A Railway Knowledge Graph in the EU Mobility Data Space

Prof Dr Marina Aguado
Data & Information Governance Team
Florence

EUROPEAN UNION AGENCY FOR RAILWAYS

DIGIT

GHENT UNIVERSITY imec

EUROPEAN UNION AGENCY FOR RAILWAYS

A real data story about the FAIRification of Mobility Register based information from a public institution **into** a building block of the EU *FAIR** Mobility Data Space

The ERA Knowledge Graph:
EU Transport Authority using knowledge graph technology to share European railway topology and vehicle data

EU DATA PLATFORM SPARQL Query Service

https://linked.ec-dataplatform.eu/sparql/

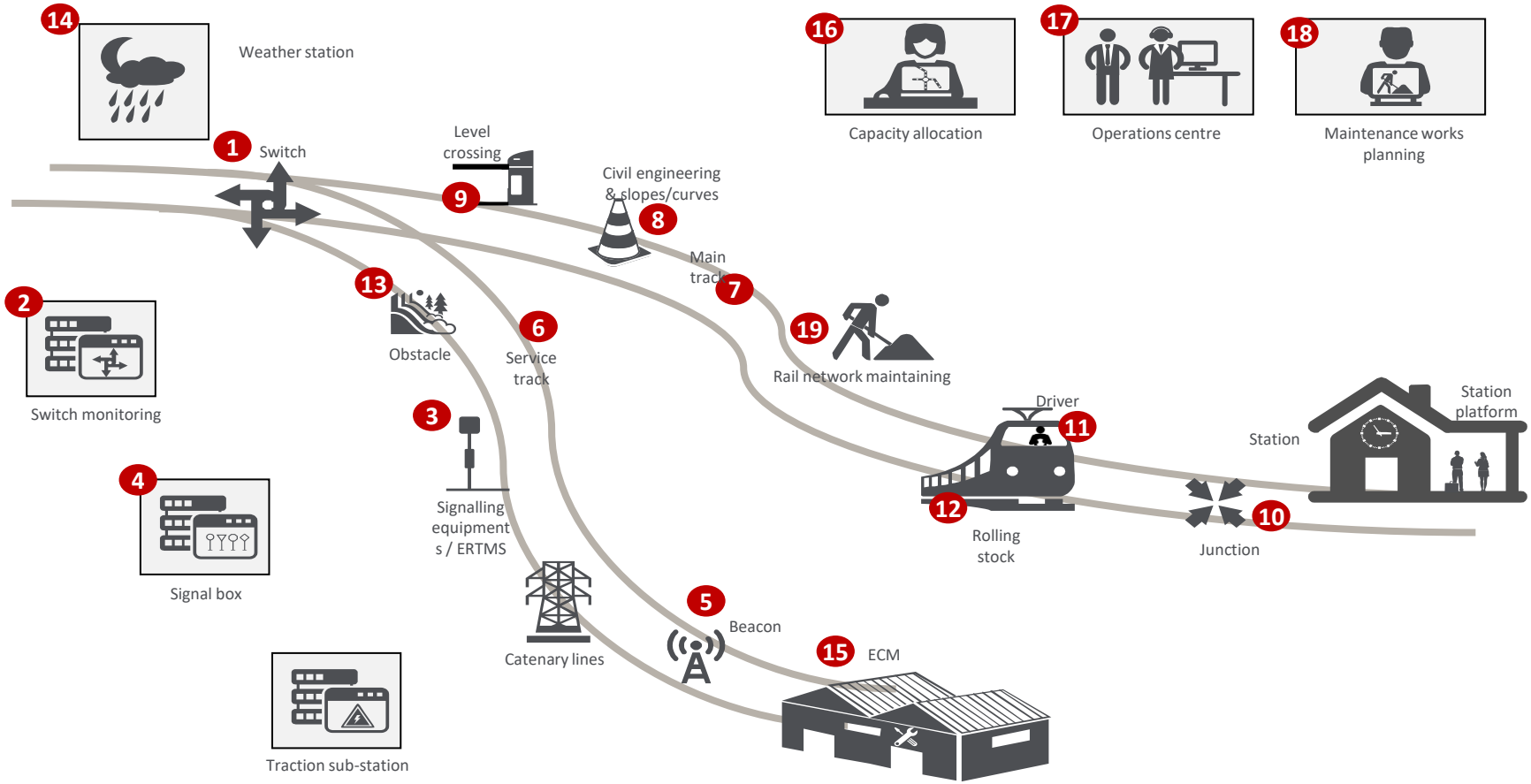
http://era.europa.eu/knowledge-graph

* **F**indible **A**ccessible **I**nteroperable **R**eusable

EUROPEAN UNION AGENCY FOR RAILWAYS

The railway challenging ¨data ¨ ecosystem

Multiple actors request and trust data from third parties to perform their business



EUROPEAN UNION AGENCY FOR RAILWAYS

Towards a Single European Railway Area

- The EU has one of the **densest railway networks** in the world
- However national railway systems across the EU vary
- Since 1990s EU has adopted **four legislative railway packages:**
  - to open the railway market to **competition**,
  - to increase the **interoperability** of national railway systems

  ultimately to define the framework for a **Single European Railway Area**.

