision. We hope that they will be successful.

Bernhard Seybold & Burkhard Franke

September 27th, 2013

Contact: trafIT solutions gmbh
Heinrichstraße 48
8005 Zurich (Switzerland)

Web: http://www.trafit.ch
E-Mail: info@trafit.ch

UIC: Erika Nissi
E-Mail: nissi@uic.org

railML: Vasco Paul Kolmorgen
E-Mail: coordination@railml.org

2 Management Summary

The goal of this feasibility study is to investigate whether there is a path towards a common topological model and corresponding data exchange format, back-up by the UIC and, if possible, with a larger railway community. In the past, there have been many projects to build railway infrastructure models. Most of them were done within one national company and there is not much data exchange between them and if so it is a bilateral exchange.

In recent years, there were also various initiatives to create models on an international level. Some of them are driven by organizational bodies such as the EU and the ERA. Still, however, the holy grail of infrastructure has not been found yet.

The ERIM (*European Rail Infrastructure Masterplan*) working-group of the UIC has been working towards a common infrastructure master plan for several years. When they got in touch with railML, a community that maintains railway exchange formats for timetable, infrastructure and rolling stocks for ten years, both parties saw the potential for a fruitful collaboration towards a UIC RailTopoModel.

However, it was unclear whether the approach was technically feasible and if so at what cost. In order to answer those questions, this study was launched.

In this study, we focused on analysing the feasibility of a UIC-driven topological model to be used with members of the UIC.

In the first step, several existing models – both from national Infrastructure Managers and European directives were investigated. An evaluation structure was set up and all models were investigated against these criteria.

At the end of the model investigation, we came to the following conclusions:

- 95% of features in topological model are compatible, due to the fact that iron network is similar in every country.
- However (topological) models are often built for specific usages.
- Therefore a systemic (not depending on any usage) and scalable core model would the most appropriate.
- This core model would need to support data at different levels of detail (micro, meso, macro, corridor)

In the second step, requirements were formulated. The requirements were gathered in close collaboration with the ERIM group and the railML coordinators to whom we communicated for comments our analysis concerning the existing models and our own conclusions from previous works.

The overall 15 requirements have been structured in the following categories:

- Content requirements (5)
- Functional requirements (4)
- Structural requirements (3)
- Organizational requirements (3)

As the final step, it was investigated how well the current railML data definitions suited to fulfil the requirements. It was soon obvious that the railML user community has a lot of valuable experience suited well for building the foundation of the UIC RailTopoModel.

However, also some gaps were identified (see Chapter 8.1):

- The railML format has no clearly described model
- There is no “established” user support, only a community of users.
- There is no suitable extension mechanism leading to poor adoption and breaking adoptions of the standard
- There is no business plan

This gap analysis lead to the definitions of six work packages dealing with Model, Format, Tools, Organisation, Instructions and R&D.

The workload for these work packages was estimated to be around 100 man-months. In order to steer the project into the right direction, an iterative approach with 4 project phases and the appropriate project team was defined;

As a result of this study, here are our final recommendations summarized:

- The UIC RailTopoModel should be a **minimal core model** allowing national or functional extensions
- For interoperability, do not strive for a centralized database but for **standardisation of model and corresponding exchange format**
- Offer a model for railways who do not have yet their own model or who wish to improve their existing model(s)
- Realize UIC RailTopoModel and exchange format in a **phased approach** with concrete use cases. Upcoming projects gain a lot of efficiency by common standards
- Converge current stand-alone efforts into a combined effort

These recommendations aim at the following vision for the UIC RailTopoModel:

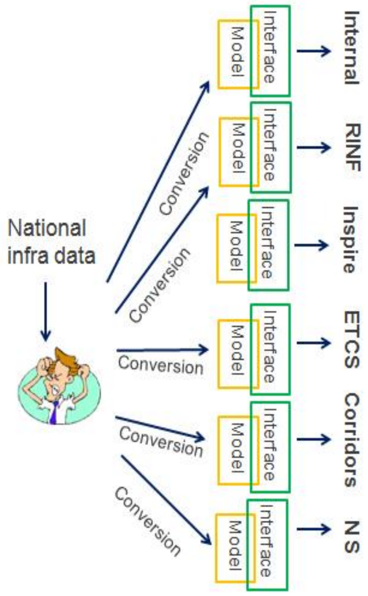
- The UIC RailTopoModel and corresponding data exchange format will be available for railways. However, they don't prevent from using the existing models and formats. Model and interface specification maintained as **open standard by railML consortium**, providing
    - documentation, tools, services, web presence
    - an active community (forum, meetings)
- The increased **interoperability** when exchanging infrastructure data, allowing to focus on data contents instead of formats. Standardised formats reduce data treatment costs and increase competition while reducing vendor lock-in.
- Efforts for infrastructure modelling and exchanging are **coordinated and centralized** leading to state-of-art models.
- **Adaptations** (extensions) happen in a coordinated and pre-defined way

3 Introduction

<u>The background</u>

This study was carried out on behalf of the ERIM (*European Rail Infrastructure Masterplan*) Task Force within UIC (International Union of Railways) from May to August 2013.

At the very beginning the ERIM Task Force members, all having data management responsibilities within their Companies, complained the fact that they were increasingly often **required to convert their national infrastructure data in different formats** to satisfy multiple business needs and legal obligations. They started to organise bilateral visits to understand how their neighbouring IMs organised their data management.



They learned that their individual Companies were all developing very similar solutions for their **central data repositories,** to be built on a **topological network description.** They exchanged experience and ideas on their topological data modelling works – a challenging issue as railway topology was to be declined in several levels of details to satisfy different business needs and processes.

And last but not least, the ERIM Task Force Members realised that they were all using or planning to use the **railML data exchange format.** This open source format was initiated in 2001 and over the years the railML users had defined specifications to exchange data in the fields of Infrastructure, Interlocking, Rolling Stock and Timetabling.

As these railML specifications (available at www.railml.org) have been developed on a voluntary / open source basis, they are not complete to fulfil all the (growing) needs for railway data exchange. Subsequently, several railways and ETCS suppliers have adopted the railML specifications as starting point and are currently developing, within their Companies and for their Companies, additional specifications on top of the railML® specifications.

In this context two of the ERIM Task Force Members realised their individual Companies had initiated comprehensive programs with their ETCS supplier (which appeared to be the same!) to complete the railML specifications for their bilateral ETCS data exchange. This was the last drop – the entire ERIM Task Force realised that they couldn't continue blindly wasting resources without reacting.

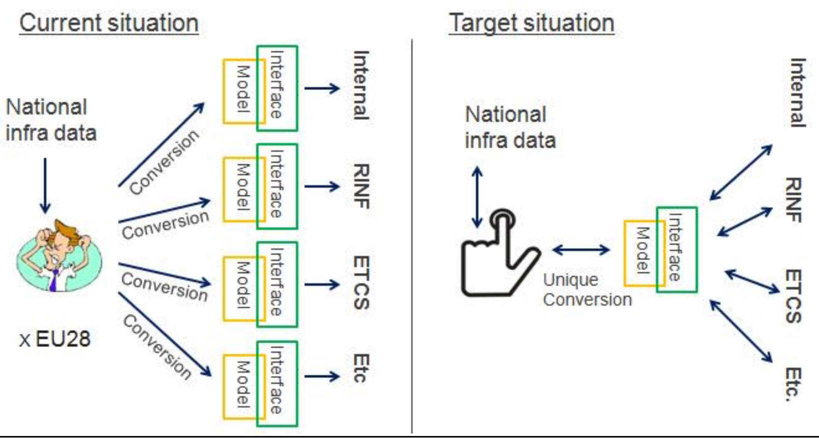
Indeed, they were all obliged to find solutions to comply with the same EU legislative requirements (RINF, Inspire….) and they were all developing rather similar solutions at national level. And most probably there were many other railways facing the same situation. This was the starting point to launch this feasibility study within the UIC ERIM activity.

The vision

Firstly, it was considered important to **build on the existing works** of EU / national data models and railML data exchange format. Indeed, the data model and the data exchange format are complementary and closely interconnected – their combination is needed for large scale data exchange.

Secondly, the Task Force wanted to establish whether a **universal data exchange scheme** would be feasible. In other words, the data exchange scheme would describe only the **characteristics** of the basic infrastructure elements (track, signal,..), independently of any end purposes and processes. On top of this systemic core model, additional data layers / modules could be built in order to satisfy specific end usages.

Thirdly, the data exchange scheme should be available for all potential users. The multiplication of the different data models and formats is, indeed, due to the absence of any **available and commonly recognised data scheme** which could be taken "from the shelf", if so desired, for network data **collection** (e.g. by EU) or data provision (e.g. from IMs to ETCS suppliers).



4 Methodology

In this study, the following steps were carried out:

1. Analyse the existing models (from national IMs and EU) and take into account the experience with the widely used exchange format railML.
2. Gather requirements that a UIC RailTopoModel and the corresponding data exchange format should fulfil
3. Analyse the gap between the requirements and railML.
4. Identify the necessary work packages based on the existing railML standard
5. Propose a roadmap and estimate its cost.

