ERA Ontology

Solving data silos in railway domain

EUROPEAN
UNION
AGENCY
FOR RAILWAYS

Photo: Adobe Stock / by_Tarasylo

EUROPEAN UNION AGENCY FOR RAILWAYS

ERA in a nutshell

The **objective of the Agency**:

- ✓ Enhancing the level of interoperability
- ✓ Developing a **common approach to safety**
- ✓ Issuing **authorizations** for the placing on the market of railway vehicles and vehicle types
- ✓ Issuing **single safety certificates** for railway undertakings.

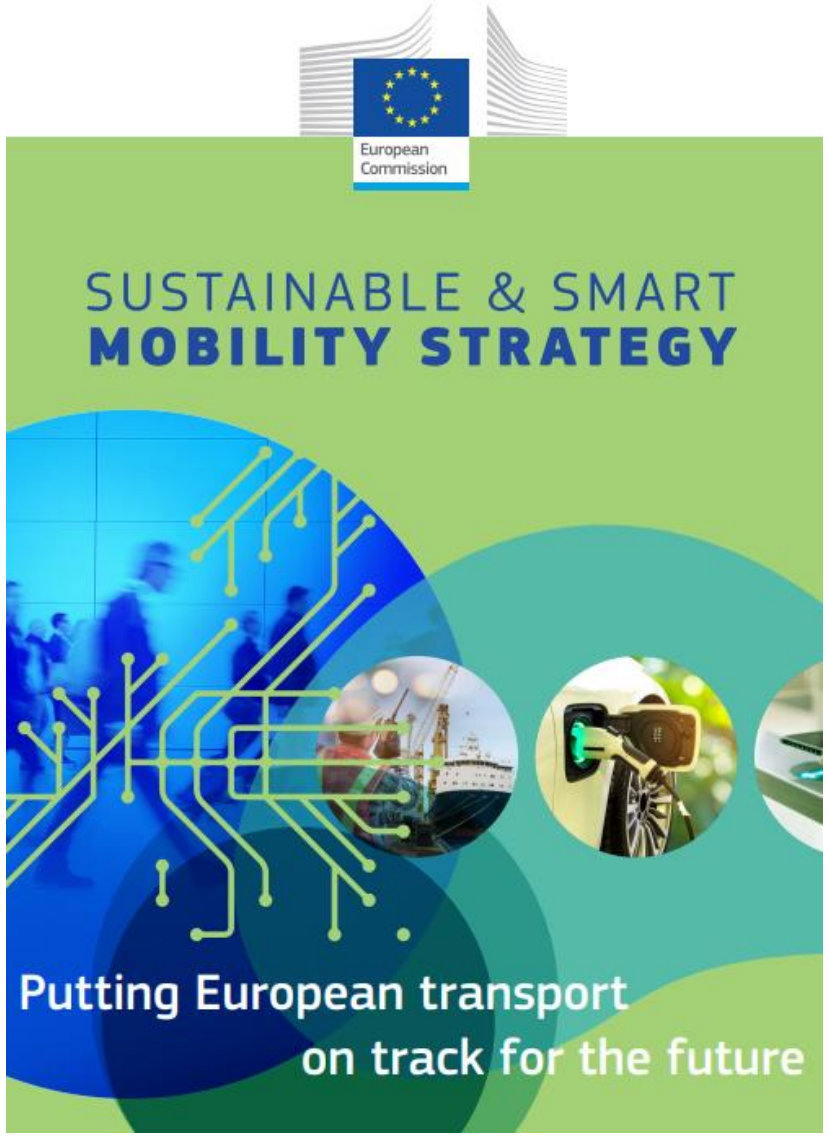
Legal Framework: Agency Regulation (Regulation (EU) 2016/796 of the European Parliament and of the Council of 11 May 2016 on the European Union Agency for Railways)

EUROPEAN UNION AGENCY FOR RAILWAYS

How did it start ?

EUROPEAN UNION AGENCY FOR RAILWAYS

The Digital rail and Green freight TSI revision package falls within the EU policy objectives



EU objectives:

- **a European Green Deal,**
- **an economy that works for people,**
- **a Europe fit for the digital age and**
- **a stronger Europe in the world.**

EUROPEAN UNION AGENCY FOR RAILWAYS

The Journey towards Data Centricity

2019

Siloed data bases /registers

Around 12 DBs application oriented

2021

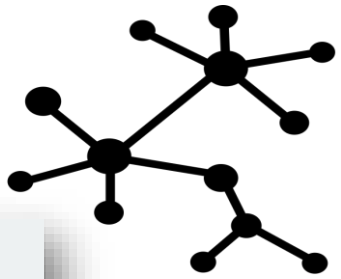
Management Board
Linked Data
Mainstream Decision