Single European Railway Area Convention
Brussels
safe - connected - affordable

Safety & Interoperability in the Single European railway network **SERA**

Making railways work better for a digital society

4TH Railway Package and Registers

Business to Government Data Exchange

B2G

Direct data exchange between:

Stakeholders, Member States **&** *The Agency*

(National Registration Entities, National Safety Authorities, Infrastructure Managers, Operators, Wagon Keepers, Notified Bodies…)

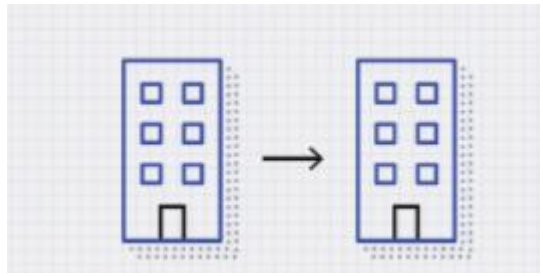
The Agency´s data ecosystem

Railway Sector

Business to Government Data Exchange

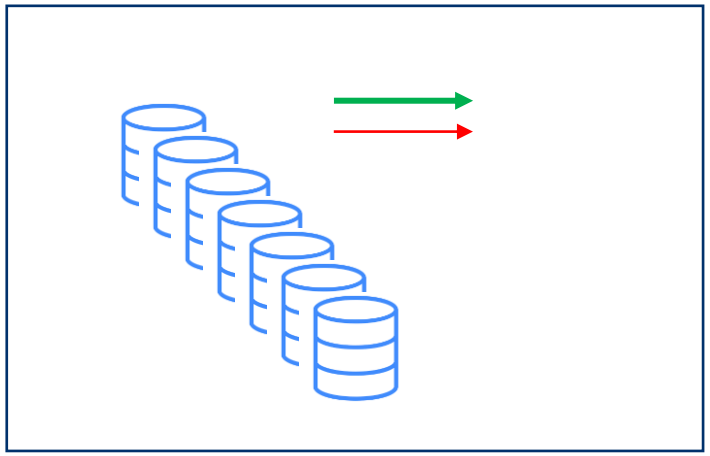
B2G

Railway Business to Business Data Exchange **B2B**



- TAF/TAP (IM, RU exchange)
- ERTMS (Supplier IM, RU)

BASE Registers



1st Meta Data – Common Ontology – Data Catalogue - Reference Data

2nd Data Access

Digital Transformation of our legal documents : Agency as vocabulary provider

Legal Basis
4th Railway Package:

Working Party
Harmonization on terms
plus expert knowledge

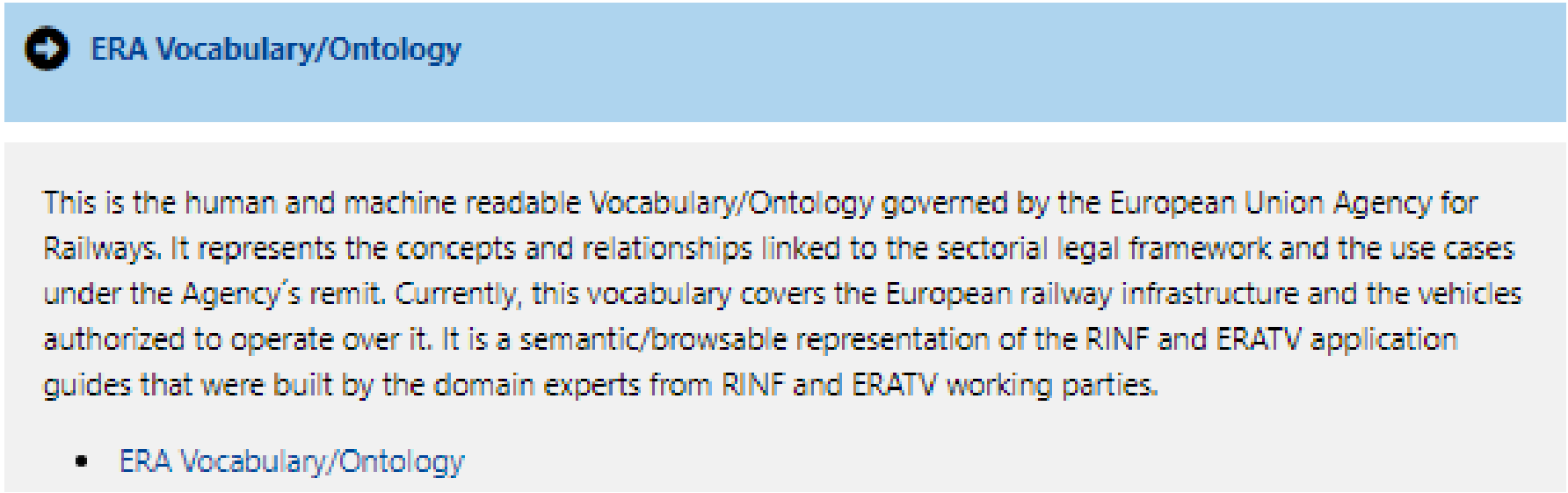
The Agency acts as a neutral agent and as a data intermediary with a leading role in the field by validating, curating, storing and publishing register based information so it can be reused and exploited by the sector to enhance data interoperability between the different players.

The Agency´s aim is to harness the harmonisation in terminology achieved in the legal railway documents -via domain experts contribution-, and to expose it in digital terms (core vocabulary/ontology, controlled vocabulary, reference data management).