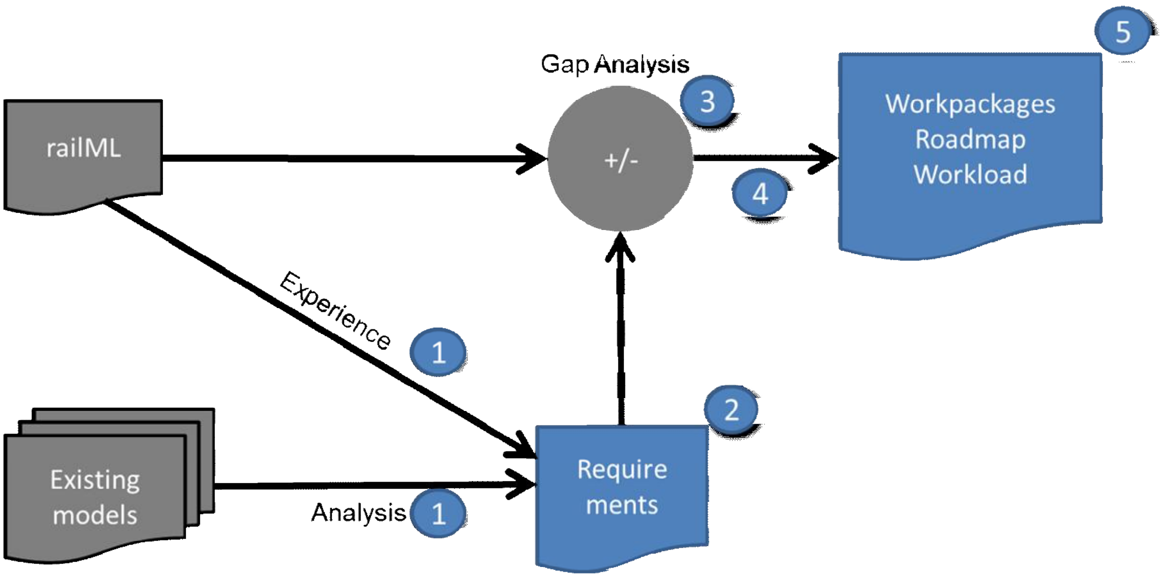


Figure 1 Methodology

The work was structured in four one-month phases. At the end of each phase, the intermediate results were presented to the ERIM group. This led to fruitful discussions and a common understanding which subsequently formed the foundation for the next phase.

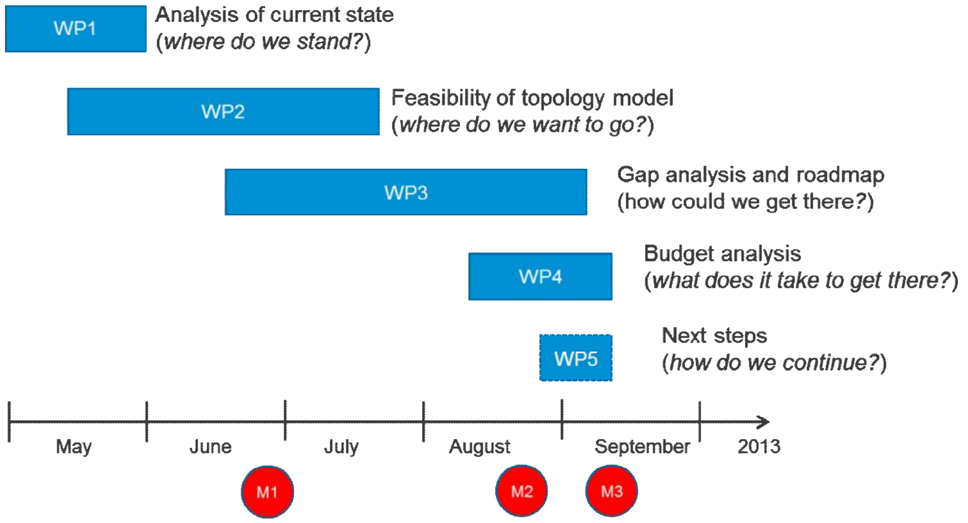


Figure 2 Timeline of the study

The final results have been presented on the 17th of September 2013 at UIC during the ERIM conference called "UIC RailTopoModel and railML – The foundation for an universal Infrastructure Data Exchange Format".

5 Use Cases

Interfaces \ Field	Technical	Operational	Legal
Internal **Between departments**	Standardised data exchange between technical departments (e.g. engineering + capacity allocation) often using different IT technologies and definitions ➔ synergy effect	Standardised data exchange between planning and monitoring of operations e.g. timetabling and real-time circulation tracking ➔ synergy effects	Improved monitoring of the network condition, via dedicated 'dashboards' providing network data summaries ➔ easier & faster data transfer and processing
National / Business **Between partners**	Standardised data exchange between IMs and their business partners, such as ETCS suppliers and maintenance sub-contractors ➔ savings in data production and transmission ➔ less vendor lock-in	Standardised data exchange between IMs and RUs (e.g. for track possessions). ➔ reduced operational costs Standardised intermodal communications ➔ enhanced railway market share	Ability of RUs to determine permissible train characteristics (esp. braking) on any infrastructure, as required by EU legislation (esp. TSI OPE) ➔ time savings, less errors Standardised data provision to national administrations such as land registers, regions, ministries. (Example: multiannual MS-IM contract as per 2012/34, art. 8 and 30) ➔ improved quality; scalable level of detail; improved credibility of rail
International **Between countries, organisations, EU**	Standardised data model / exchange on which ETCS-, IT- and other industries can design their products ➔ from taylor-made to inexpensive mass market solutions	Standardised data exchange within corridors and between organisations (RNE, ...) ➔ no need to develop multiple data conversion interfaces Information exchange concerning station accessibility ➔ contribution to TSI PRM objectives	Standardised / unique data provision to legal obligations; NS, RINF, Inspire, EU Freight corridors, TEN-T network ➔ Savings in data conversions and reduction of administrational burden

As base for the study, the following use cases were identified and structured in a general framework for this feasibility study.

6 Existing Models

Many different topological infrastructure data models and interfaces have been created over the years, either to fulfil national railway needs or to support EU directives. Indeed, in the absence of any commonly agreed standard for (international) data exchange each railway or EU initiative has been obliged to create its own data model and interface, often from scratches. Subsequently IMs are constantly requested to convert their data according to these different interfaces and data usages generating poor data quality and high data management costs.

In this study, several topological data models have been investigated to understand their converging and diverging points. This analysis provides the basic understanding of the current state-of-art and the feasibility of a common data model in the future.

The following models have been considered in a more detailed manner:

- RINF (ERA)
- Inspire (EU JRC)
- ARIANE (RFF, France)
- InfraNet (Infrabel, Belgium)
- Banedata (Jernbaneverket, Norway)
- RINM (Network Rail, United Kingdom)

In addition, knowledge of UNO, (SBB, Switzerland) and InfraAtlas, (ProRail, Netherlands) was taken into account together with the vast experience with infrastructure modelling in the railML consortium.

In the following part, we provide only snapshot illustrations of these comprehensive models. For a more thorough understanding, the reader is asked to contact the model owners for more detailed documentation.

6.1 Register of Infrastructure (RINF), EU directive

- Purpose: General description of the rail networks within EU 28. National Register Entities (NRE) are requested to submit quarterly rail infrastructure data to ERA.
- Interfaces: common xml interface (under construction)
- Supports routing at micro and macro level
- Member State dataset with validity period
- Use of Linear Reference System and GPS Coordinate System

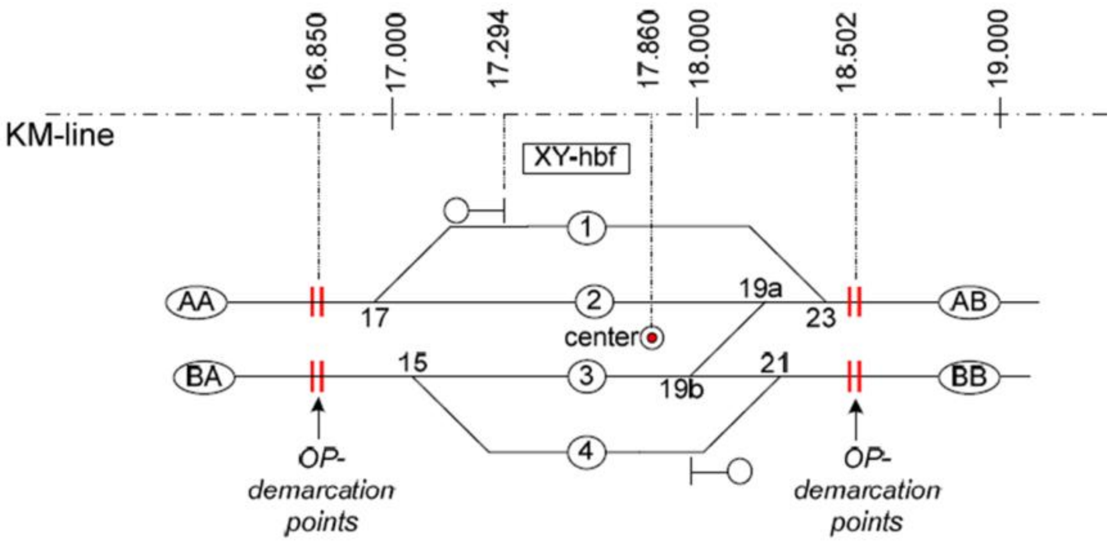


Figure 3 RINF: Modelling

[Source: RINF Key Notions, Concepts and definitions of the RINF CI- model, page 7]