EUROPEAN UNION AGENCY FOR RAILWAYS

RINF - Register of Infrastructure

Connected into Linked Data

2022

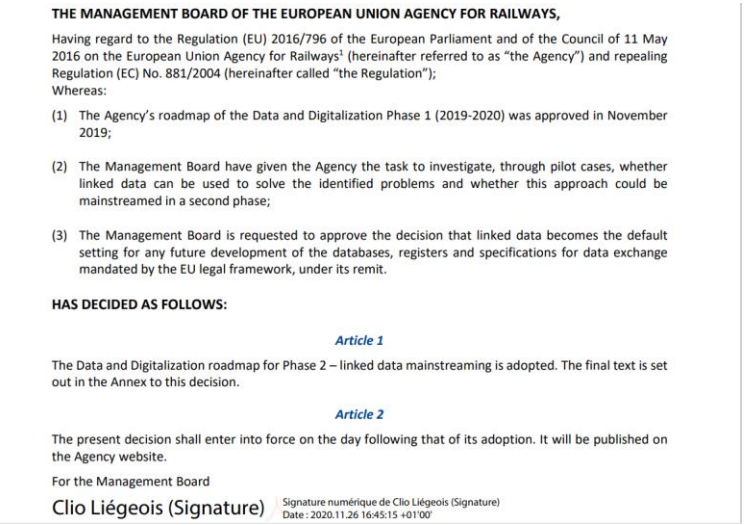


Register of Railway Vehicles

2024

European Railway Occurrences Data *

Vehicle Authorization Data *



THE MANAGEMENT BOARD OF THE EUROPEAN UNION AGENCY FOR RAILWAYS,

Having regard to the Regulation (EU) 2016/796 of the European Parliament and of the Council of 11 May 2016 on the European Union Agency for Railways[1] (hereinafter referred to as "the Agency") and repealing Regulation (EC) No. 881/2004 (hereinafter called "the Regulation");
Whereas:

(1) The Agency's roadmap of the Data and Digitalization Phase 1 (2019-2020) was approved in November 2019;

(2) The Management Board have given the Agency the task to investigate, through pilot cases, whether linked data can be used to solve the identified problems and whether this approach could be mainstreamed in a second phase;

(3) The Management Board is requested to approve the decision that linked data becomes the default setting for any future development of the databases, registers and specifications for data exchange mandated by the EU legal framework, under its remit.

HAS DECIDED AS FOLLOWS:

Article 1

The Data and Digitalization roadmap for Phase 2 – linked data mainstreaming is adopted. The final text is set out in the Annex to this decision.

Article 2

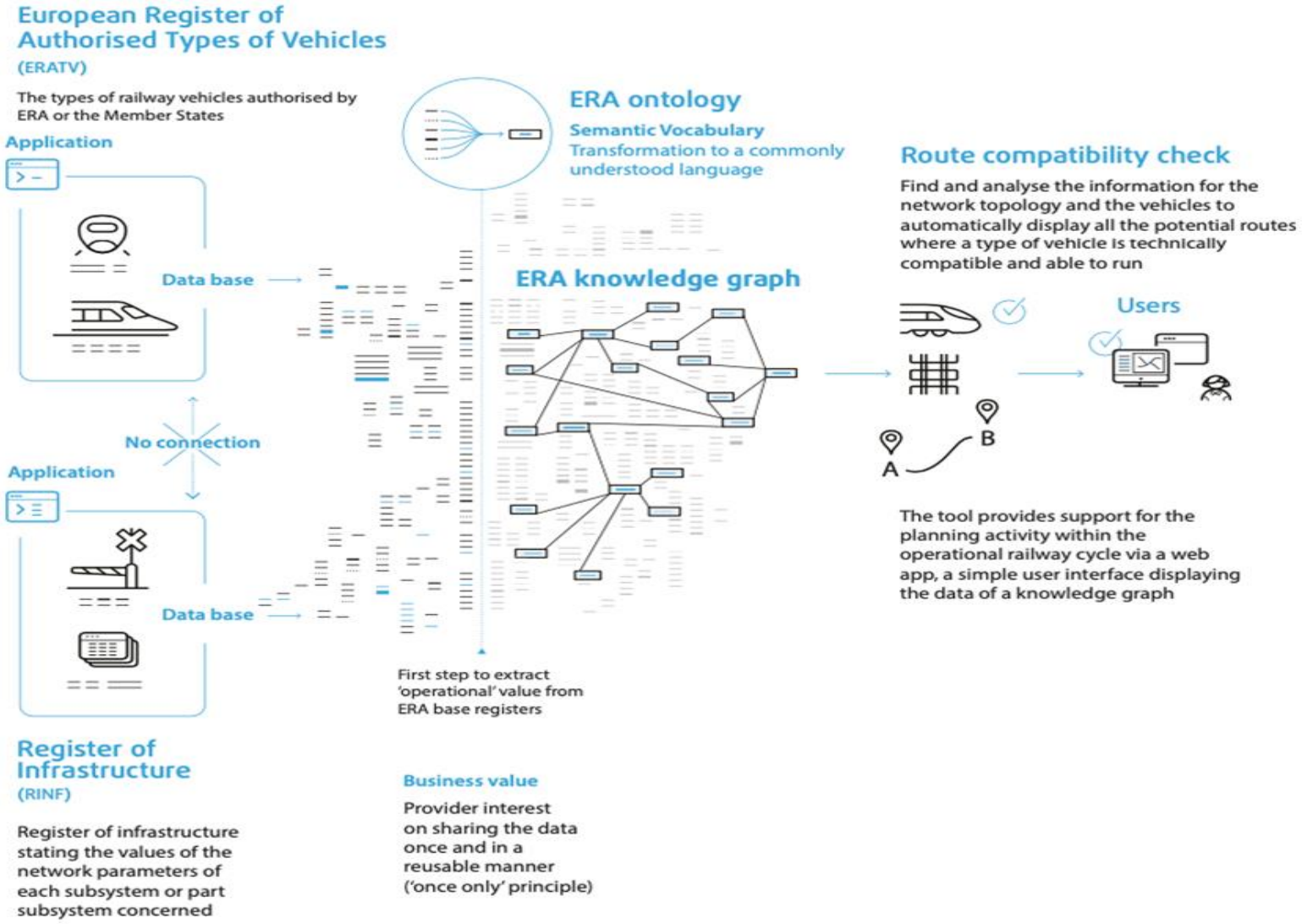
The present decision shall enter into force on the day following that of its adoption. It will be published on the Agency website.