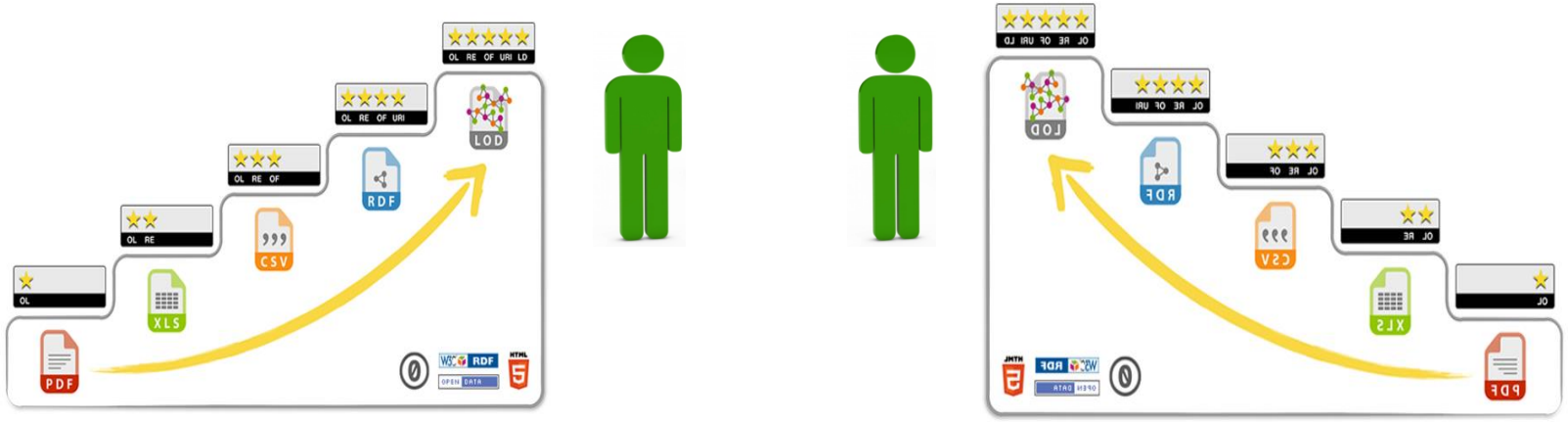
The ERA vocabulary/ontology

Railway digitalization also means to provide a digital identity -accessibility- to all the railway assets, a browsable **U**niform and unique **R**esource **I**dentifier (URI), with a secured control access when needed

Just launched the human and Machine readable (=NOT PDF) version of the RINF Application Guide. More flexibility to adapt to new amendments and/or parameters and controlled vocabulary



Railway digitalization also means to provide a digital identity -accessibility- to all the railway assets, a browsable **U**niform and unique **R**esource **I**dentifier (URI), with a secured control access when needed

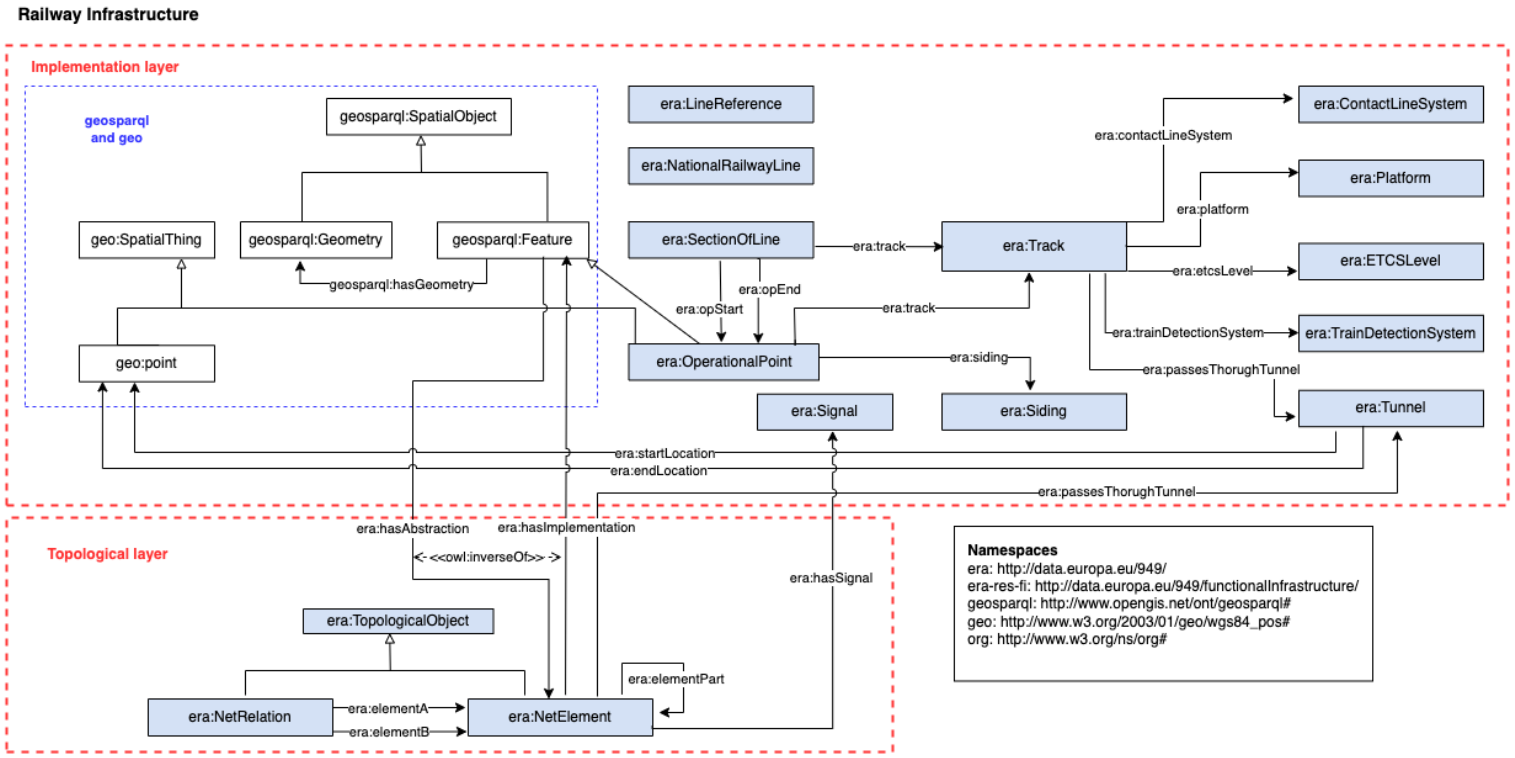


EUROPEAN UNION AGENCY FOR RAILWAYS

ERA Vocabulary

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

Semantic vocabulary defined by ERA to describe the concepts and relationships related to the European railway infrastructure and the vehicles authorized to operate over it.



