Building Trustworthy Railway Data

Achievements from the DIM Project at Bane NOR

30.05.2025 Linnea Olsen

This is Bane NOR

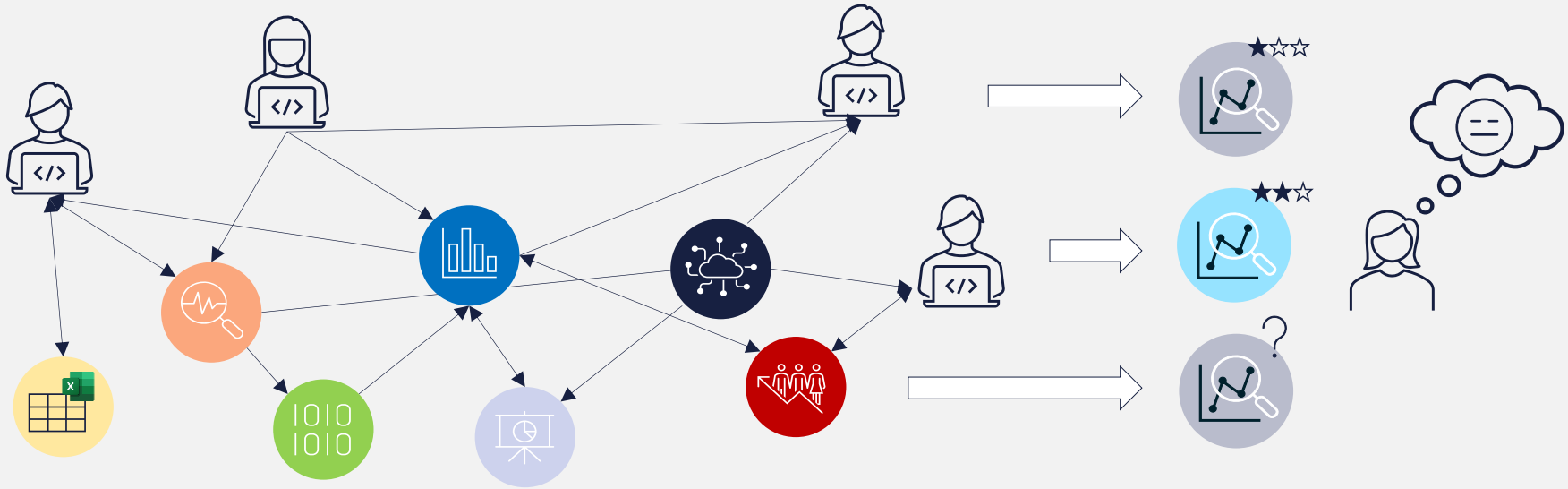
Bane NOR is responsible for operating, maintaining and building the railways in Norway. We are responsible for the infrastructure on the railway throughout the country

- We are a state enterprise, owned by the Norwegian State through the Ministry of Transport, and established as part of the Railway Reform on 1 January 2017.
- We are responsible for the Norwegian rail infrastructure. Bane NOR is in charge of operations, maintenance and construction of railways throughout the country.
- This includes **4200 kilometres** of tracks, **335 stations and stops** and more than **4300 properties.**
- We have a **staff of 5.200 employees**, and our main office is located in Oslo.

Why DIM?

Challenges with the data foundation

Datainnsamling og -distribusjon	*Bruk av data*



- ! Manual collection of data and information from a variety of sources.
- ! Data is delivered uncoordinated from various parts of Bane NOR to data consumers.
- ! High risk of errors, inaccurate, and outdated data.

BANE NOR

Purpose and deliveries

Purpose: Exchange of information about the traffic characteristics of the railway infrastructure and train routes in various formats

Use cases:

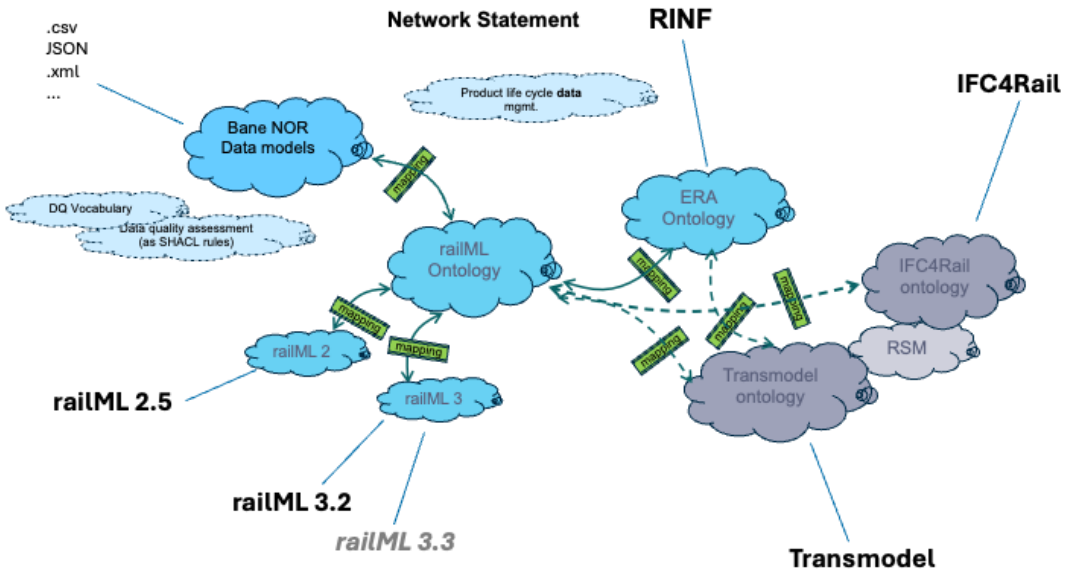
Run time calculation	• September 2025
RINF	• March 2026
Conflict analysis	• April 2026
Future infrastructure	• May 2026

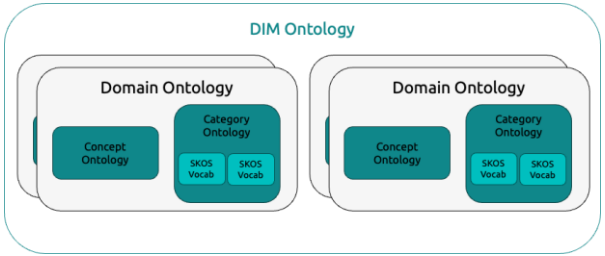
Identified future usecases:

Capacity planning	• To be decided

We are building a knowledge graph

Railway sector models





DIM coordinates models with railML.org for the reuse of mechanisms and efficient certification.

"think big, start small"

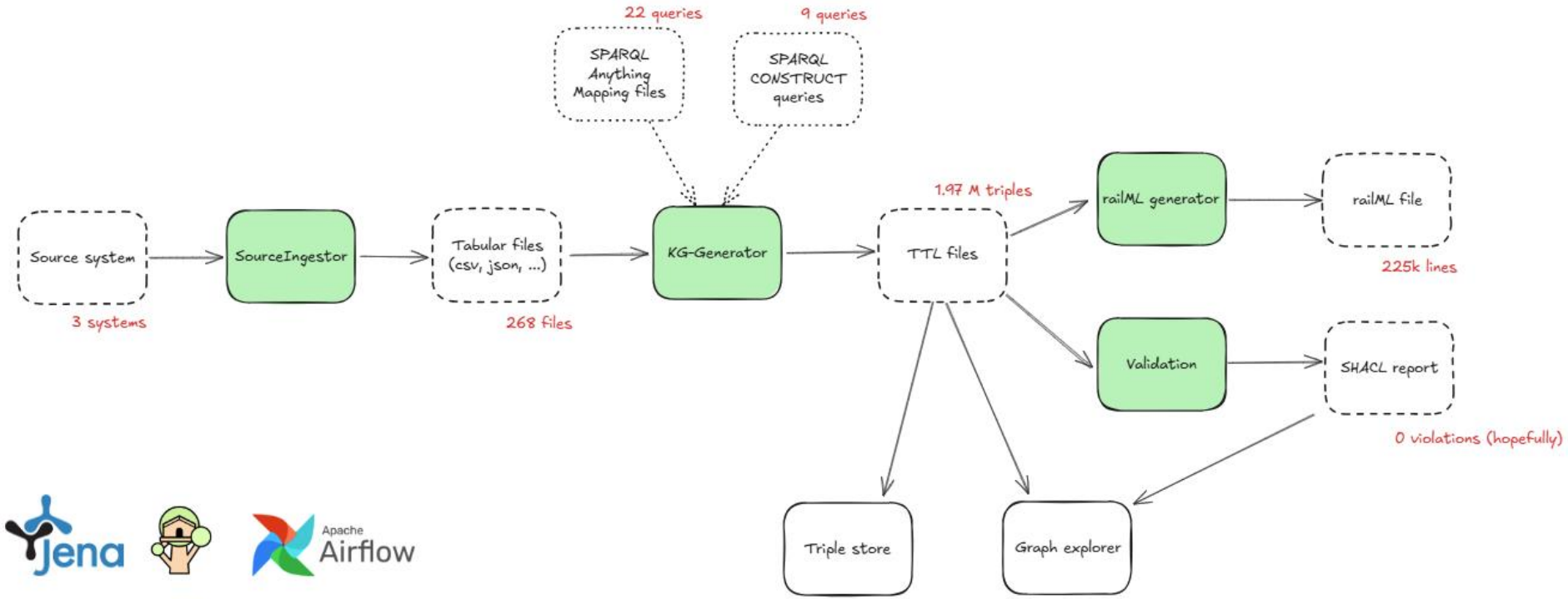
BANE NOR

What We’ve Achieved So Far

What DIM Have Achieved So Far

- **Agile** development: **monthly push** to production
- **Data model established** based on railML & ERA ontology
- **Dataset with topology produced** with functional object types and data based on TRASÉ, Linjedatabasen and Banekart
- **railML 3.3 data production** in place (the first iterations are in place)
- **Knowledge graph** with 1.97 million tripples
- **Continuous feedback loop** with The Norwegian Railway Directorate

Runtime process



What's in the knowledge graph

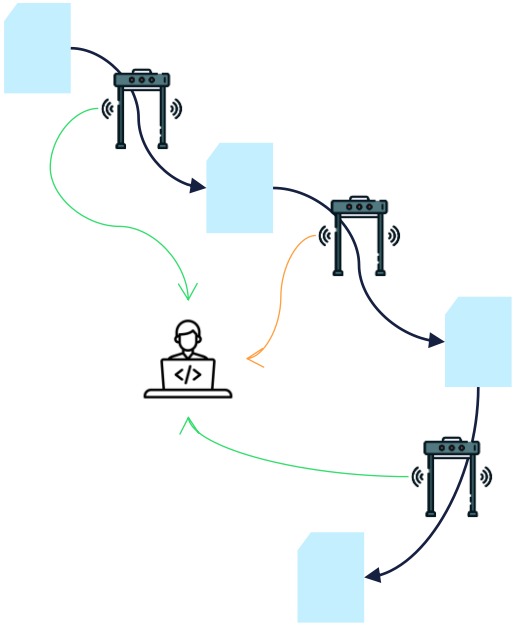
Class	Count	Class|	Count	Class	Count
gsp:Geometry	96169	rtm:RTM_NetElement	9113	era:PlatformEdge	924
rtm:NetEntity	84485	era:Track	8084	railml3:BufferStop	622
rtm:GeometricCoordinate	71482	rtm:Relation	7894	railml3:OperationalPoint	522
rtm:SpotLocation	63545	railml3:Length	7789	railml3:LevelCrossingIS	506
railml3:Designator	29642	railml3:Track	7789	rdfs:Container	357
era:InfrastructureElement	27194	era:KilometricPost	7774	time:Instant	158
rtm:IntrinsicCoordinate	18226	era:Signal	7061	bno:BaneSubnetwork	33
rtm:AssociatedNetElement	15969	era:LevelCrossing	3267	rtm:AreaLocation	26
railml3:GradientCurve	14392	railml3:SwitchIS	2984	railml3:Level	3
rtm:LinearLocation	12560	railml3:SignalIS	2585	railml3:ElementState	2
era:LinearElement	9113	railml3:SignalConstruction	1765	rtm:GeometricPositioningSystem	2
rtm:AssociatedPositioningSystem	9113	railml3:HorizontalCurve	1679	railml3:Network	1

So... What about data quality?