For the Management Board

Clio Liégeois (Signature) Signature numérique de Clio Liégeois (Signature) Date : 2020.11.26 16:45:15 +01'00'

EUROPEAN UNION AGENCY FOR RAILWAYS

A successful PoC – A Real-world Use Case



EUROPEAN UNION AGENCY FOR RAILWAYS

Article 7a

ERA vocabulary

'ERA Vocabulary' means a Technical Document issued by the Agency pursuant to Article 4(8) of Directive (EU) 2016/797, establishing human and machine-readable data definitions and presentations and linked quality and accuracy requirements for each data element (ontology) of the rail system.

The Agency shall ensure the ERA vocabulary is maintained to reflect regulatory and technical developments affecting the rail system.

COMMISSION IMPLEMENTING REGULATION (EU) 2019/777 of 16 May 2019

EUROPEAN UNION AGENCY FOR RAILWAYS

Moving Europe towards a sustainable and safe railway system without frontiers.

OSS Login Search

THE AGENCY & YOU | DOMAINS | LIBRARY | EVENTS & TRAINING | NEWS & PRESS | CAN WE HELP YOU?

Home > News > ERA takes action to reduce fragmentation in the European rail data landscape

ERA takes action to reduce fragmentation in the European rail data landscape

Published: 06 December 2023 **Updated:** 07 December 2023

The European Union Agency for Railways (ERA) has taken significant steps towards its mission of fostering a sustainable and safe railway system without frontiers. On 6 December 2023, ERA announced the official launch of the ERA Railway Factsheets, an intuitive tool aimed at unlocking data and addressing fragmentation within the European railway data landscape. Importantly, ERA presented its fundamental work on linked data and the ERA knowledge graph: the data architecture that enables efficiency gains for the entire European railway sector.

Press release

Page content

The ERA Factsheets offer a comprehensive overview of the rail sector across European countries with one-click access. Covering aspects such as transport statistics, railway organisations, rolling stock, infrastructure, safety, market structure, and the legal framework, the tool unlocks data from ERA and other organisations, providing a broad and accurate representation of the European railway landscape.

Whereas ERA Factsheets bring fragmented data together in an accessible format, ERA's linked data project changes the fundamentals on how railway data is defined and shared. Today, railway data is typically kept in isolated and incompatible databases. Linked data is central to ERA's strategic plan to overcome challenges such as data gaps, unaligned definitions, and accessibility barriers. It sets the foundation for more efficient data sharing between railway actors across Europe and novel applications. ERA presents the substantial progress that has been achieved and several future use cases.

Through linked data and the Factsheets, ERA is taking a crucial step towards increasing the efficiency of railway operations, filling knowledge gaps, and facilitating informed policymaking.

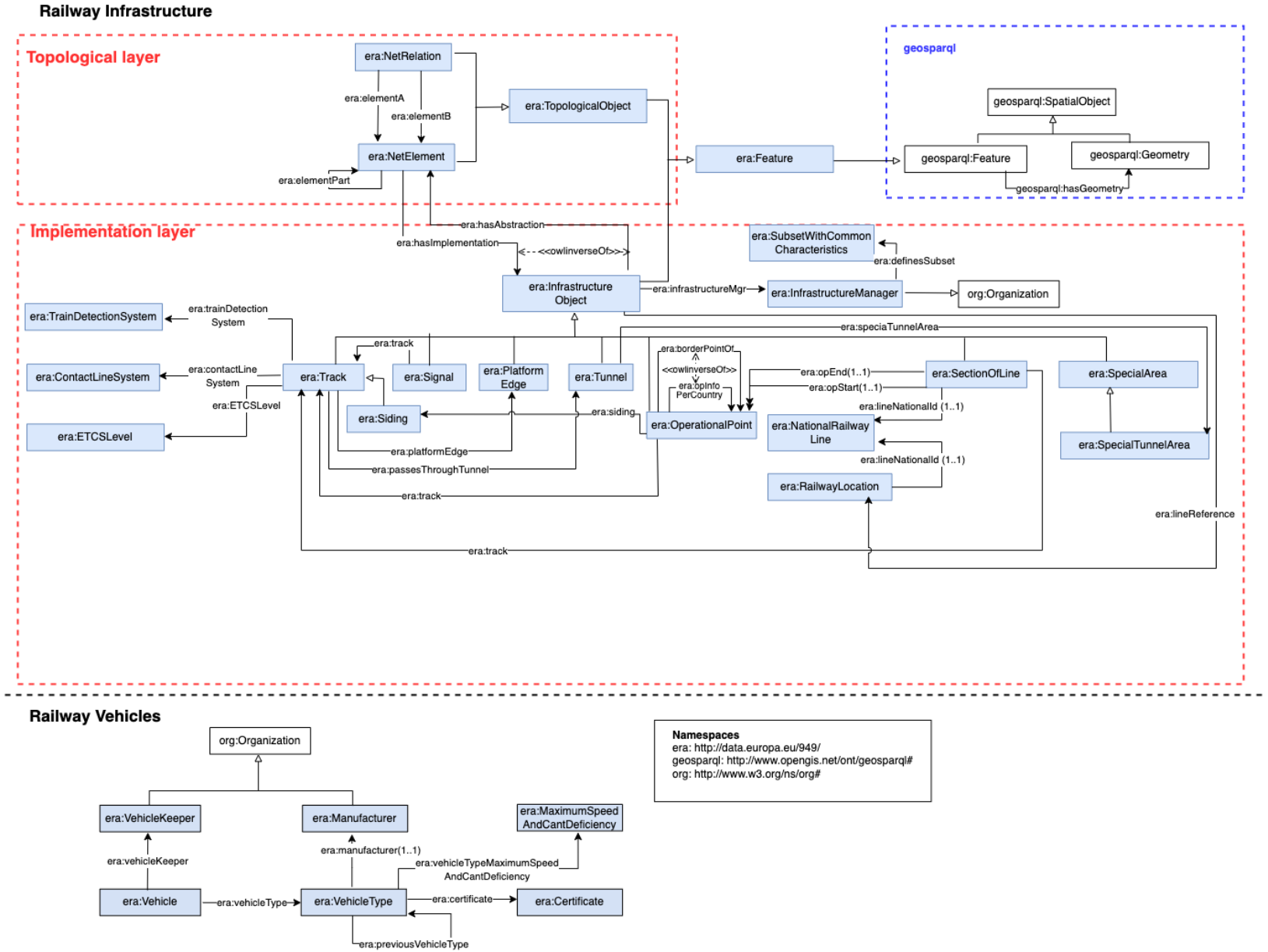
> Linked Data and the ERA Railway Factsheets contribute to our mission to create a sustainable and safe railway system for Europe. They provide tangible solutions to historical challenges in the European railway data landscape. By addressing fragmentation and enhancing accessibility, we aim to streamline operations and facilitate evidence-based policymaking to empower European railways.
>
> — Josef Doppelbauer, ERA's Executive Director