Ontology Specification

ERA vocabulary

Release 2021-05-31

This version:
https://git.fpfis.eu/datateam/ERA/era-vocabulary/-/releases/v2.4.0

Latest version:
https://git.fpfis.eu/datateam/ERA/era-vocabulary/-/releases/v2.4.0

Previous version:
https://git.fpfis.eu/datateam/ERA/era-vocabulary/-/releases/v2.3.2

Revision:
v2.4.0

Authors:
Julián Rojas, (Ghent University - imec)

Contributors:
Dylan Van Assche, (Ghent University - imec)
Ivo Velitchkov, (DG DIGIT)
Marina Aguado, (ERA)
Pieter Colpaert, (Ghent University - imec)
Polymnia Vasilopoulou, (ERA)
Edna Ruckhaus, (UPM)
Oscar Corcho, (UPM)
Wouter Beek, Triply

Publisher:
European Union Agency for Railways

EUROPEAN UNION AGENCY FOR RAILWAYS

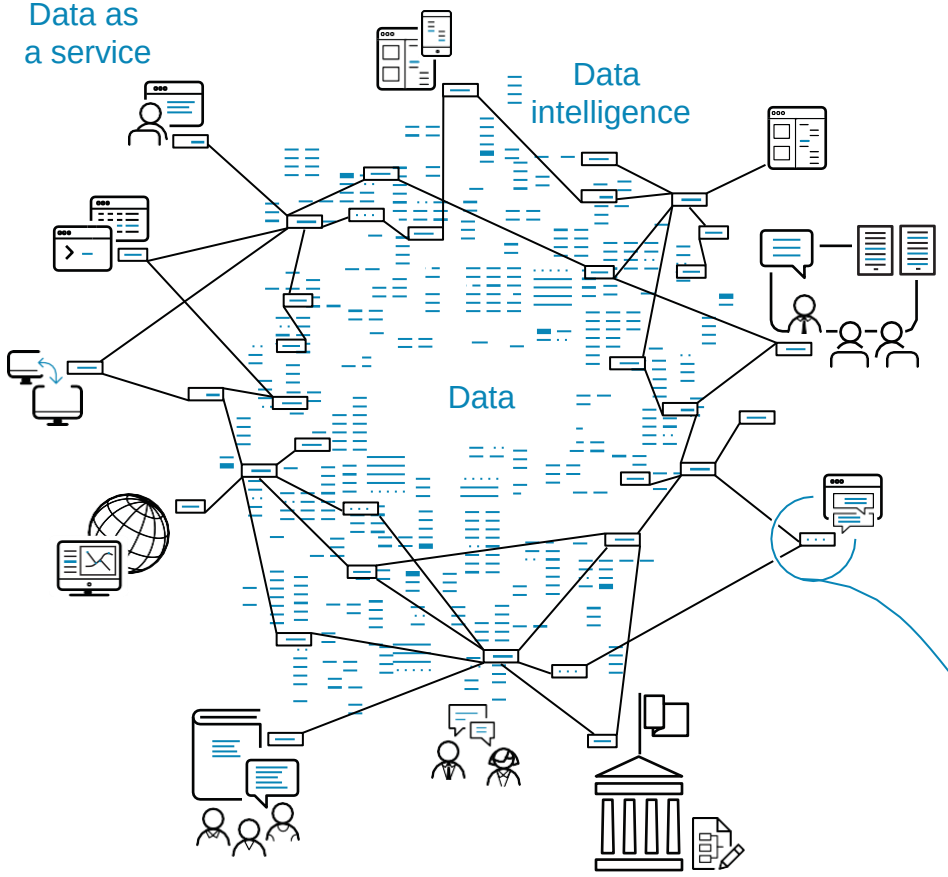
Agency role as *a neutral agent*

Despite the data may be private, shared or open, using data from multiple sources in the railway and multimodal exchange requires **interoperability at several levels**.

The Agency has started to play the active role of **vocabulary provider** - governance of the core ontology linked to the uses cases under its remit, management of the reference data -master data and taxonomies, enumerations..- linked to sectorial legal framework and uses cases under its remit . *Being a public institution and a neutral agent, the Agency is the best fit for this purpose.* [Commission High Level Group]

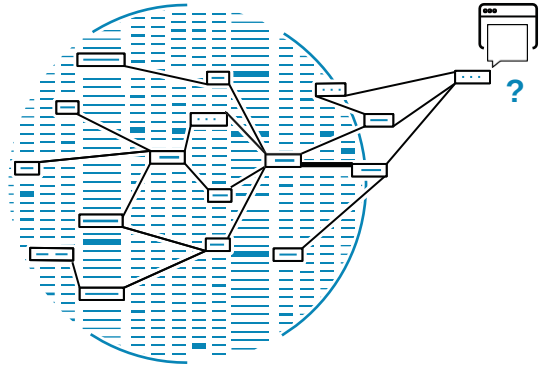
ERA route towards a data driven and data centric organisation

Data Centric Organization



From "siloed" to connected data

Web of data
not web of documents
URI granularity



New queries, previously unthinkable, are now possible (federated queries)

Agency´s digital roadmap:
from siloed to connected data

Strategic Calendar

2019: Data & Digitalization Roadmap phase 1

2020 : Interoperable Data Project

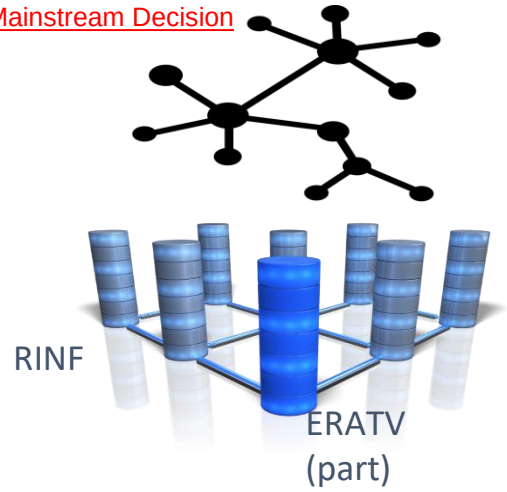
2021 : Linked Data MainStream Decision phase 2

Common Situation

No data exchange between different Information Systems

Silo & application centric approach

Linked Data Mainstream Decision



Connected data

Federated queries

New Insights

Automatic data quality control

Data centric approach

less heavy IT CAPEX and OPEX

DIGIT

EU data Platform

@data-interop.era.europa.eu !!