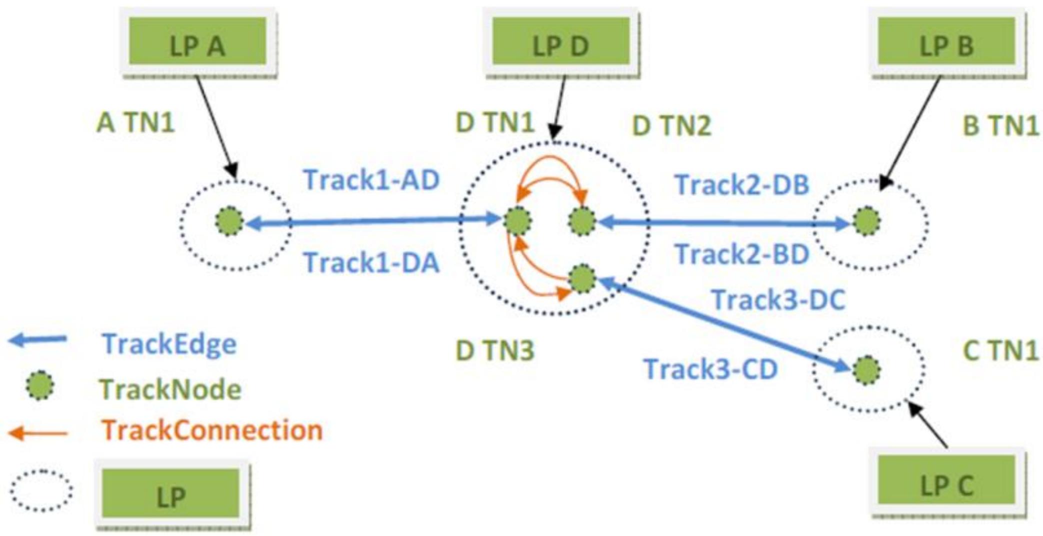


Figure 4 RINF: Track connection on micro level

[Source: RINF Conceptual and Implementation models, page 23]

6.2 INSPIRE, EU Directive

- Purpose: General description of 32 environmental related themes, including transports, within EU 28. The data Specifications for transports contain a thematic layer on rail transport networks.
- Interfaces: GIS based Geoportal
- Model: GIS-based, contains nodes and links. Node/link model can be interpreted as macro or micro level (flexible but also ambiguious).

Requirement 10	In a Transport Networks data set which contains nodes, these nodes shall only be present where Transport Links connect or end.

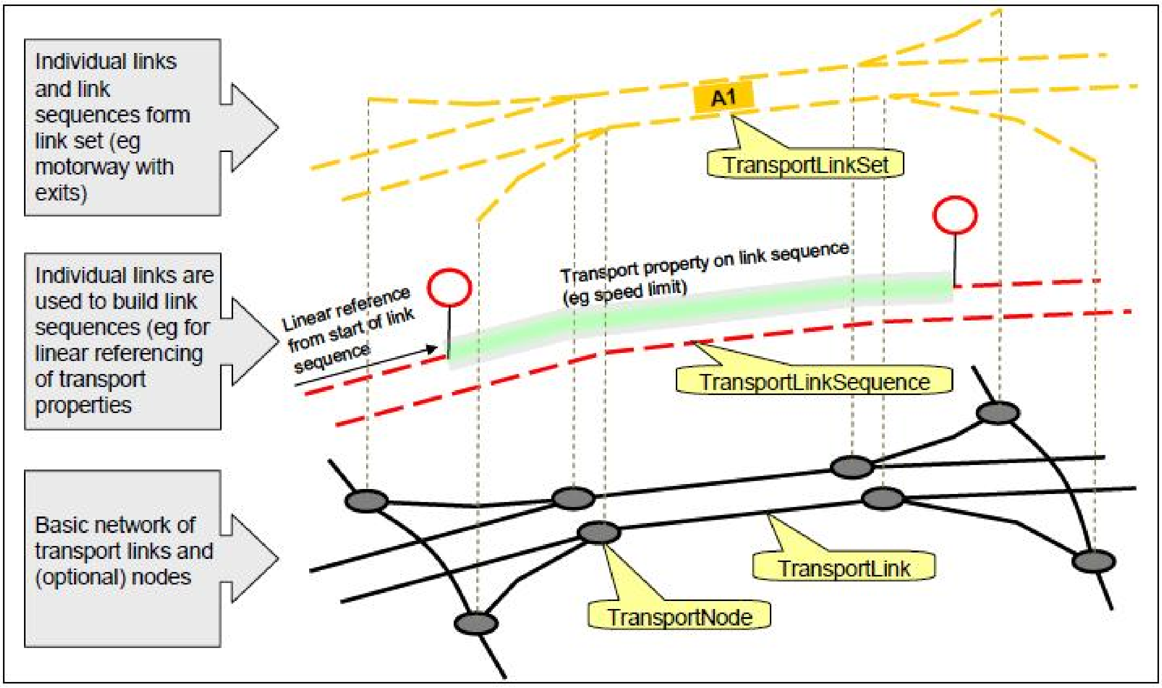


Figure 5 Inspire: Data specification on transport networks

[Source: Inspire data specifications on transport networks- Guidelines, page 24]

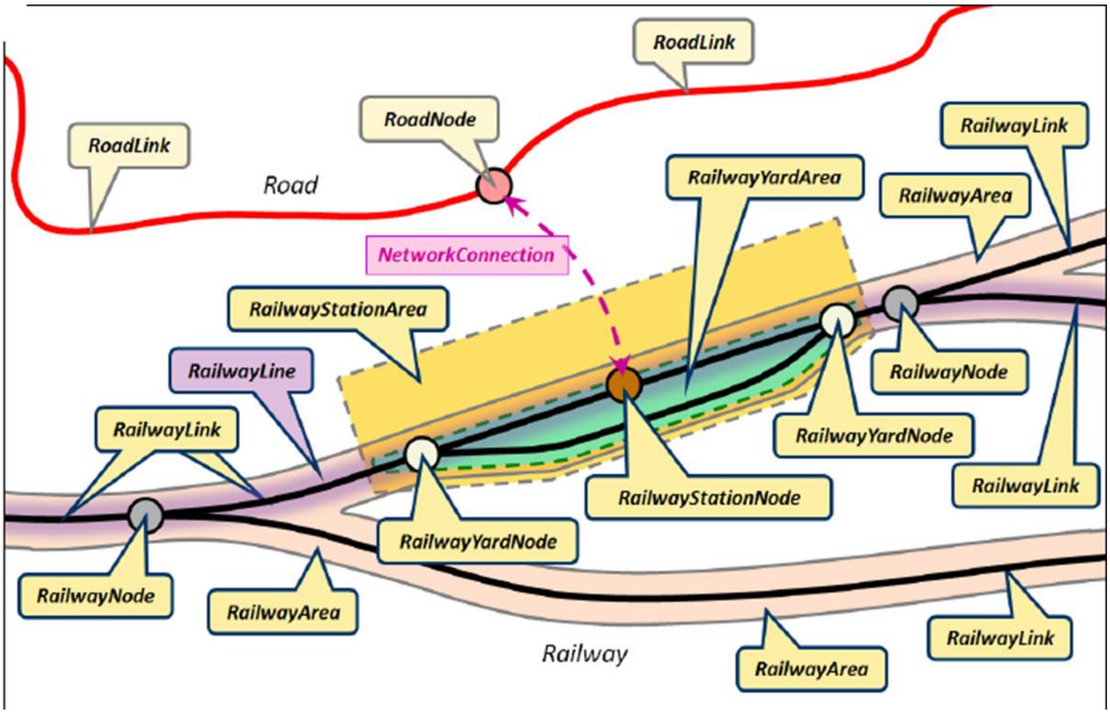


Figure 6 Inspire: Elements forming the rail transport network

[Source: Inspire data specifications on transport networks- Guidelines, page 70]

6.3 ARIANE, RFF, France

- Purpose: General network description
- Interfaces: text, json, xml
- ARIANE Model: Connectivity graph (dual graph)
  GAÏA Database: One unique common database for all french railway businesses and activities. Multilevel and aggregation (tracks, lines, corridors, ...), supports technical components and characteristics, physical paths and logical routes, includes natively multi-referencing (geo, linear) and geometry, time scales and business segmentations.

View of the new graph of the railway network (dual graph) :

- Each infrastructure asset (track section, signal, junctions,...) is shown as an endpoint to which the asset characteristics are linked
- The edges link the endpoints. They represent the connexity relations between the infrastructure assets

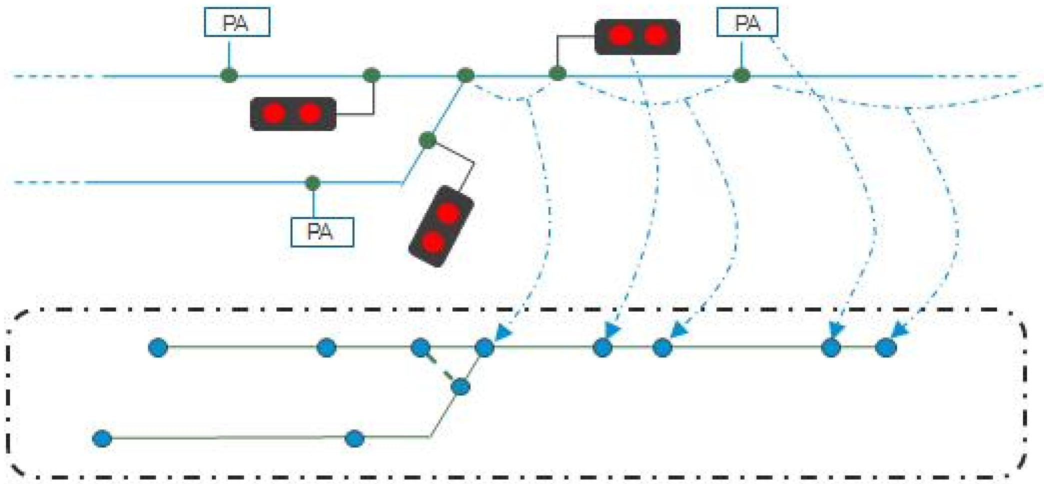


Figure 7 ARIANE: Dual graph model of the railway network

[Source RFF: Ariane, a new model for describing the railway system, slide 6]