The ERA Railway Factsheets are set to be regularly updated, reflecting the commitment to ensure that the information remains current and relevant over time. ERA encourages stakeholders, innovators, and citizens to explore the Factsheets, available on ERA's website, and provide feedback to contribute to the continuous improvement of this pivotal tool: https://www.era.europa.eu/content/era-railway-factsheets

EUROPEAN UNION AGENCY FOR RAILWAYS

ERA Ontology Overview – v3.0.0

https://data-interop.era.europa.eu/era-vocabulary/



EUROPEAN UNION AGENCY FOR RAILWAYS

KG GENERATION PIPELINE

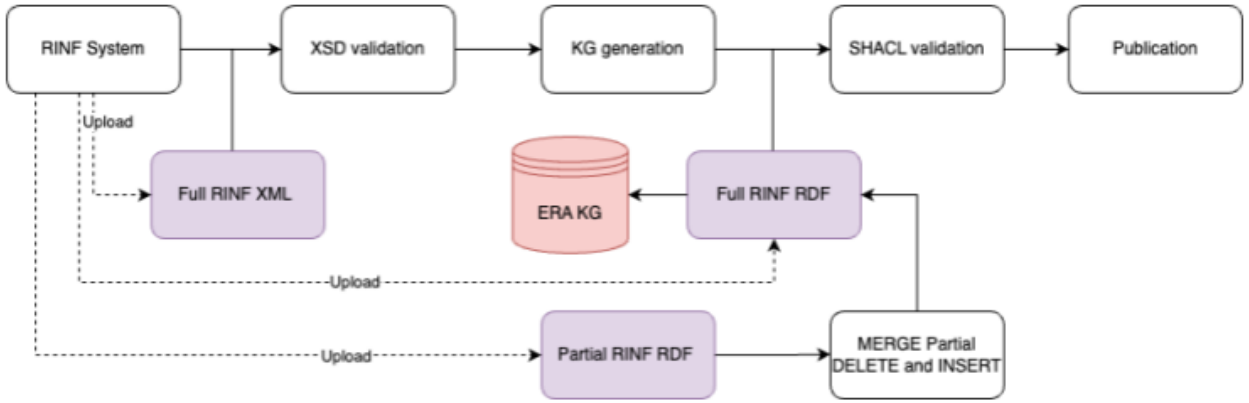
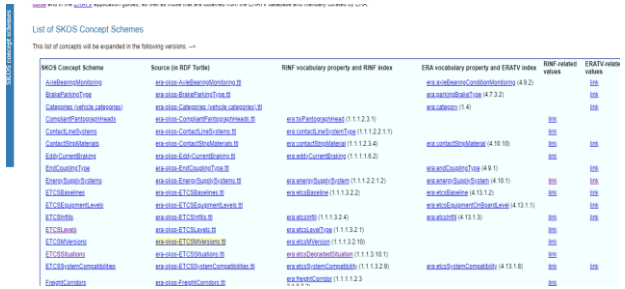


Figure 4. Validation (XSD-based and SHACL-based), KG generation and publication steps for datasets.

Demo dataset manager full xml (youtube.com)

ERA vocabulary (europa.eu)