EUROPEAN UNION AGENCY FOR RAILWAYS

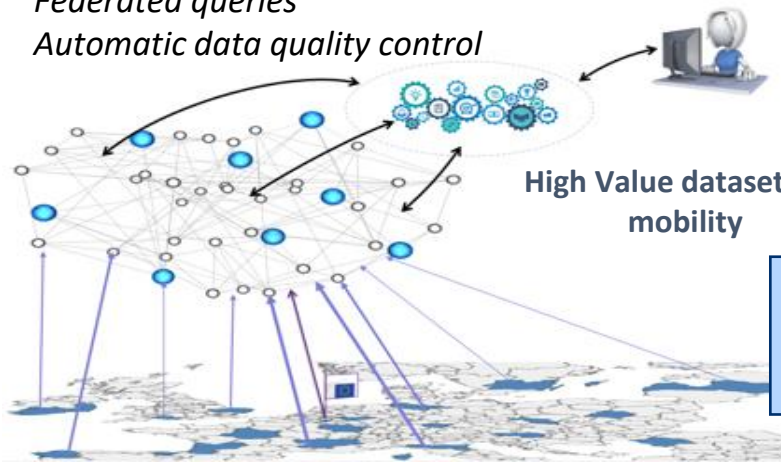
Agency´s digital roadmap:

"Fairification" of our railway datasets

Connected data

Federated queries

Automatic data quality control

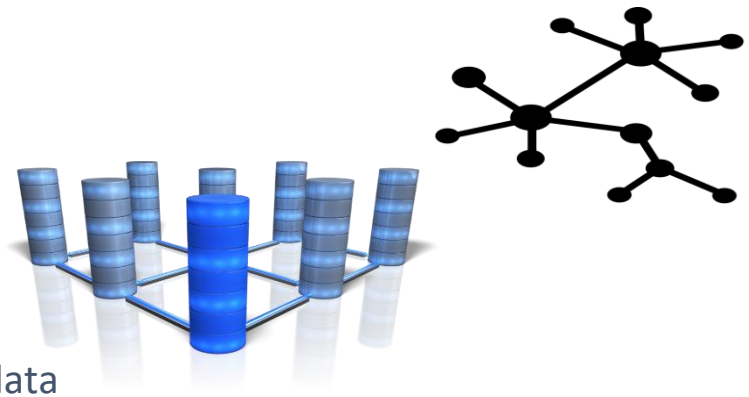


High Value datasets for mobility

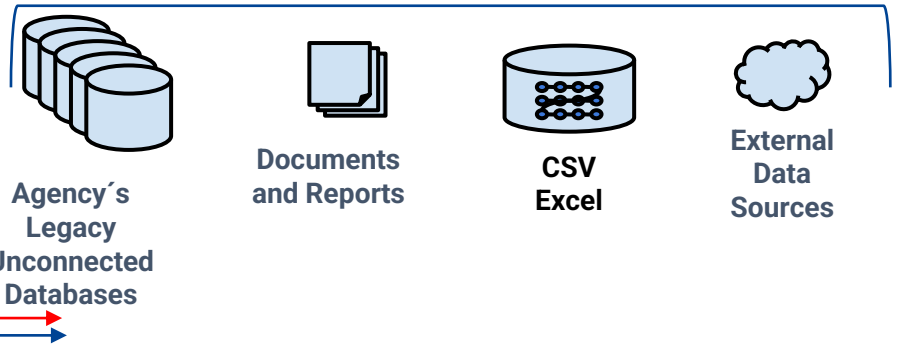
Common European Mobility Data Space

Mobility

DTLF | ITS NAPS | Vehicle data | Aviation | Road Safety



ERA in house data



ERA Knowledge Graph + SPARQL endpoint

- ~26 millions triples

Ontology

22 classes

318 properties

Reference Data = Controlled vocabulary

54 SKOS managed taxonomies

DIGIT

GHENT UNIVERSITY imec

EUROPEAN UNION AGENCY FOR RAILWAYS

Agency´s digital roadmap:

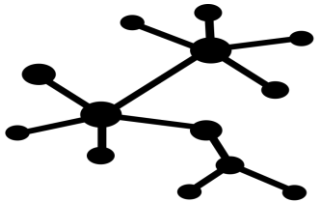
from siloed to connected data

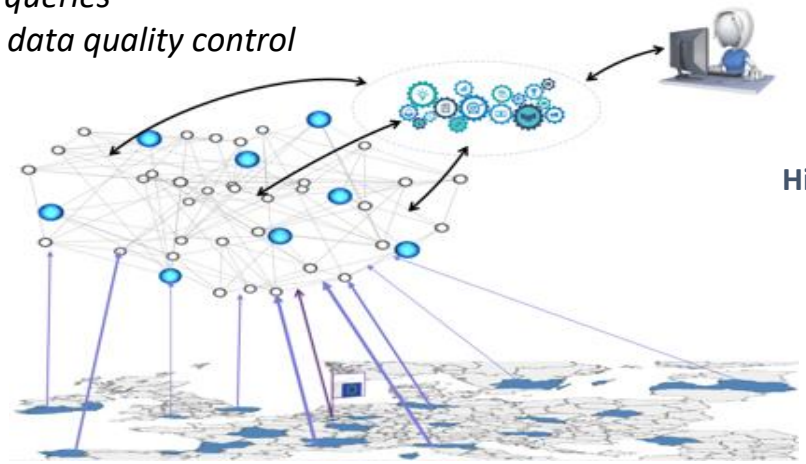
*"**Fairification**" of our railway datasets*