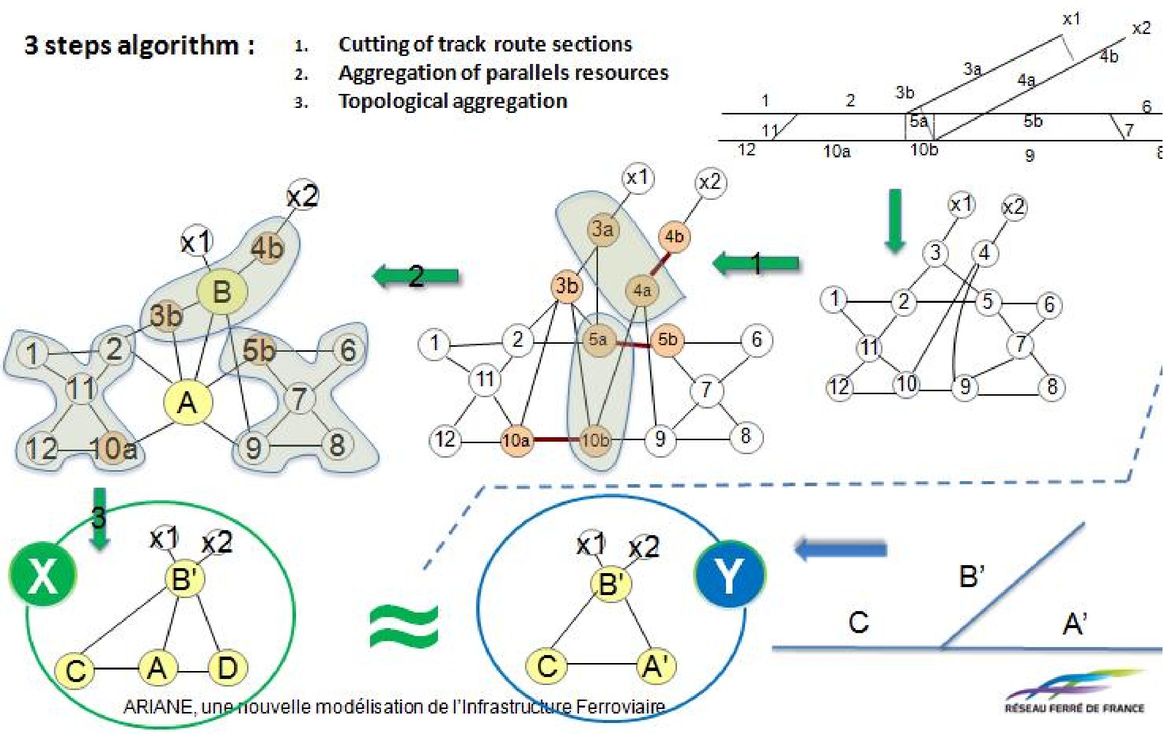


Figure 8 ARIANE: 3-step-algorithm from track to line

[Source RFF: ARIANE, Aggregation process From tracks to lines and Dense areas]

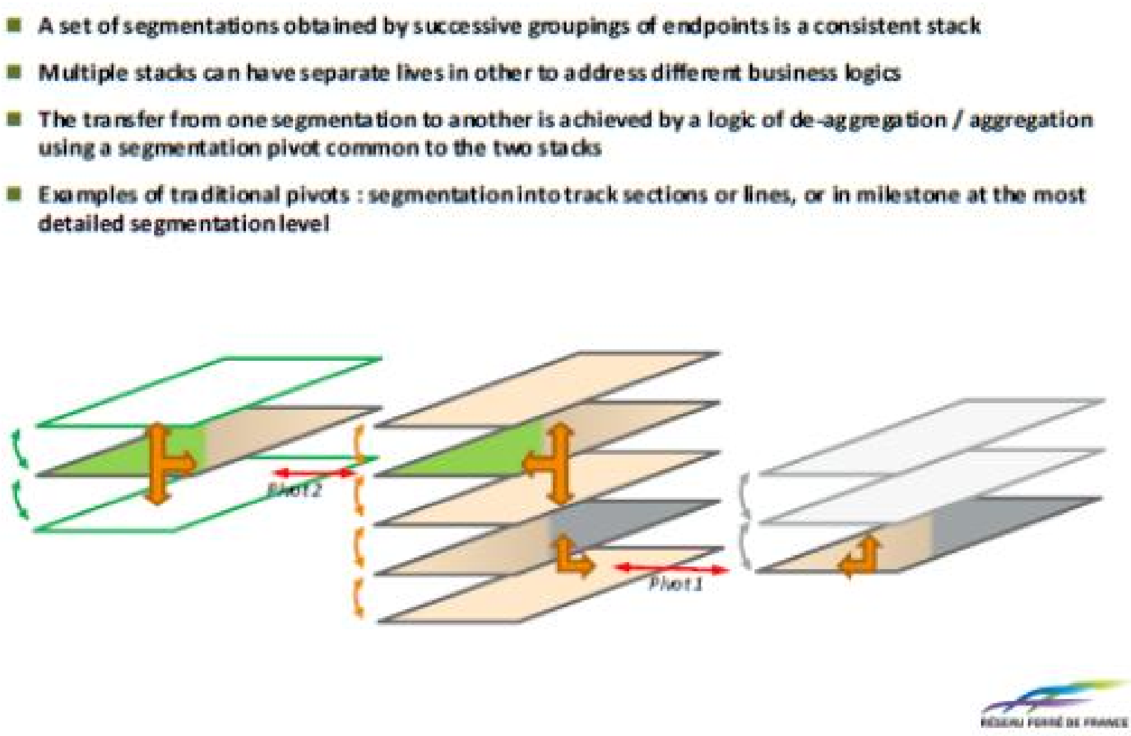


Figure 9 ARIANE: Correlation between different segmentation levels

[Source RFF: Ariane, a new model for describing the railway system, slide 9]

6.4 InfraNet, Infrabel, Belgium

- Purpose: General network description
- Interfaces: xml
- Specialty: topology graph with node, each node has a detailed graph describing the driveable paths and is connected to the outside via ports

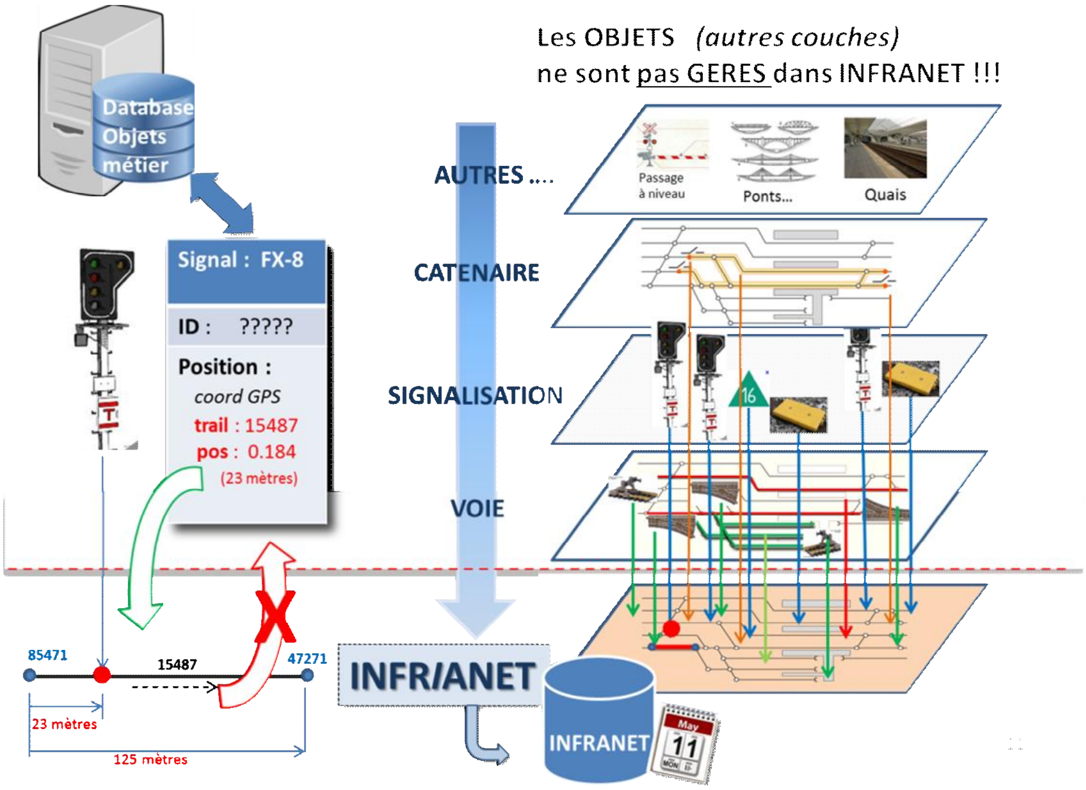


Figure 10 InfraNet: Different levels of details

[Source: Presentation "GIS, InfraNet, Georamses", Infrabel, slide 11]