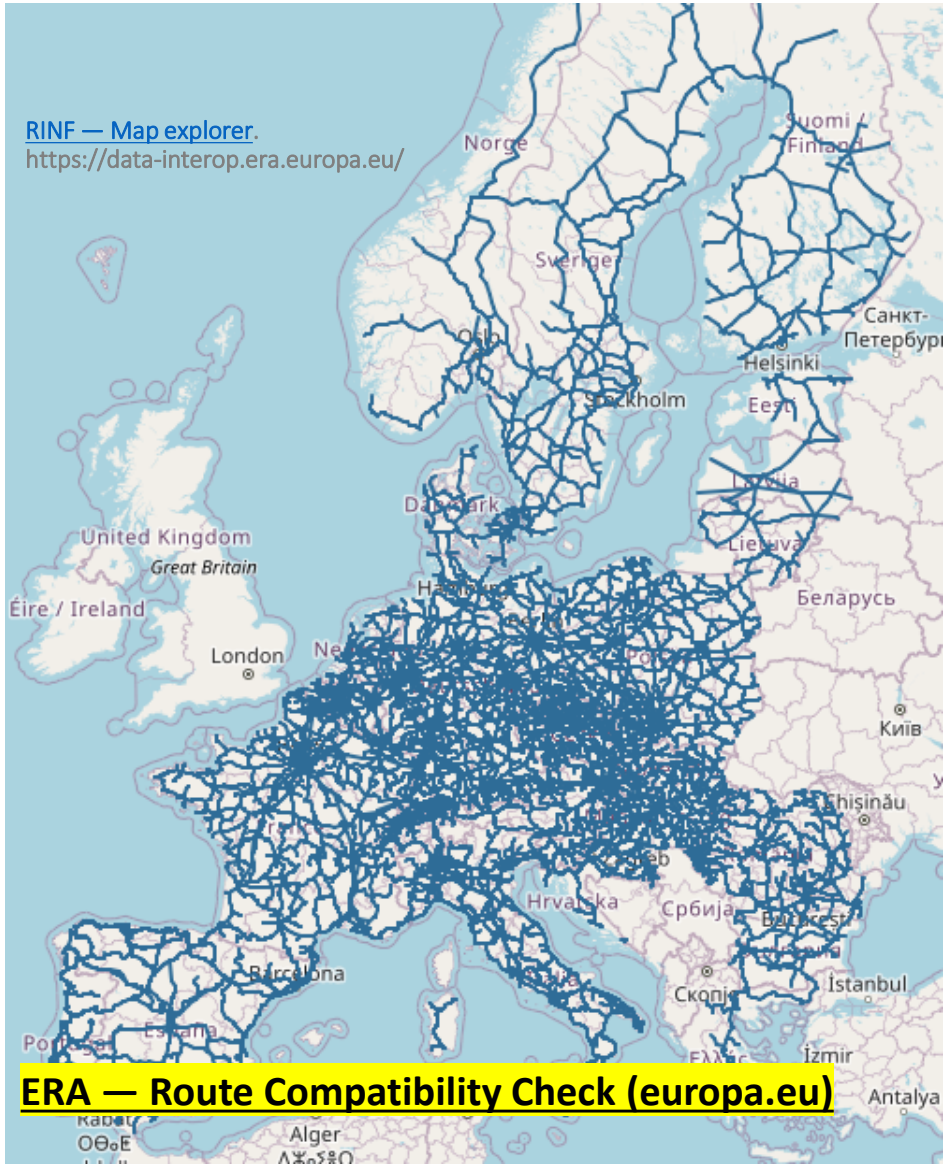


SKOS Concept Schemes used in the ERA vocabulary (europa.eu)

SHACL
Validating RDF graphs

Link to ERA eng. rules

EUROPEAN UNION AGENCY FOR RAILWAYS



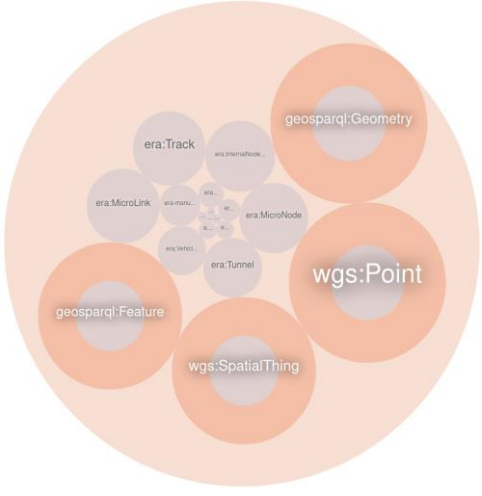
ERA — Route Compatibility Check (europa.eu)

Ontology Specification

EUROPEAN UNION AGENCY FOR RAILWAYS

ERA vocabulary. Version 3.0.0

This version:
https://data-interop.era.europa.eu/era-vocabulary/
Previous version:
https://zenodo.org/record/7775344
Version:
v3.0.0 (released on 2023-03-29)
Publisher:
European Union Agency for Railways
Download serialization:
Format JSON LD | Format RDF/XML | Format N Triples | Format TTL
Browse SKOS thesauri:
Format HTML
Download SHACL shapes:
Format TTL



ERA KG endpoint

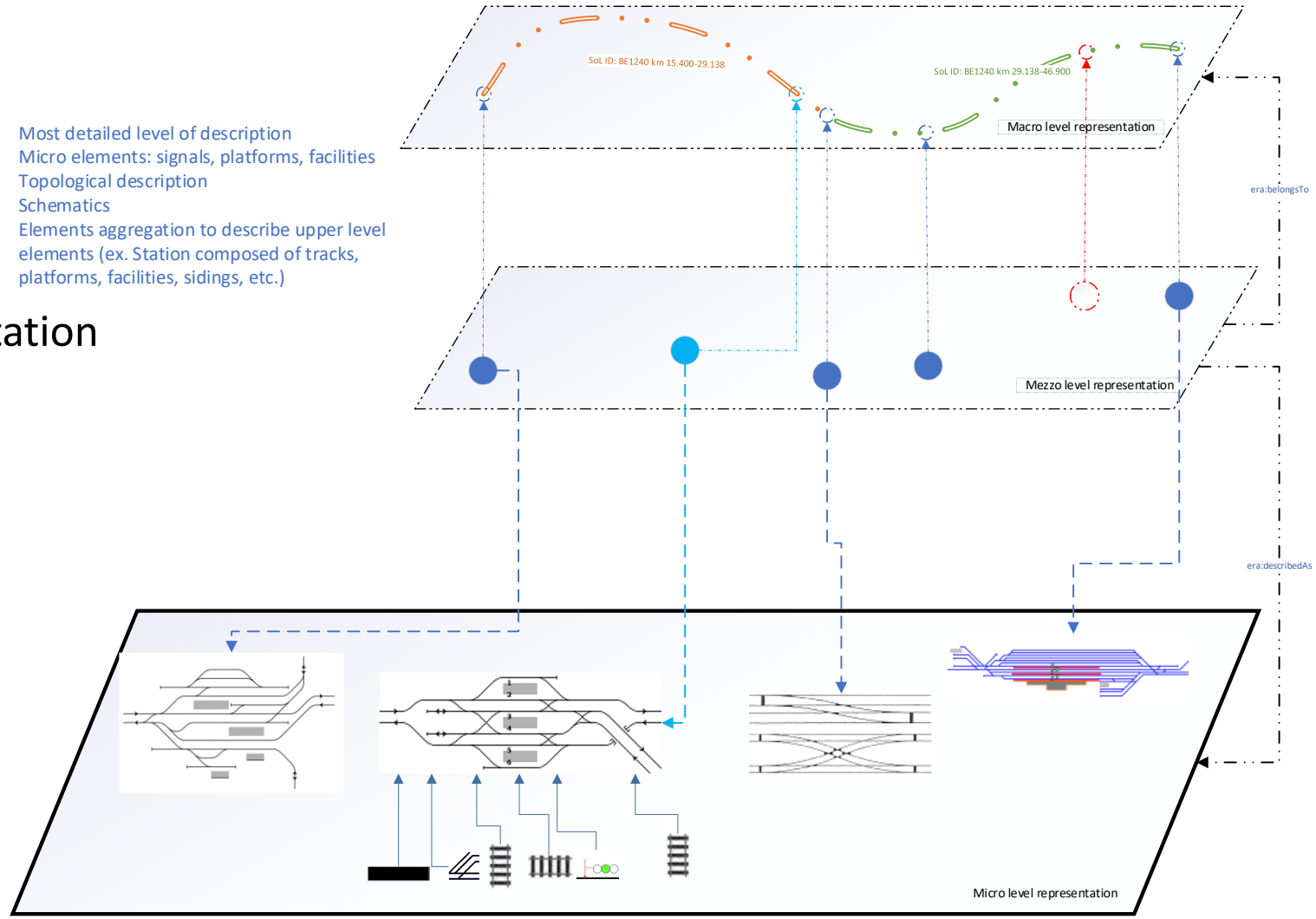
- More than 47 million triples
- More than 31k lines of mappings
- More than 100 SHACL shapes
- +270k track segments described
- +50k stations described
- +50k geo-referenced objects (lat/long)
- +2k Vehicle Types described
- 27 countries covered (EU countries)