Connected data

Federated queries

Automatic data quality control





High Value datasets for mobility

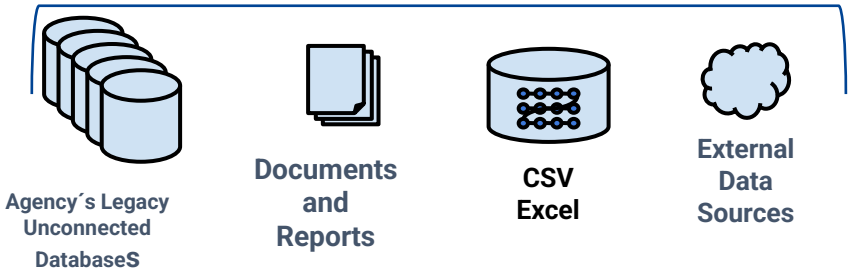
Mobility

DTLF

Common European Mobility Data Space

[ERA Knowledge Graph]

ERA in house data



TEN-T Trans-European Transport Network

railML

ProRail

INFRABEL Right On Track

OpenStreetMap

INSPIRE Infrastructure for Spatial Information in Europe

FEDeRATED NETWORK OF PLATFORMS

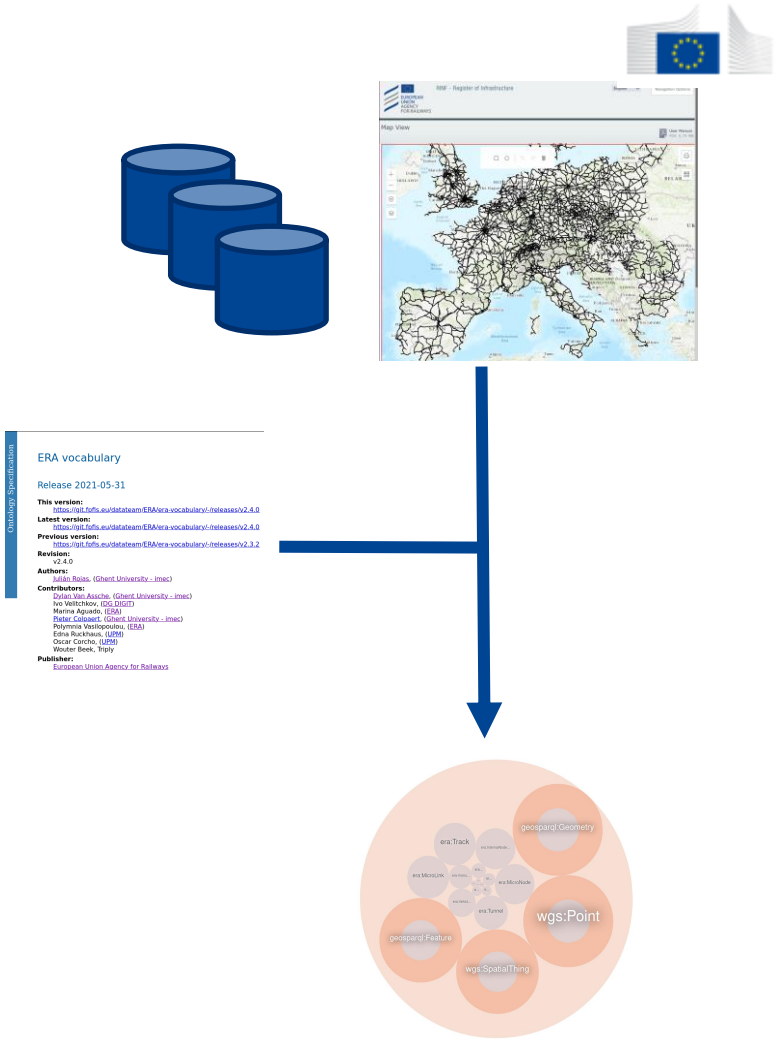
towards seamless federated MULTIMODAL data exchange

ERA's current environment

ERA collects railway infrastructure data from all EU member states

ERA provides semantic definitions in the form of an ontology

ERA publishes a Knowledge Graph that brings semantic interoperability across multiple data sources