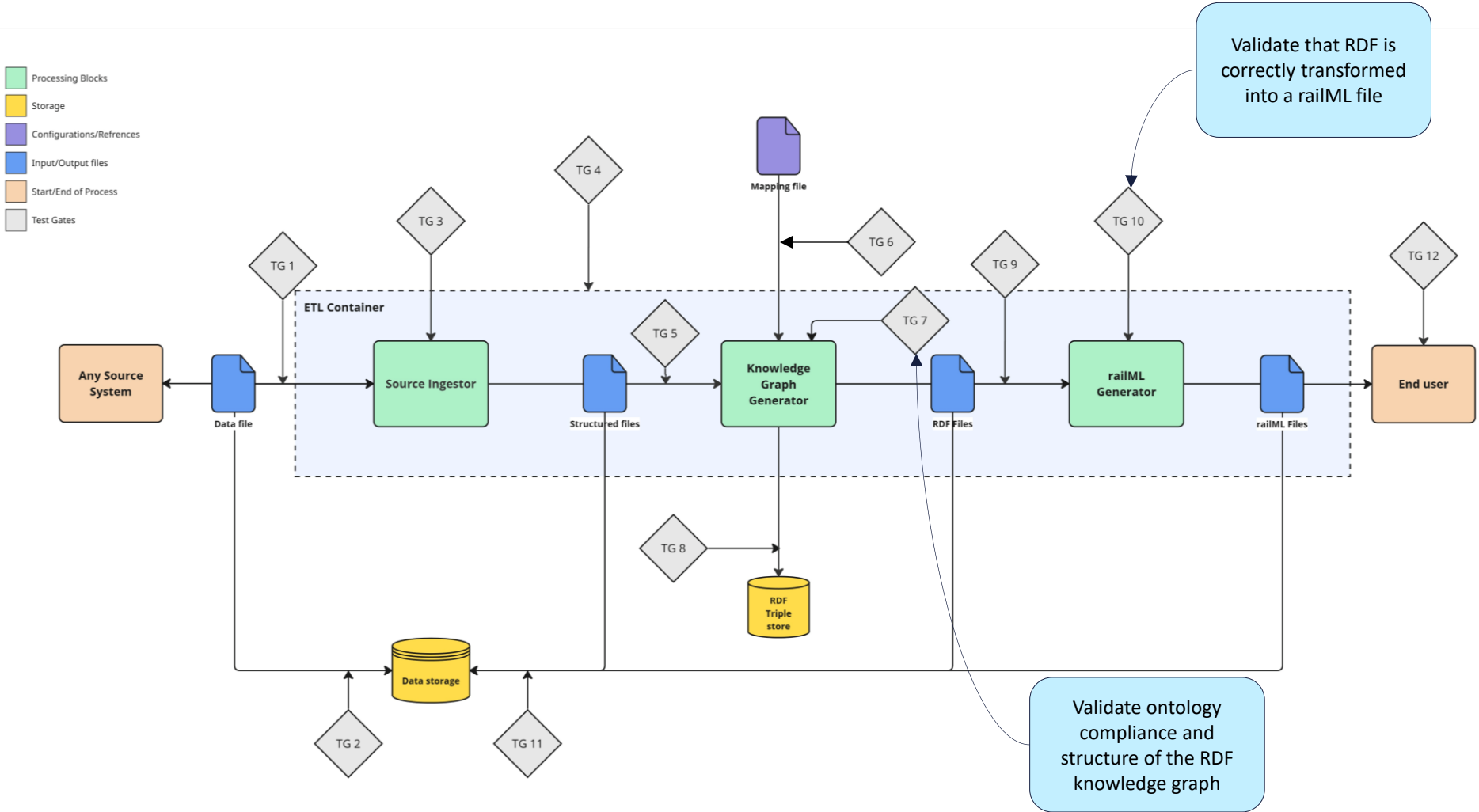
What We’ve Achieved So Far

Continuous Testing with Test Gates

> *Quality is not an act, it's a habit*



- *Building quality into both the project and process*
  - Test gates serve as multiple feedback loops
  - Part of the deployment and development process
  - Implemented basic automated tests


Processing Blocks
Storage
Configurations/Refrences
Input/Output files
Start/End of Process
Test Gates
Validate that RDF is correctly transformed into a railML file
TG 1
TG 2
TG 3
TG 4
TG 5
TG 6
TG 7
TG 8
TG 9
TG 10
TG 11
TG 12
Mapping file
ETL Container
Any Source System
Data file
Source Ingestor
Structured files
Knowledge Graph Generator
RDF Files
railML Generator
railML Files
End user
RDF Triple store
Data storage
Validate ontology compliance and structure of the RDF knowledge graph

This isn’t just about technology.
It’s about improving and modernizing the railway so more people can take more trains.

By building trust in data, we lay the foundation for better decisions, more efficient operations, and a more sustainable future