EUROPEAN UNION AGENCY FOR RAILWAYS

Data modelling requirements

- Different levels of granularity of data

Most detailed level of description
Micro elements: signals, platforms, facilities
Topological description
Schematics
Elements aggregation to describe upper level elements (ex. Station composed of tracks, platforms, facilities, sidings, etc.)



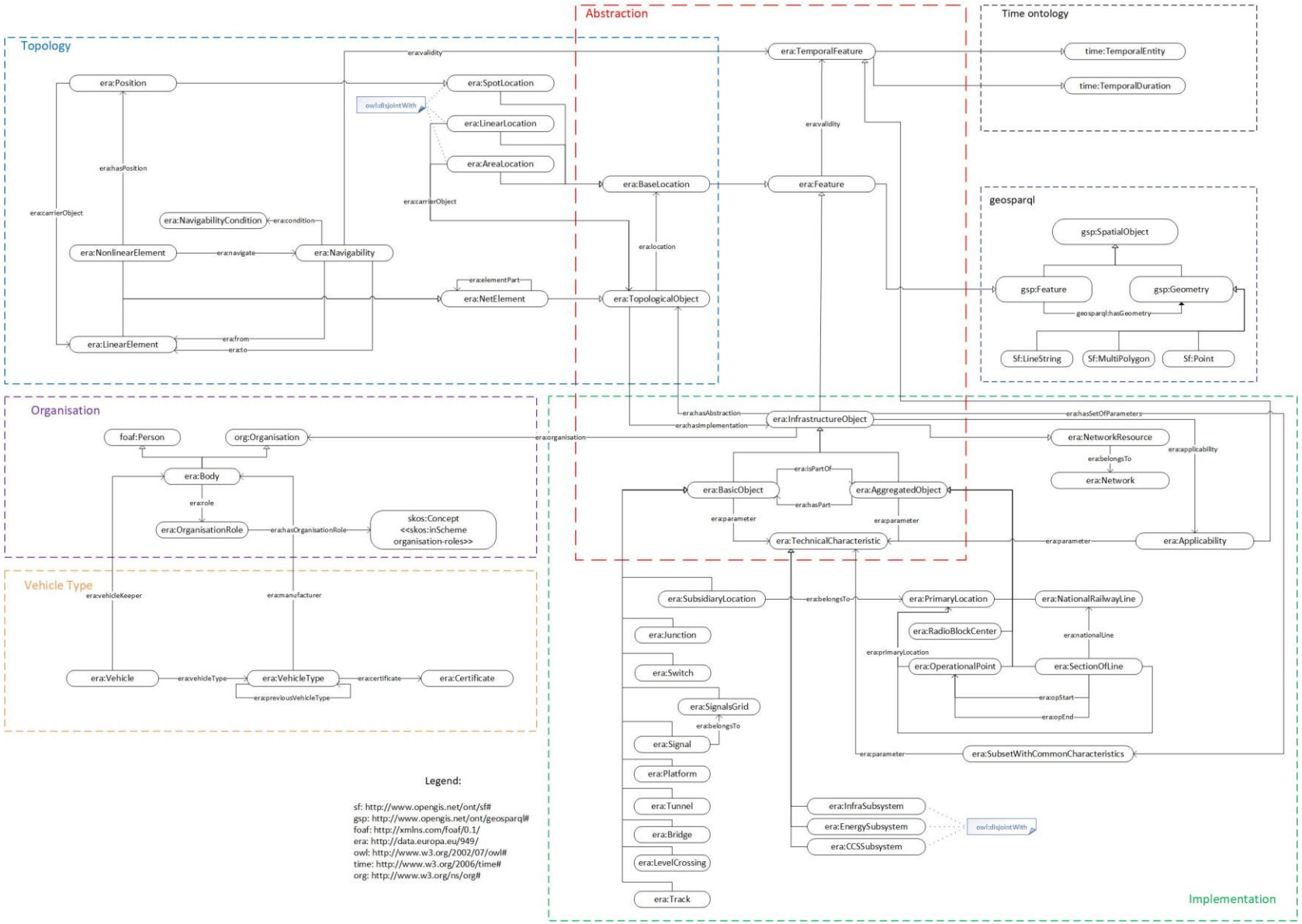
- Topological and topographical representation of railway infrastructure

