PUBLIC SUMMARY OF DELIVERABLE

D6.3 – Proactive Rail Infrastructure Release 1

Project Acronym	TT
Project Title	Transforming Transport
Grant Agreement number	731932
Call and topic identifier	ICT-15-2016
Funding Scheme	Innovation Action (IA)
Project duration	30 Months [1 January 2017 – 30 June 2019]
Coordinator	Mr. Rodrigo Castiñeira (INDRA)
Website	www.transformingtransport.eu

Document fiche	
Authors:	Aida Carnicero Costa [INDRA], Phil Flynn [Thales UK], Soledad Hurtado [INDRA], Joanna Thornton [Thales UK]
Work Package:	WP6
Task:	T6.2-T6.3
Nature:	R
Dissemination:	PU

This document is a public summary of a confidential deliverable of the TT project. It serves as a summary of the release 1 demonstrators and provides links for external actors to connect to the TT pilot leaders if they are interested in more information.

ACKNOWLEDGEMENT

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement № 731932.

Motivation

TransformingTransport (TT) demonstrates, in a realistic, measurable, and replicable way the transformations that big data can bring to the mobility and logistics market. TT is structured into 13 different pilots in 7 pilot domains, which cover areas of major importance for the mobility and logistics sector in Europe. The **Proactive Rail Infrastructure** domain focuses on the predictive maintenance of railway assets.

The two pilots of this domain are designed to look at the following goals using big data techniques:

- Improving safety;
- Improving reliability;
- Improving cost efficiency;
- Improve capacity.

Both pilots strive to provide functionality that predicts the failure of mainline railway assets with sufficient foresight that preventative maintenance can be scheduled and performed. Improving the maintenance of rail assets consistently and predictably improves the safety case and other sub goals of the project such as cost efficiency and minimising disruption.

Understanding the history of an asset could also lend itself to predicting the asset's future behaviour. By using big data techniques to process the various sources of disparate data this will allow the data scientist to perform analysis and visualise correlations. These insights will be presented to the infrastructure owner to allow them to perform a basic forecast of asset behaviour and plan an intervention if required. The validated results form the basis of a strategy for the data scientist to analyse the data in greater depth with the aspiration of predicting asset patterns.

The two pilots for this programme of work utilise infrastructure owners' data, namely Network Rail and ADIF for the lead and replication pilot respectively. Each of these Infrastructure owners has differing business goals and ambitions but can be generalised as improving operational efficiency, customer experience, and new business models. The infrastructure owners for both pilots reside in two different physical locations.

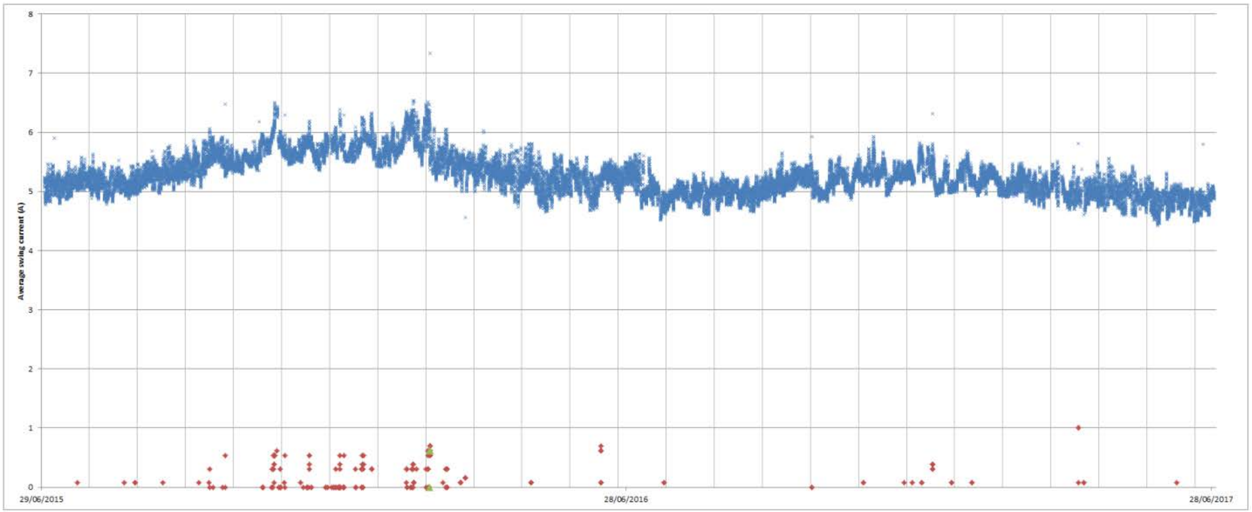
The key difference between the locations, other than geographic location itself, is the type of railway line each operator is using. The Network Rail railway line of interest is classed as 'mainline', whereas the ADIF railway line is categorised as a 'high-speed' line.