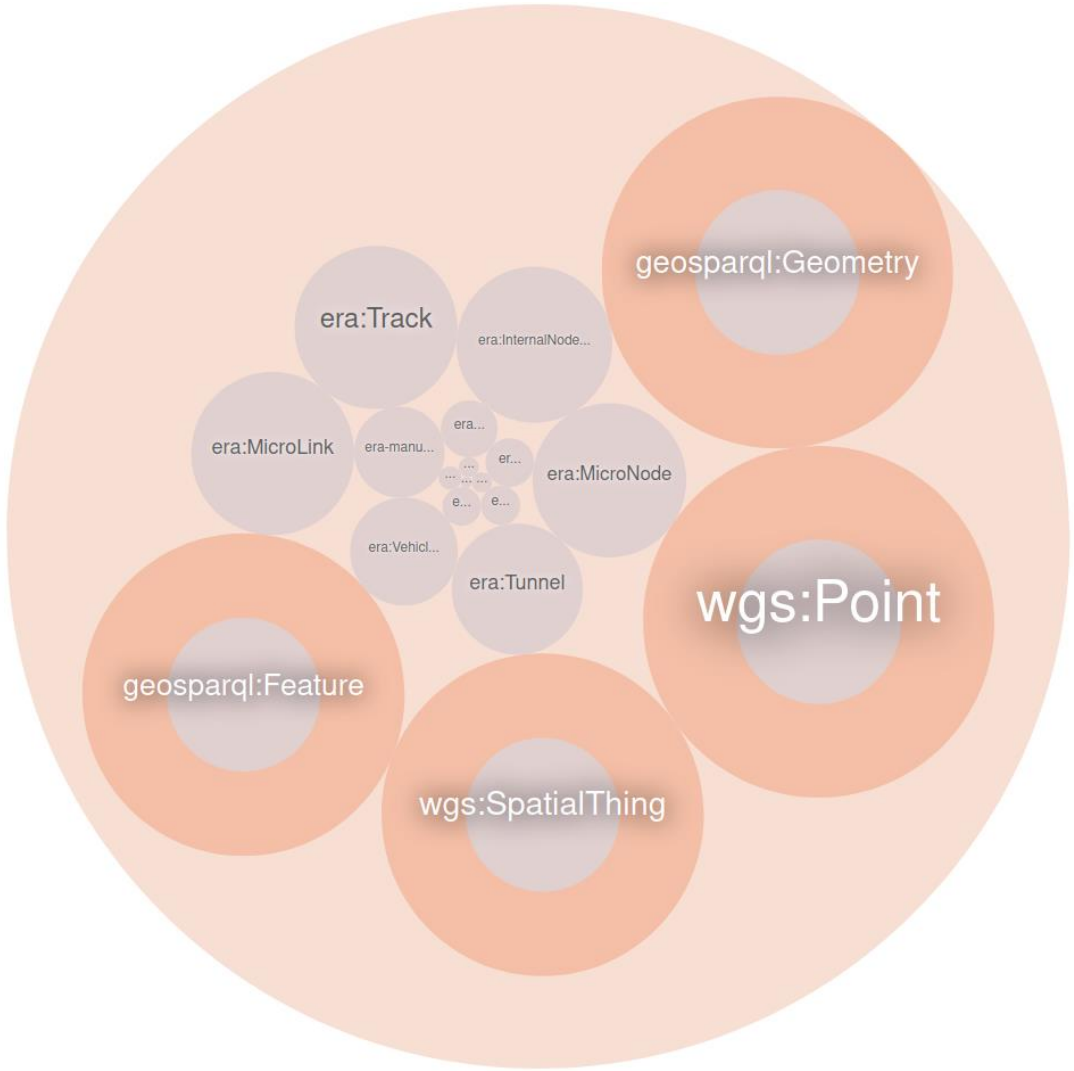


EUROPEAN UNION AGENCY FOR RAILWAYS

ERA Knowledge Graph

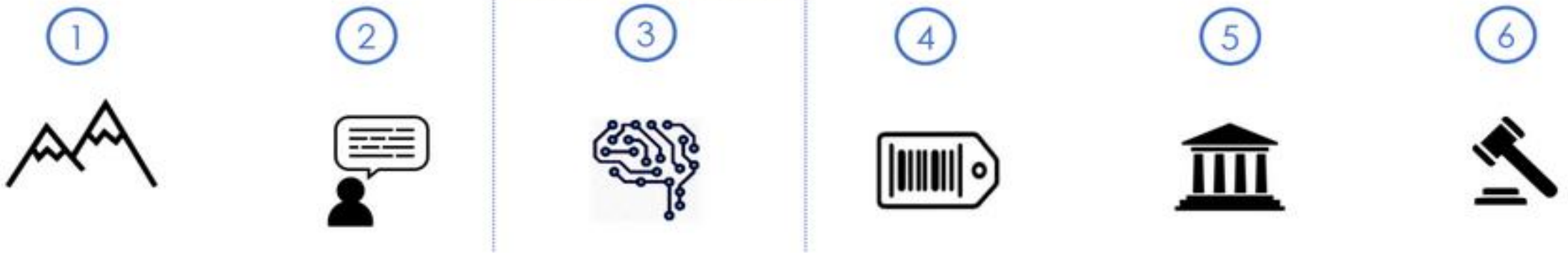
https://linked.ec-dataplatform.eu/sparql

- A total of ~30 million data statements (aka triples)
- +270k Track segments described
- +50k stations (aka Operational Points) described
- +50k geo-referenced objects (lat/lon)
- +2k Vehicle Types described
- 27 countries covered



EUROPEAN UNION AGENCY FOR RAILWAYS

Agency´s Six Pillar approach towards Data Centricity



Vision and Principles	Use Case Matrix	Architecture and Pipeline @data.europa.eu	Data Assets Management	Data Culture	Data Governance
100% Completed in the ICT & Data Strategy	On going to discuss RCC & Matrix	100%	Ongoing end March 2023	Ongoing InHouse Training & Sector Innotrans Workshop	Ongoing - Project Kickoff March 2022 with TopQuadrant

* Based on Maturity model by Enterprise Knowledge Graph Foundation

EUROPEAN UNION AGENCY FOR RAILWAYS

The Key principles

Principles to unlock access to Agency´s data for staff and sector benefit:

FAIR principle (make your data Findable, Accessible, Interoperable and Reusable)

The once-only principle an e-government concept that aims to ensure that citizens, institutions and companies as data providers must provide information to the authorities and administrations only once.