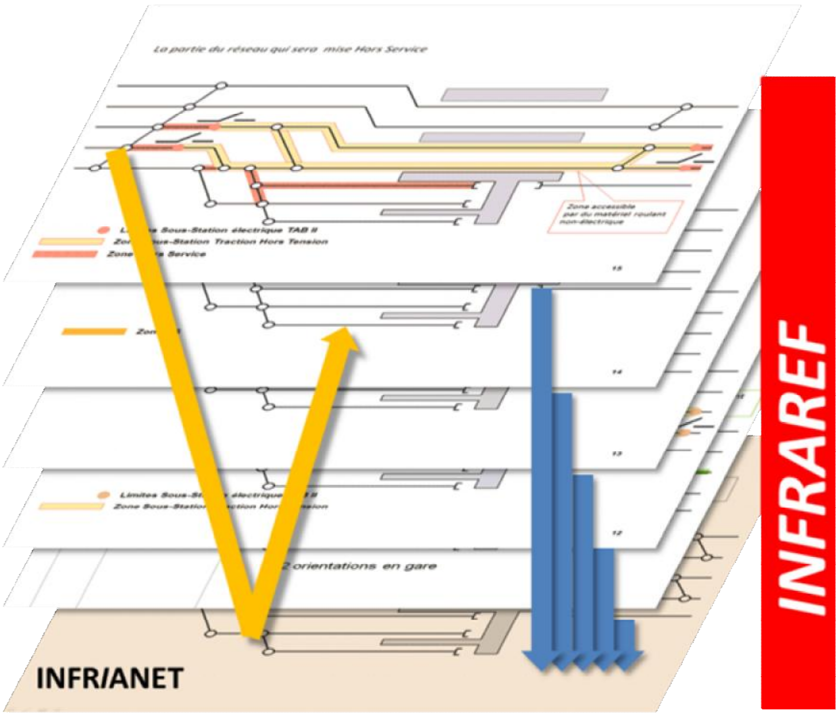


Figure 11 InfraNet: Vision transversale

[Source: Presentation “GIS, InfraNet, Georamses”, Infrabel, slide 15]

6.5 PPROD / EADB / ADB, ÖBB, Austria

- Purpose: General network description with focus on
  - PPROD: assets
  - EADB: signalling
  - ADB: data / radio networks
- Interfaces: DataBase to DataBase exchange
- Micro and macro level (integrated via special table)
- Central object is asset and its derivations (e.g. track, switch, sound protection walls)

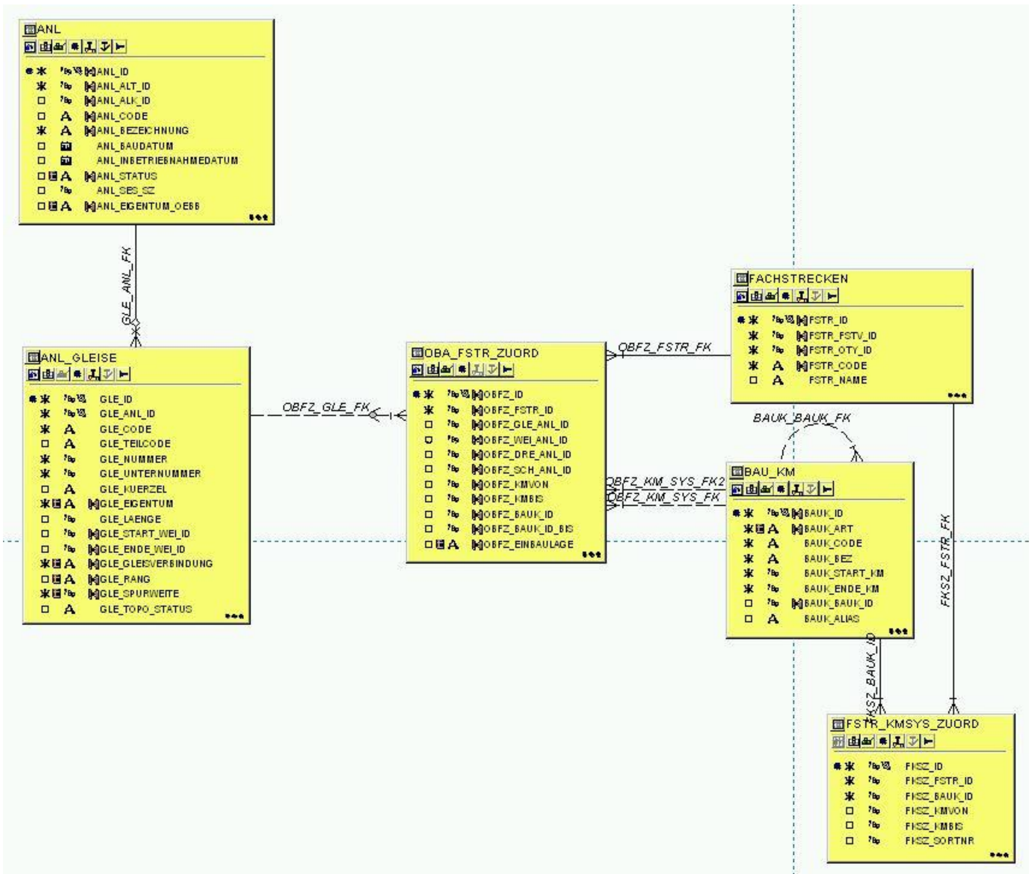


Figure 12 PPROD: Asset modelling

6.6 Banedata, Jernbaneverket, Norway

- Purpose: General network description and maintenance of infrastructure objects
- Microscopic level
- Interfaces: xml (railML), csv, xls
  railML interface is intended

- One common database containing information about all infrastructure objects (also in binary formats, e.g. drawings)

6.7 RINM, Network Rail, United Kingdom

- Purpose: General network description
- Currently under development
- Interfaces: xml and others (via FME)
- Network graph based on track-centreline at micro level. Macro level being designed.

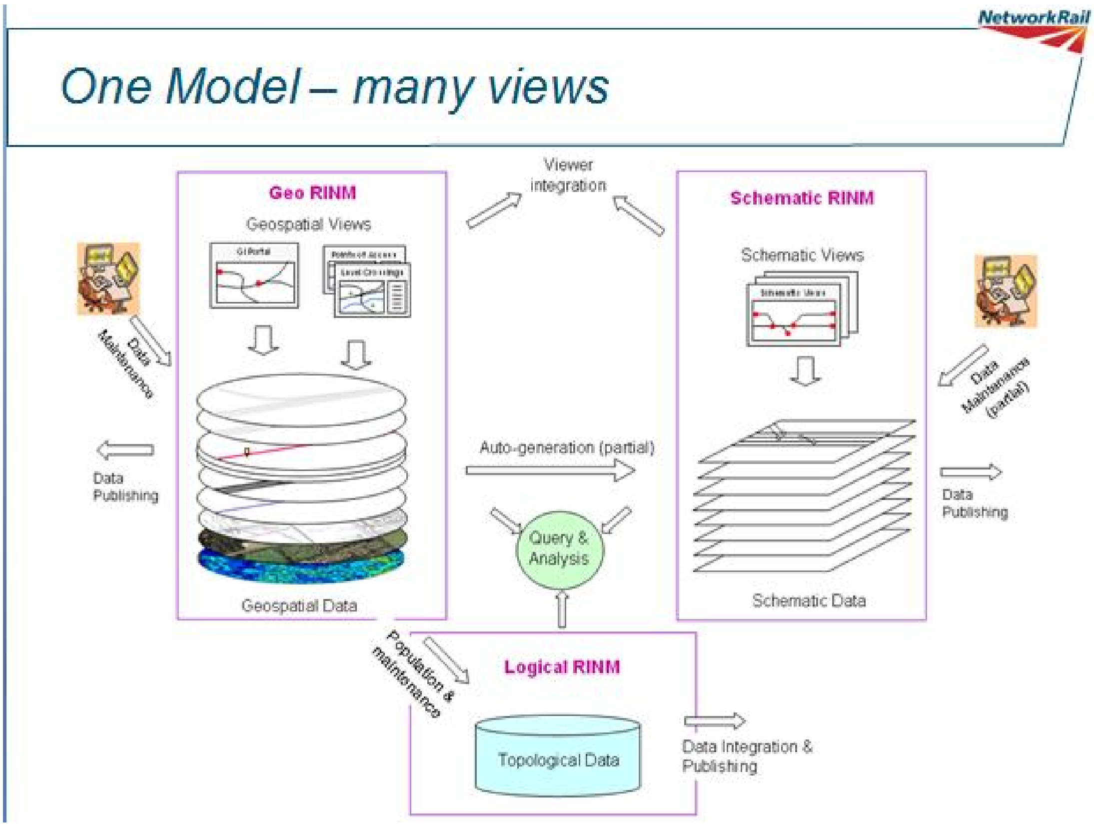


Figure 13 RINM: One model, many views

6.8 Observations

The analysis of the aforementioned models led to the following observations:

- **95% of features** in these topological models are **compatible**, as the iron network is basically similar in every country.
- However (topological) models are often build for specific use cases.

- Therefore a **systemic approach** and scalable **core model** are needed. To build the model and format**, precise requirements** need to be defined.