Both pilots have produced demonstrators using real historical data provided by infrastructure owners. The demonstrators will prove the ability to work with large quantities of historic data as well as handling 'live' streaming data. They demonstrate the business benefit of the application of big data tooling to rail infrastructure data in terms of increasing life expectancy of assets, reducing failures, better scheduling maintenance activities thus saving costs and improving safety.

Predictive Rail Asset Management

The Predictive Rail Asset Management pilot has advanced the data analytics of 4 asset types namely, points machines, Track Circuits, Train Track Interface and Overhead Lines. Most progress has been made in the analysis of point machines and for this reason the Release 1 has focused on the identification of health trends for point machines, specifically the health attributed to a feature known as stall headroom analysis. The maximum effort that the point machine is able to exert has been calculated and can be used to compare with future swings to determine a score of risk or health i.e. how close to stalling is the point machine and how much risk is there in continuing to operate this device without maintenance.

Research has thus been focussed on identifying relationships between swings and actual faults with the intention that if trends can be identified that lead to a fault then the same trend can be used to prevent a fault from occurring. In the figure below, the blue 'x' series represents average swing current over time, the red diamond series represents alarms automatically raised, and the green triangle series are actual faults.



Swing current over time and alarms

Development of the prototype for this deliverable is still in its early stages however it is felt that even the information presented at this time is of great use to the end user. The prototype hopes to provide reassurance to human operators in their business cases for prioritising maintenance of one asset over another. It does this by not only providing high level summaries and key values, such as the remaining useful life of an asset, but by also providing the data behind that calculation so that confidence and reasoning can be assigned to a decision.

Contact for more information about the pilot:
Phil Flynn, phil.flynn@uk.thalesgroup.com

Predictive High Speed Network Maintenance Pilot

The Predictive High Speed Network Maintenance pilot has focussed on three use cases namely: track profiles, point machines, and traffic optimization. However, for Release 1 this has focussed on predictive maintenance of both track profile degradation and point machine degradation. Track degradation has been analysed through different acceleration measurements taken through regular inspections. Point Machines have been analysed using data such as train circulation, maintenance works and static track information.