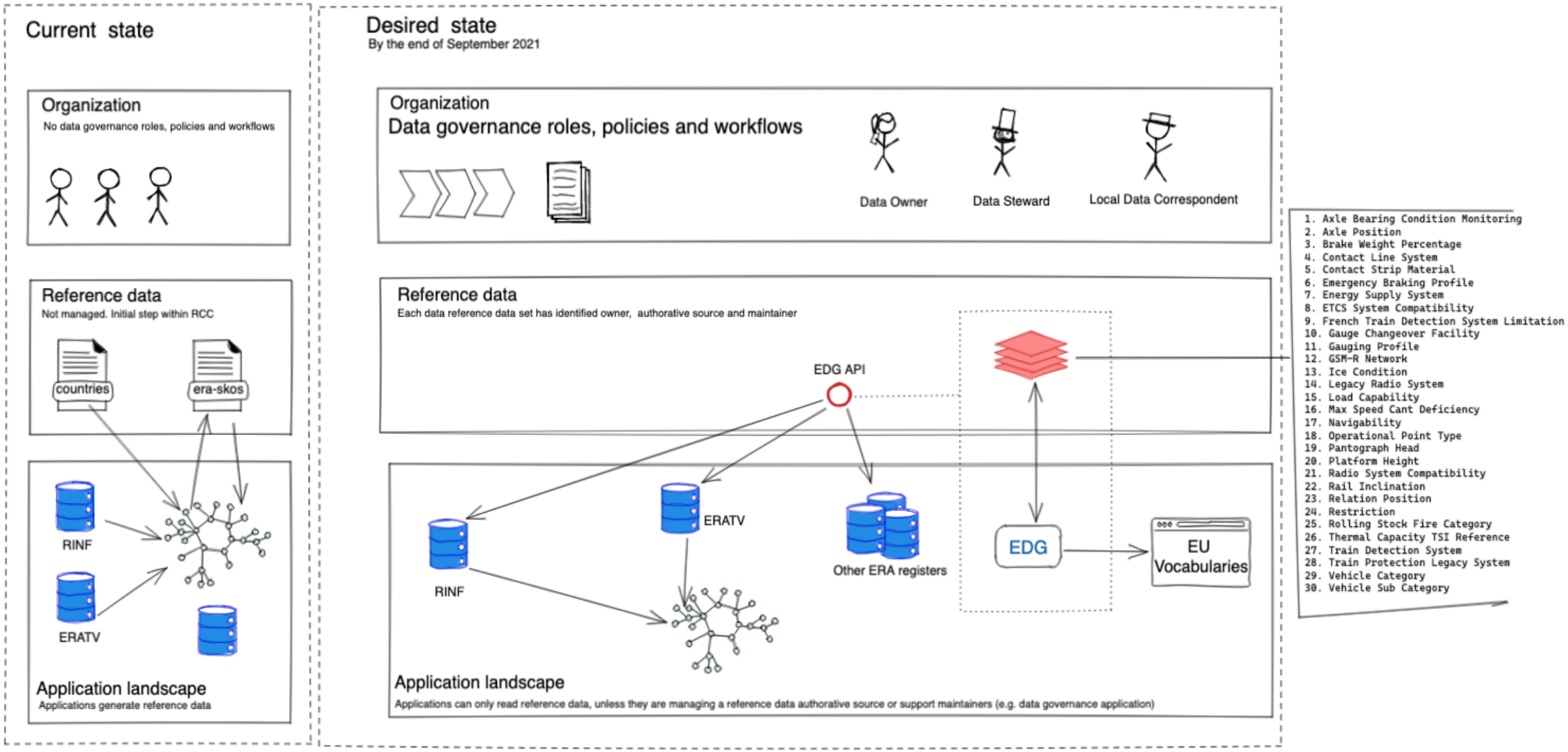
The adoption of these principles is aligned with the European Commission Data Governance and Data Policy and the European Interoperability Framework.

to implement the ‘need to share’ principle of the Von Der Leyen Commission we need to remove obstacles to internally sharing, combining and reusing the Agency´s data assets.

EC data policies are to be implemented on a ‘comply-or-explain’ basis.

EC data policy will impact Member States and our stakeholders through the Data Governance Act and the Implementing Act on high-value datasets for mobility (according to art. 13 ad Annex I of the Open Data Directive (2019/1024)).

Towards Governance and better Management of our Reference Data



The Route Compatibility Check use case

European Register of Authorised Types of Vehicles (ERATV)

Types of railway vehicles authorised by ERA or the Member States

Application

Data Base

Data models are hidden in the application layer

No connection

Application

Data Base

Register of Infrastructure (RINF)

Register of infrastructure, stating the values of the network parameters of each subsystem or part subsystem concerned

ERA Ontology

Semantic Vocabulary
Transformation in a common understanding language

Business value
Provider interest on sharing the data once and in a re-usable manner (once only principle)

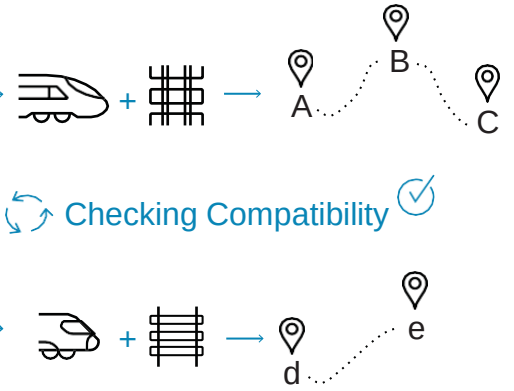
ERA Knowledge Graph

First step to extract "operational" value from ERA Registers

Route Compatibility Check

Find and analyse the information for the network topology and the vehicles...

...to automatically expose all the potential routes where a type of vehicle is technically compatible and able to run



Users

The tool provides support for the planning activity within the operational railway cycle via a web app, a simple user interface consuming the data of a knowledge graph

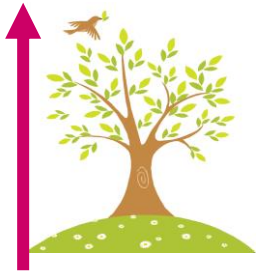
EUROPEAN UNION AGENCY FOR RAILWAYS

A real data story about the FAIRification of Mobility Register based information from a public institution **into** a building block of the EU ***FAIR*** Mobility Data Space

Common European Mobility Data Space

Mobility

Single European Railway Area Convention
Brussels
safe - connected - affordable



High Value datasets for mobility

EUROPEAN UNION AGENCY FOR RAILWAYS

Dissemination

SEMIC 2020

EUROPEAN UNION AGENCY FOR RAILWAYS webinars

European Data on Rails: the Linked Data Project

28 October 2021 16.00 [CET]

ENDORSE.

THE EUROPEAN DATA CONFERENCE ON REFERENCE DATA AND SEMANTICS

KGC

The Knowledge Graph Conference

Speakers

David Dowey

Head of Data, Information and Knowledge Management Sector in DG-DIGIT

Ivo Velitchkov

Linked Data Consultant and Trainer working with DG-DIGIT

Henk Mulder

Head of Digital Cargo at IATA. Team lead for the Semantic Modelling in DTLF

IATA DIGITAL TRANSPORT & LOGISTICS FORUM

EUROPEAN UNION AGENCY FOR RAILWAYS

Making the railway system work better for society.

Follow us on ERA_railways

Discover our job opportunities on **era.europa.eu**