From this, the following conclusions were drawn:

- One unique model covering all aspects is not feasible
- **Core model** for iron network should be defined with an extension mechanism for predefined (common) extensions and (personalised) user specific extensions. See Figure 17 for extension mechanism.
- Topology should be the foundation of this core model
- Topological foundation should support the basic levels of detail, see Figure 14:
  - micro: detailed level (e.g. track geometry and signalling)
  - meso: track level (e.g. train dispatching)
  - macro: line level (e.g. timetabling)
  - corridor: international level (e.g. cartography, economical analysis)

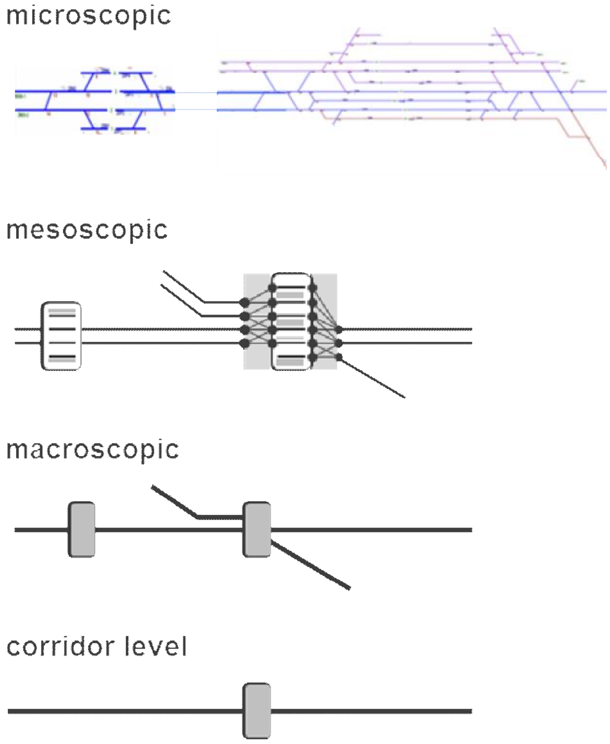


Figure 14: Levels of detail

- Data availability and precision may differ vastly between railways → The model should allow the user to choose what is the appropriate level of detail to start with or which levels he wants to use.

The aforementioned observations were based on an evaluation structure for each investigated model. During that process, however, we soon discovered that the information was almost trivially equal in the topology section (marked yellow in the table on the next page) and rather different and often incomparable in the other sections.

This confirmed our view that it is wise to concentrate on a core topology model. Further detailed analysis was stopped and we concentrated to formulate the requirements (see the next Chapter). The spreadsheet is hence incomplete but it is still given below as it served its purpose.

Model characteristics	Name of model(s)		ARIANE	INFRANET	Banedata	PPROD / EADB / ADB	RINM	RINF	INSPIRE	InfraAtlas	UNO
	Purpose		general network description		Register of Infrastructure and maintenance of asset	general network description	general network description	general network description	general network description	general network description	general network description
	In use since				2004	~1993 redesign with respect to GIS is planned	Currently being developed			~2002	~2006
Technical characteristics	Format (text, xml, database, …)		database	database	Oracle SQL-database	SQL database	Database			database	SQL tables + views
	Data exchange		text, json, xml	xml	XML, data exports (csv, xls)	typically not via xml but direct DB to DB exchange, LAND-XML for track geometries	XML & others (via FME)			text exports (csv)	
	Type of modelisation		connexity graph (i. e. dual graph at micro level		objects with attributes	central object is "Asset" and its specialisations	Network graph based on track-centreline at micro level. Macro level being designed.			graph with objects and tracks	
Topology elements	Integration of macroscopic and microscopic data in the model	C3	Native	yes	no only micro	micro, macro (integration via special table)	Micro level implemented, macro level being designed.			yes	???
	use of coordinates for location of objects?	C4	Native	yes	yes some objects, but not used for all	yes (via mile posts), objects reference via mile posts + mileage	Yes, coordinates are the master loction, but linear references also stored against many objects		yes	no	yes
	geometry information (for display)	C5	Native	yes	no	yes (polylines shapes for assets, GIS-geometry for lines only)	Yes.			no	???
Macroscopic objects and their properties	nodes	C1	yes	yes (OPs)	*lines as object in micro model*	yes	Being designed	#tracks,		yes	yes
	sections	C3	yes	not yet	*linesections as objects in micro model*	yes	Being designed	#tracks, gauge, length		yes	yes
Microscopic objects and their properties	track	C3	yes	yes	yes	yes	Yes			yes	yes
	joint	C3	yes	no	yes switches	yes	Yes			yes	yes
	signal	C3	yes	yes	Name, ID, position	yes (EADB)	Yes			yes	yes
Non-Topology elements	gradients	C3	yes	yes	Curvature points (asset)	yes	Yes			yes	yes
	curves	C3	yes	yes	Curvature points (asset)	yes	Yes			yes	yes
	speed limits / speed changes	C3	yes	yes	yes (are assets)	yes	Yes			yes	yes
	tunnels, bridges	C3	yes	yes	yes (are assets)	yes (are assets)	Yes			yes	
Features	Connectivity of topology of network		Native	yes, port notion	no, not in Banedata, but railML 1.0 model based on extract from banedata, ambitions to supplement this in Banedata	yes	Nodes (switches) connected by links (track centreline)	Network built of nodes and trails (inbetween nodes = edge)		basic switch with L, R	
	Characterisation of feasible movement at switch	C2	Native by typing edge = connection	yes, detailed movement in each node	no, not in Banedata, but railML 1.0 model based on extract from banedata, ambitions to supplement this in Banedata	yes	Future development			basic switch with L, R	
	Management of aggregation / disaggregation	F3	Native with composite pattern applied to different instance of graph (contraction/split of vertex	no	no	is not a requirement	Future development			no (borders as terra incognita without id of neighbor)	
	Transverse analysis of segmentation between different level of instance of network	C2/C3	Native with composite pattern	no	no		Future development				
Functional	Unique references	F1	yes	yes	yes	yes	Being designed			yes	
	Validities, versions, variants	F2	yes	yes	To some extent	no	Being designed			externally	
	Validations	F4	in progress	yes	To some extent	no	Being designed			no	
Structure	Modules, layers	S1	native		native	extension by adding new tables referencing old ones	New layers/tables can be added.			implicitly (can add new table for new information)	
	Normalization, univocal, stability	S2	object model, univocal, upgradeable		yes ?	primarily the current situation is contained	Yes, but under development.			on database level	
	Standards	S3	national units depending on attribute (km/h km, m)		national units depending on attribute (km/h km, m), ongoing work with defining standards	national units depending on attribute (km/h km, m), LAND-XML	national units but also support legacy linear referencing based on mileposts			no externals (national units, eg cm as height)	national units depending on attribute (km/h km, m)
Specialities	Remarks			topology graph with node, each node has a detailed graph describing the driveable paths and is connected to the outside via ports	One common database containing information about all infrastructure assets	PPROD: assets EADB: signalling ADB: data / radio networks					

7 Requirements

Before defining the specific requirements for the topological model or for the data exchange format, it is important to make a clear distinction between model and format and how they work together.

- A **model** defines (in UML) how to describe infrastructure objects and their attributes in different topological representations.
- A **format** is (one of many possible) representations of model objects, typically in text-format. , for exchanging model objects.
- Several models could use the same exchange format to share data, provided that there is an **adapter.**

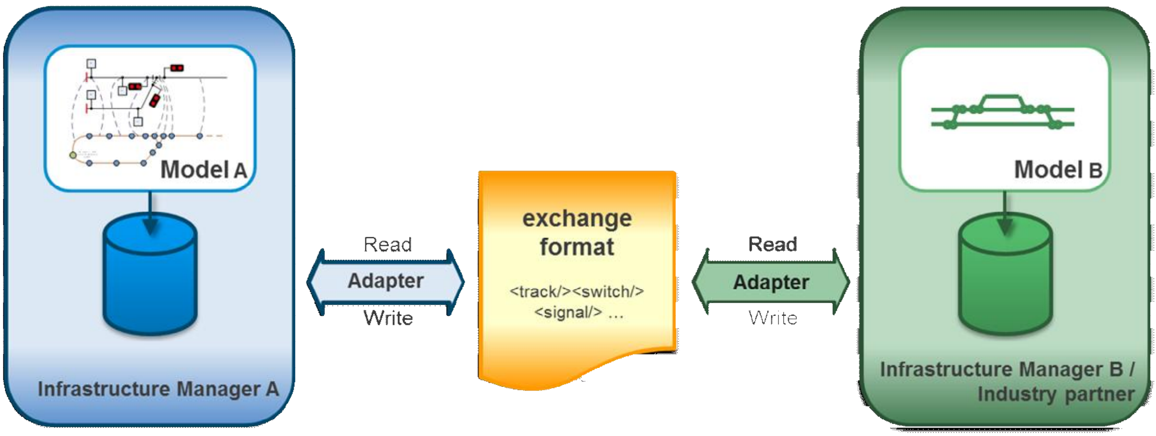


Figure 15 Model, format, adapter

The requirements for the model [M] and format [F] need to be completely independent from any end-usage of data or existing interfaces / tools. The defined requirements will help the railway community to **build up the universal scheme** for data exchange and **evaluate existing models and formats** against them.

The requirements are structured in four groups, as presented in the following sections. In addition, we also discussed and decided some **non-requirements** and they are reported in this study to inform that these items were taken into account in our analysis.

7.1 Content Requirements

7.1.1 C1: Contains topology [M]

- The model should support the **logical representation** of the iron network in a graph format (nodes and edges).

7.1.2 C2: Contains driveable paths [M]

- The model should support **path descriptions** and train routing

7.1.3 C3: Integrates micro, meso and macro and corridor topologies [M]

- The model should be one and unique whilst integrating **network topologies at several levels of detail**. These levels need to be interconnected and the **aggregation** and **disaggregation** between them should be feasible within this unique model.

7.1.4 C4: Contains reference systems [M]

The module containing the reference systems should be optional and refer to the topology. The topology itself is not aware of reference systems. Hence, objects of reference system refer to topology objects and not vice versa.