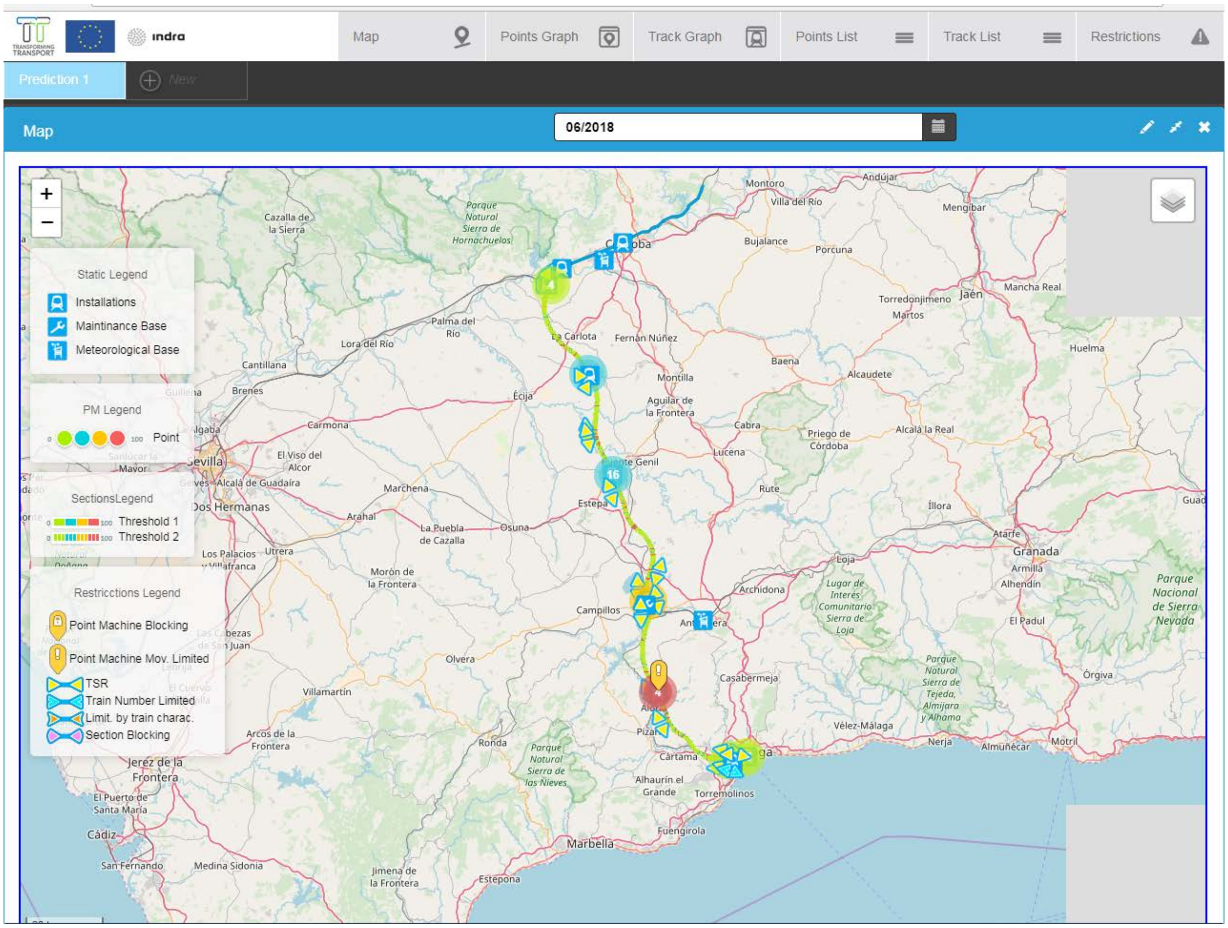
Very promising results have been obtained regarding predictive data of the track sections with a score of more than 85%. The resultant test score for predictive data of the point machines has been <85% and is understood to be due to the variables that are used as an input for these models are not correlated adequately with the variable target. As more data characteristics are included in the analysis it is expected that the test score will significantly improve during the next phase of analysis.

Good progress has been made, demonstrating the efficacy of varying statistical methods and machine learning models. There is more work to be done to further improve these methods but the success rate is already high enough to be useful and focus resource and effort, if not trusted fully.

Significant research has gone into understanding the data provided in great detail, starting with the relationships between dynamic and geometric track profile inspections; exploring threshold intersections; distributions of data within data sets; and frequency, correlation, signedness and spatial analysis, and ending with machine learning algorithms. A number of dashboards have been also created to display the vast array of data in an easily digestible and navigable format whilst still being very powerful.

In addition, a new module has been developed to generated new knowledge to improve the use of the infrastructure from traffic management point of view, operational restrictions are generated automatically based on predictive information about incidents identified in the infrastructure as a result of applying Big Data technics.

The figure below shows one of the view of the dashboard where the information is located on a map. Specifically, the information that is represented is related to predictions about incidence in tracks and point machines and Operational Restrictions, each one is represented with a particular icon, the colour is determined according to a value (depending of the risk a probability of the incidence).



Failures Predictions Map – Track Profiles & Point Machines

Contact for more information about the pilot:
Aida Carnicero Costa, acarnicero@indra.es
Marina Soledad Egea González, msegea@indra.es