- Model the temporal constraints for the technical characteristics

EUROPEAN UNION AGENCY FOR RAILWAYS

Integration of Micro-Level

Use cases covered:
- Route book data compilation
- Tracks connectivity and navigability
- Better geographical representation
- Telematics locations data integration
- Future planning for infrastructure

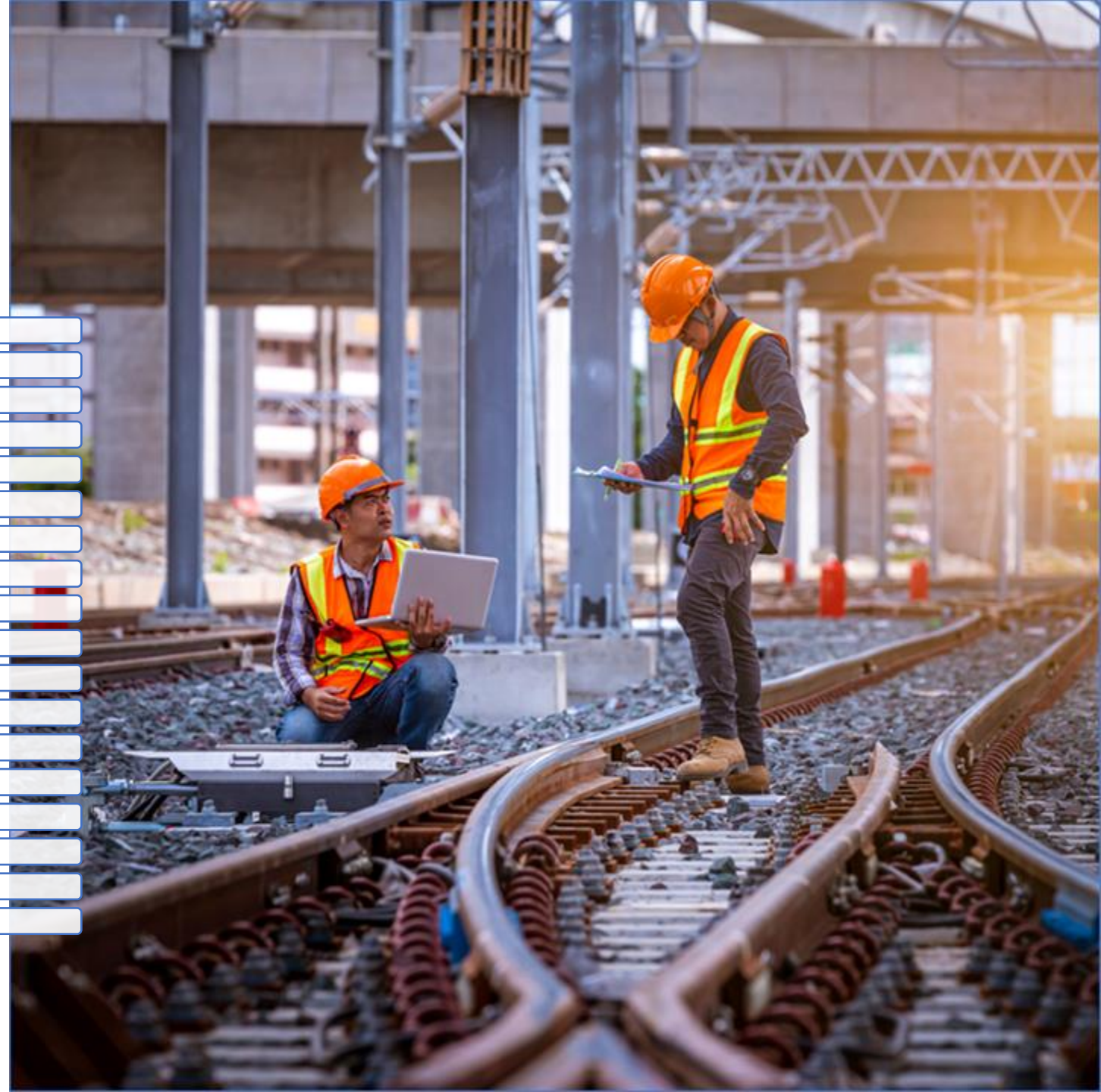


[Interoperable-data/ERA-Ontology-3.1.0: Extended version of the ERA Railway Infrastructure Ontology (github.com)]

EUROPEAN UNION AGENCY FOR RAILWAYS

Competency questions

- CQ 1 •How many tunnels are there per member state?
- CQ 2 •Which are the distinct values for a particular parameter P of a track in member state C?
- CQ 3 •What is the tunnel length per member state?
- CQ 4 •What is the railway length per member state? (takes into account sections of lines and tunnels)
- CQ 5 •What are the platform lengths in ascending order? (What is the variety of platform lengths in Europe?
- CQ 6 •What are the parameters of a certain entity E in the vocabulary
- CQ 7 •Which are non TSI compliant legacy signaling systems existent in the different countries?
- CQ 8 •What are the number of tons per track?
- CQ 9 •Which are the energy supply system values (voltage frequency) per SOL per member state?
- CQ 10 •What is the network of an infrastructure manager IM?
- CQ 12 •Which are the diesel vehicle types?
- CQ 13 •Which are the tracks with a contact line system type that is not electrified
- CQ 14 •What is the track length of Trans-European Transport Network (TEN-T) lines per member state?
- CQ 16 •How many tracks are equipped with GSM-r per member state (rate of deployment)?
- CQ 18 •Which are the tunnels in Belgium that are not category (rolling stock fire category) A nor B?
- CQ 19 •How many countries are using local and non-international Gauges (gauging profile)?
- CQ 20 •What are the available values for the GSM-R version per member state?
- CQ 21 •What are the available ETCS levels per member state?