Several reference systems need to be supported simultaneously:

- Geo-coordinates and national projections in GIS environment
- Linear referencing systems (called mileage or km-points) for each individual line
  - Mileage systems are historical reference systems
  - May have more than one per country / IM
  - Must support jumps in mileage
- Rail addresses (buildings, tunnels…)
- Screen coordinates

s

7.1.5 C5: Contains geometries [M]

- Model exact shape of entities with geometries
- The model should support the **mathematical descriptions of geometry** and shapes of railway entities, such as transition curves (clothoides)
- Support 0-dim, 1-dim, 2-dim shapes for larger (external) entities (station, region, IMs, etc.)

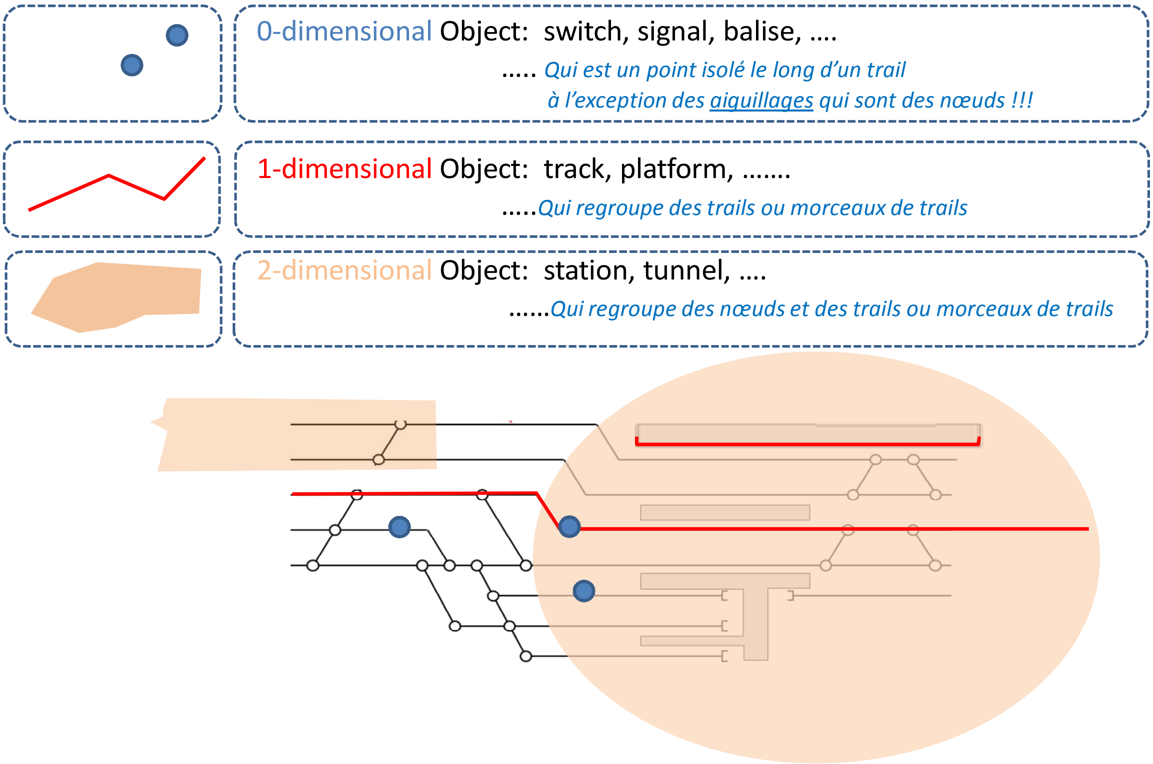


Figure 16: Shapes [Source: Infrabel]

7.2 Functional Requirements

7.2.1 F1: The format should allow the objects to be uniquely referenced [F]

- The format should support identities (surrogate keys) and allow references from outside
- The format should define the scope of **uniqueness**
  - In terms of time: how long should the same object have the same id (for lifetime of object or only for one exchange of data)
  - In terms of extent: file-wide, IM-wide nation-wide
- The format should support **extensions** (see S1) dealing with 0-dim, 1-dim or 2-dimentional structures within the topology.
  - Express locations references and location points (0-dim)
  - Define reference, for instance in 1-dim with trail and offset

7.2.2 F2: The model should support validities, variants and versions [M]

- validities: when an object is in operation / active / usable (and when not)
- variants: alternative states of model for the same time horizon
- versions: model states evolving over time (different versions, such as 1.1, 1.2)

7.2.3 F3: Supports partitions and unions [M]

- The model should allow the **division** in parts and the **reunification** of these parts again
- The model should allow the definition of **borders, interfaces, identifiers**
- The model should support different **scopes** of time and extent (see F1)
- The model should allow the creation of **artificial border nodes** where national models and data need to be split or combined (e.g. for RINF)
- The model should support **connector mechanism**, adding pieces of infrastructure using references

7.2.4 F4: Validations [F]

- The format should support **syntactical and semantic correctness**. Ideally from syntactic towards semantic correctness.
- Syntactical correctness can be checked easily when using xml.
- The model should imply as much semantic correctness as possible
    - Use of enumerations instead of strings for context info
    - Large variety of independently optional attributes should be avoided. Instead, sets of alternative attribute combinations should be defined.
    - Information for specific, additional purposes is added via extensions
- The **completeness** of the format is dependent on each use case. Therefore it would be helpful to define **specific profiles** for recurrent use cases such as "network statement", "running time calculation". The same could apply for some support tools.

7.3 Structural Requirements

7.3.1 S1: The model should support **extensions** for customer modules and layers [M]

- The model should contain a common **kernel** with clear semantics.
- The core model should allow the creation of **customer extensions** around the kernel for modules / layers that are not needed by all users.
- The model should allow the customer models / extensions to **reference the common core model.** However, the core model MUST NOT reference the customer extensions.
- The model should allow IT applications to select and read only those (core) layers that they have been programmed for, without having to deal with the rest of the data.

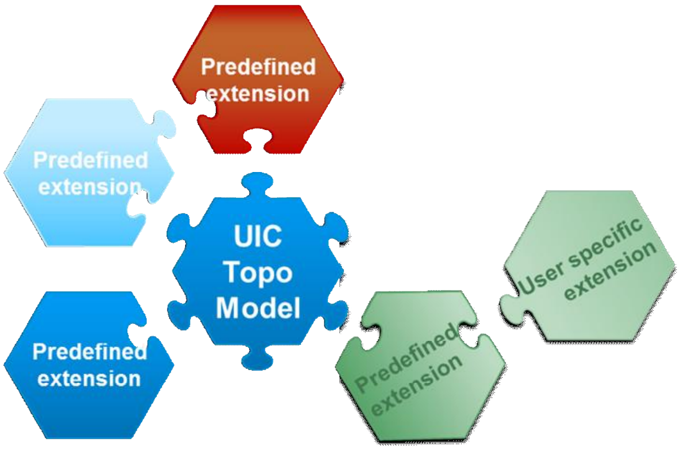


Figure 17 Extension mechanism

7.3.2 S2: The model should ensure the normalization and stability of network description [F]

- Network description needs to be strictly **normalized** to allow better comparisons
- Model should ensure 100% **univocal** description of network and individual items, leaving no room for interpretation
- Identity of items should be stable over lifespan of the model since these **ids are used also outside** of the model (→see F1).
- Versioning of model and format should not jeopardise the backward and forward **compatibility**.
- An idea for future is to build a **journaling format** including information on past and future modifications.

7.3.3 S3: The model should use existing standards whenever possible [M]

- The use of existing **libraries / tools for standards** will prevent from mistakes and oversights. This will facilitate also cross analysis and –referencing with other data sources.
- To mention some of these standards: Dublin Core for metadata, ISO units, UIC codes, design practices (e.g. workflow status)
- Support the **base data creation and reuse,** such as geographical information on station center points or - areas.

7.4 Organizational Requirements

7.4.1 O1: Endorse the use of open source standards

- Establish the topological model and data exchange format as open standard
- Encourage railway community to **contribute and adopt** this standard to facilitate further works and to build on existing works (not to invent the wheel...)
- Use the standard as a platform for the user community to exchange experience and **best practices.**
- Ensure that the **documentation** on the standard and support tools is easily available and understandable.
- Ensure the **independency** of vendors and national specificities.

7.4.2 O2: Enforce the compliant use of the Standard

- Prevent the dilution of the standard by **forbidding local dialects** under the "brand name" of the standard. In other words control that dialects don't claim using the standards if the core model features have been changed , see S1 Extensions)
- Exercise quality control through **certification process**

7.4.3 O3: Support Common Conventions

- In addition of the use of existing standards (see S3), the topological model and data exchange format could **support the creation of new conventions** (e.g. for namings) and definitions (e.g. for topological entities, such as "track" which is currently used for different meanings).

7.5 Non (or semi) Requirements

The following issues were discussed at length have been taken into account when defining the recommendations.

7.5.1 NR1: Human Readability

- The description of the topological model should be human-readable
- The data exchange format is designed for machine-to-machine communication and does not have to be human readable and understandable.
- However, standardized visualization tools will be provided to:
  - Check the validity and correctness of data before exchanging it
  - Visualize topological model to understand the content of the exchange files

7.5.2 NR2: Trade-Offs

- The model and format need to be designed without any consideration to the **size of resulting exchange files**. Thus savings in bandwidth or disk space should not be an issue!
- The model should not be designed in respect of the **calculation efficiency**. Thus it is up to performance-critical tools to convert the model structure into an optimized structure for them.

7.5.3 NR3: Data integrity

- Data integrity (completeness) is handled by using standard mechanisms (encryption, signatures, hashes) outside of the model and format
- Redundancy is to be avoided

8 Roadmap

8.1 Gap Analysis: the current situation

- Most currently used models are **not systemic / generic**, but rather have evolved over time for special purposes and processes and / or for national legacy systems.
- However, **complementary modelling approaches** are being developed at present by several actors, providing suitable "building blocks".
- **railML® provides consolidated base** for an exchange format
  - Indeed, railML benefits from large user community with solid IT experience, existing documentation and finally complementary implementation areas including definitions for Infrastructure, Interlocking, Rolling Stock and Timetables.
  - However, topological model and support tools which would facilitate the railML implementation are still missing.

8.2 Recommended Approach

- Use the **most advanced** modelling practices from national and EU models
- Concentrate the international data exchange scheme only on **few, but integrated representations of topology** (micro, meso, macro, corridor levels)
- Define an univocal, stable and up-to-standards **core model**
- Build a solid **extension mechanism** for coordinated customer models
- Take benefit from **railML format** and organization
- Bundle stand-alone efforts into **coordinated development** and working group

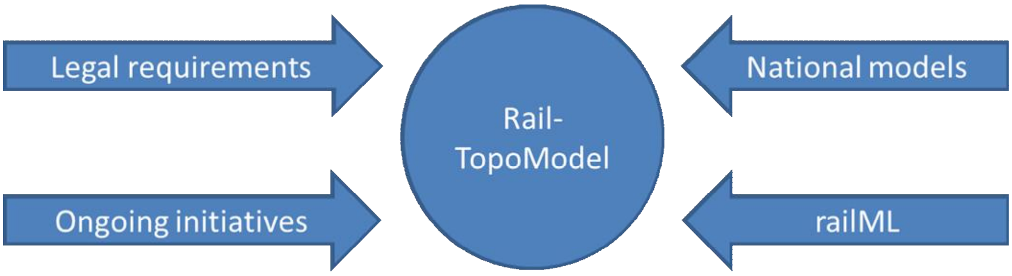


Figure 18 Recommended approach

8.3 Work phases

The work packages are structured into four phases, starting with the basic principles and ending up with the refinements of the complete data exchange scheme. The validation of each development phase by a concrete use case would allow a **stepwise consolidation before entering into the next phase.** At the time of writing this report, two possible test cases were identified, namely data provision to RINF and data exchange between IMs and ETCS suppliers.

- Phase 0: Concept and Basic Principles
  - Get a clear view of the model and design principles
  - Create a model draft as basis for cases, initiatives, third parties
  - Initiate new organization and start lobbying work for the Model
  - Duration: 3 calendar months, Workload: 8 man-months
- Phase 1: Base Components
  - Build the essential components of model / format (for iron network) and first tools (validation, viewer)
  - Design extension concept
  - Duration: 9 calendar months, Workload: 24 man-months
- Phase 2: Completion
  - Complete and finalize model, format and remaining tools (topology editor and migration tools)
  - Duration: 9 calendar months, Workload: 40 man-months
- Phase 3: Refinements
  - Custom extensions
  - Further work packages identified in earlier phase

Phase / Test Case	**Quarter**	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12
Phase	**Concept and Basic Principles**			Duration :			3 months						
0	Define the model and design principles			Workload:			8 man months						
	Create draft model for testing												
	Initiate new organisation and lobbying												
Phase	**Base Components**						Duration :			3 months			
1	Build essential component of model / format / tools						Workload:			8 man months			
	Design extension concept												
Test case	**RINF-compliance** (RINF)					*Additional workload ?*							
A	Define a model and data exchange for RINF												
Phase	**Completion**												
2	Finalise model / fomat / tools	Duration :			9 months								
		Workload:			40 man months								
Test case	**ETCS + Interlocking**												
B	Apply railML format for real ETCS project								*Additional workload ?*				
Phase	**Refinement**	*Additional workload ?*											
3													

8.4 Project Team

The foreseen project set-up is organized around a small core group doing the main work which is supported by several groups.

- Core team
  - 4 persons working on full time basis (50% - 100%) and ideally located in same office(s)
- Experts
  - 10 - 15 people
  - Meetings every 2 - 4 weeks for feedback and questions
- Technical steering committee(s)
  - 4 - 5 managers
- Advisory Group
  - Decision makers from stakeholders
- Interest group
  - Infrastructure managers, railways, public
  - One or two conferences per phase

8.5 Workload estimation

The estimation for workload is based on the following assumptions:

- The estimation is based on the current **version 2.2. of railML**
- Only the workload for the **core team is included**. It comprehends the actual IT development of each work package and coordination with possible external working efforts with universities, consultants, …
- Workload for steering committee, the expert group (outside of core team) and other supporting bodies is not taken into account.
- The workload for later phases is just a rough estimation which should be refined after the first phases.

The estimated workload for each Work Packages is given in the next section, adding up to a total of workload of 105 man-months.

8.6 Work Packages

8.6.1 WP 1: Model

- Create model (UML model), extension concept, documentation, examples

- WP 1.1 UML class model for the topological model
- WP 1.2 Concept extension mechanism
- WP 1.3 Documentation and sample data
- WP 1.4 UIC leaflet

8.6.2 WP 2: Format

- Complete (or redefine) the current railML specifications (version 2.2 → v. 3.0):

- WP 2.1 Complete definitions for topological iron network.
- WP 2.2 Complete the railML specifications for reference systems (geo-coordinates, mileages)
- WP 2.3 Create definitions for geometry (new in railML)
- WP 2.4 Build extensions for remaining infrastructure
- WP 2.5 Produce documentation, tutorials and sample data
- WP 2.6 Provide base data
- WP 2.7 Provide libraries and sample codes

8.6.3 WP 3: Tools

Build support tool for model and format
- WP 3.1 Validation (testing data correctness based on profiles)
- WP 3.2 Topology visualization (viewer)
- WP 3.3 Topology editor (based on viewer)
- WP 3.4 Migration tools (railML 2.2 -> 3.0)

8.6.4 WP 4: Organization

- Build an international organization and long-term business plan based on open standard

8.6.5 WP 5: Instructions

- Support users in creating the necessary data
- Support users in maintaining the format
- Support in writing adapters between existing models and the exchange format

8.6.6 WP 6: R&D (optional)

- Additional research and development work can be set up with universities or be connected with research programs in the field of transports, for instance to improve intermodal connections.

Work-	WP-Nr	Description	Resp.	Phase 0 Q1	Phase 1 Q2-Q4	Phase 2 Q4-Q9	Phase 3 Q7-Q12	Total Q1-Q12
		WP1: Rail-topo-model	**Respon-sible**	**Phase 0 Q1**	**Phase 1 Q2-Q4**	**Phase 2 Q4-Q9**	**Phase 3 Q7-Q12**	**Total Q1-Q12**
	1.1	**UML-class model** of the UIC-topo-model	Core	2	1	2	1	6
	1.2	Concept **extension mechanism**	Core	2	2	1		5
	1.3	**Documentation** and sample data	Core	1		1	1	3
	1.4	**UIC leaflet**	UIC			2	2	4
		Total		**5**	**3**	**6**	**4**	**18**
		WP2: Exchange format (railML 3)						
	2.1	Build **format for iron** network	Core	0.5	2	0.5	0.5	3.5
	2.2	Create **reference system** (coordinates, mileages)	Core		2	1	1	4
	2.3	Create **geometry** (new in railML)	Core			2	2	4
	2.4	Build **extensions** for remaining infrastructure	Core	0.5	1	3	3	7.5
	2.5	Produce documentation, **tutorials, sample data**	Core		1	3	3	7
	2.6	Provide **base data**	Core		1	2	2	5
	2.7	Provide libraries, **sample code**	Core		1	1	1	3
		Total		1	8	12.5	12.5	**34**
		WP3: Tools	**Respon-sible**	**Phase 0 Q1**	**Phase 1 Q2-Q4**	**Phase 2 Q4-Q9**	**Phase 3 Q7-Q12**	**Total Q1-Q12**
	3.1	**Validation tool** (profiles)	External	1	3	1	1	6
	3.2	**Topology visualization**	External	1	3	3	1	8
	3.3	**Topology editor** (based on viewer)	External		1	6	3	10
	3.4	**Migration tools** (2.2 -> 3.0)	External		2	2	2	6
		Total		2	9	12	7	**30**
		WP4: Define Organization						
	4.1	Create **business model** based on open standard	Core		1			1
	4.2	Define and set-up **organization**	Core		1	1		2
	4.3	Define **service level**	Core		1	1		2
	4.4	Create **certification process**	Core		1	1		2
	4.5	Define **release cycles**	Core			1		1
		Total		0	4	4	0	**8**
		WP5: Instruction & training						
	5.1	**training concept**	Core			1		1
	5.2	**train the trainers** (material + actual training)	Core / External			0.25		0.25
	5.3	**training** (assumption 5 adopters)	External				1.25	1.25
		Total		0	0	1.25	1.25	**2.5**
		WP6: R&D						
	6.1	**R&D**	Universit.			6	6	12
		Total		0	0	6	6	**12**
		Total for all workpackages		**8**	**24**	**42**	**31**	**105**

9 Terms, Jargon

Dublin Core	Method for formulating metadata, see http://www.dublincore.org
ERA	European Railway Agency, http://www.era.europa.eu
ERIM	European Rail Infrastructure Masterplan
IM	Infrastructure manager (railway)
Infrabel	IM of national network of Belgium
JBV	Jernbaneverket, IM in Norway
Network Rail	IM in the United Kingdom
Prorail	IM of national network of Netherlands
ÖBB	Österreichische Bundesbahnen, IM (and RU) of Austria
railML	Railway exchange format and community, http://www.railml.org
RFF	Réseau Ferré de France, IM of the national railway network of France
SBB	IM of national network of Switzerland
UIC	International Union of Railways, http://www.uic.org