[ERA Knowledge Graph | Data Stories]

Challenges

EUROPEAN UNION AGENCY FOR RAILWAYS

EUROPEAN UNION AGENCY FOR RAILWAYS

The Challenges

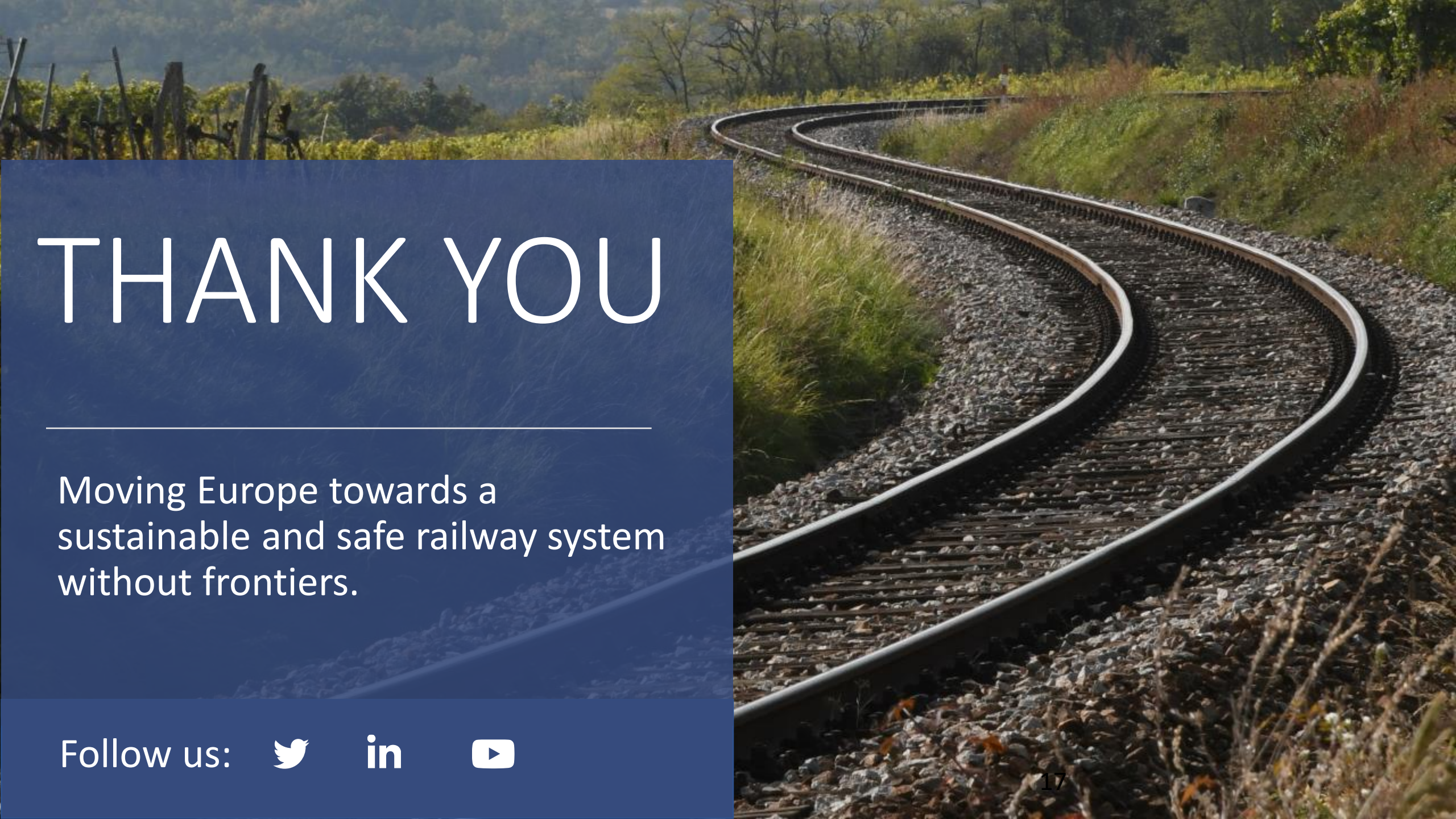
- As a regulatory body our data exchange rules have an impact in the regulatory reporting actors
- Multi-lingual aspect of the ontology, due to technical terms
- Coexistence of data compliant with different regulatory text
- Scalability issue with SHACL rules in data provision of small updates
- Building rich application functionalities on top of KG

EUROPEAN UNION AGENCY FOR RAILWAYS

Take Away Message

Use of Open Software	👍
Use of W3C and open standards	👍
Agnostic to different formats (check Loops Loop example)	👍
Aligned with other ontologies	👍
Aligned with existing models (UMLs, XSD, etc.)	👍
New questions to the knowledge graph formulated on demand	👍
Backwards compatible with previous (in house) developments	👍
Data more relevant than the application layer Easily extendable with other sources	👍

- Railway infrastructure managers provide data in RDF for usage.
- ERA publishes and maintains railway infrastructure using (open) semantic web technologies.
- Public
  Endpoint: https://virtuoso.ecdp.tech.ec.europa.eu/sparql
- Questions/suggestions:
  eraontology@era.europa.eu / ghislain.atemezing@era.europa.eu



THANK YOU

Moving Europe towards a sustainable and safe railway system without frontiers.

Follow us: