# USE OF ONTOLOGY FOR DATA INTEGRATION IN A DEGRADED MODE SIGNALLING SYSTEM

CHRISTOPHER MORRIS & JOHN EASTON
Birmingham Centre for Railway Research and Education, UK

ABSTRACT

Every week many thousands of delay minutes are accrued on the UK railway, a significant proportion of which can be attributed to failures of signalling sub-systems, such as track circuits. Signalling failures are both expensive for the infrastructure provider (as knock-on delays to services build up rapidly) and a source of frustration and delays for passengers, leading to increased dissatisfaction with the industry as a whole. A degraded mode signalling system, offering less functionality than the main system but via alternative channels, can help to mitigate these failures and allow railway operations to continue whilst the main signalling system is repaired. The use of ontology was a key enabler in this project, making it possible to draw together data from multiple sources and infer meaning from the data supplied. National scale signalling data was used with the ontology without impediment.

*Keywords: ontology, degraded mode signalling, integration.*

## 1 INTRODUCTION

In the financial year 2015/16 Network Rail spent £106,008,691.22 [1] compensating operating companies for unplanned delays. Every week many tens of thousands of delay minutes accrue on the railway, and of these many thousands are attributed to signalling failures. The delay caused to passengers as a result of such failures degrades customer experience and contributes to negative public perceptions of the railway. The Combined Positioning Alternative Signalling System, (COMPASS), is a system to provide a degraded mode signalling system with the primary objective of reducing the impacts associated with failures of the main signalling system. Additional benefits of such a system include improved vehicle positioning relative to the existing track circuit-based system, leading to improved passenger information and the potential for use as a low cost primary signalling system on lightly used lines.

Implementing this system required the combination of several different data sources and the client stated it would be beneficial if a system could show ease of integration with other data sources as they became available. There is a consensus in the literature [2]–[4] that ontology can be a good tool for data integration with in the rail domain, and it was employed in this study.

Ontology has different definitions dependent upon the field in which it is used. In philosophy the definition given by the Oxford English Dictionary: “the branch of metaphysics dealing with the nature of being” applies. This definition, whilst useful, differs somewhat from that used within computer science. In computer science ontologies are applied as a means of storing not just data, but information, that is data with meaning. In this domain the definition normally cited is:

*“An ontology is an explicit specification of a conceptualization”*

This definition alone requires further explanation, which is provided by the author, who defines a conceptualisation as:

*“The objects, concepts, and other entities that are presumed to exist in some area of interest and the relationships that hold them”*

WIT Transactions on The Built Environment, Vol 181, © 2019 WIT Press
www.witpress.com, ISSN 1743-3509 (on-line)
doi:10.2495/CR180191

This use of "Conceptualization" is in turn taken from [5], who state that it offers an explicit view of the world. This is helpful in computer science because it allows software to work not just with data, but with knowledge. This knowledge in turn adds context to the information thus making information exchange easier.

Working with information, as opposed to data, is useful for integration because once one stores the meaning of the data alongside the data itself it is much easier to integrate that data.

### 1.1 Previous ontology-based data integration in the rail domain

As reported in [6], there is a need for better data integration within the UK rail domain: "*Over 130 information systems maintained by approximately 20 suppliers were in operation in 2011. Maintaining individual legacy systems is expensive and inefficient. Information cannot be shared or exploited efficiently and this inhibits whole-system approaches for technology-based improvements.*"

Ontology has been proposed as a means of improving data integration in the wider information science domain since [7]. At first this was hampered by the inability of the available hardware to put into practice the theoretical concepts, however as hardware has progressed, there now exist several web-scale ontologies, for example DBPedia [8] and whilst care must be taken when designing ontologies not to add unnecessary complexity hardware limitations have been largely overcome.

Other domains have used ontology for data integration for some time:

- Biomedical Research and Bioinformatics [9];
- Media [10];
- Process and petrochemical plant [11].

In all these domains use of ontology is routine and progress now centres on improving the implementation, not starting it.

In the rail domain a number of projects have proposed ontologies, most recently the IT2rail [4] project, however whilst these have produced demonstrators none of these have become industry wide solutions.

## 2 COMPASS SYSTEM DESIGN

In the tender request the customer (Network Rail) gives the purpose of the system as: *"to automate the manual processes involved in Temporary Block Working"*. Temporary block working is the current fall-back procedure whereby trains are allowed to pass through sections of track on which the signalling system has failed. Several consortia are creating products in response to this tender and it is possible that more than one will be selected.

After the project started the client altered the specification, notably the amount of automation expected was reduced and the requirements altered such that there must always be a human in the loop. Furthermore, it was clarified that the system would at no time be an *"Alternative Signalling System"*, rather it was to provide "Degraded Mode Working" and would never be used outside of those circumstances.

The COMPASS demonstrator aims to show that ontology can form the core of a fall-back signalling system, reducing the impact of signalling system failure, keeping trains moving even when the main system has failed.

The demonstration scenario assumed new equipment could be placed in two physical locations: In Cab, and in the Rail Operating Centre (here on ROC – a control centre). It should be noted that cost is also important, this system must be less expensive than a main signalling system, as such certain compromises are required. Guidance from Network Rail suggests that points (also known as switches or turnouts) should not be remotely controlled, rather the lie of the points should be detected and trains only routed where that permits. This reduces the safety criticality of the software and thus the level of (expensive) certification required. In particular the infrastructure operator wished to avoid the need for SIL level 4 certification, as such it was also requested by Network Rail that the system not issue trains authority to move without manual intervention. It is expected that improved knowledge of train location will also make possible improved passenger information.

### 2.1 Client requirements

The requirements from the client were set out in a tender documents and modified verbally, they may be summarised as:

- *"The replacement of temporary block working with a more efficient solution"*. Originally this was to be the automation of temporary block working, however in light of guidance from the infrastructure manager there will still be a manual element;
- The system shall be deployed in a limited number of predetermined areas, were signal failures risk the greatest impact. This also differs from the original specification, which required scalability up to providing a national train position database. The area need not be a plain line, but can and likely will have multiple entrance and exit signals. The area can be bi or uni directional.
- The accuracy of train location data and hence arrival times estimation should be better than is available from the existing track circuit based systems;
- *System shall be separate from the existing signalling system*;
- Be resilient to cyber-attack, physical vandalism and deliberate sabotage;
- "maintain, in memory, train location to a given time stamp for reference purposes";
- Ease of adding other datasources when they become available would be an advantage;
- *"Adopt a multi-layered solution for train location"*. That is take train position information from multiple sources.

In order as to achieve the above objectives, in particular, the adoption of a multi-layer train location solution and the improved location accuracy it is necessary to bring together data from multiple sources. Use of an ontology architecture would minimise the development effort required to accomplish this.

### 2.2 Demonstrated scenarios

Two demonstrations were performed, in response to two different operational scenarios: The first scenario demonstrates normal operation, with the ontologies and supporting tools connected to and processing data-feeds from the UK infrastructure manager. This first scenario could be extended to include providing improved customer information. The second scenario focused on degraded mode operation and showed how vehicles could transition to COMPASS signalling and pass along tracks with failed signalling.

#### 2.2.1 Normal operation – quiescent state

In the first scenario the railway is assumed to be operating normally, with the system in what is referred to as a *"Quiescent State"*. In COMPASS this is used to demonstrate the system monitoring the locations of services already running over the network, ready for use as the *"base state"* when a failure occurs. The train location data used in this scenario came from the Network Rail open data feeds, which were used in conjunction with a static map of the network, provided by an industrial partner. These two diverse data sources provide good examples of typical industry datasources that need integration in a functioning industry wide system.

In normal operation the demonstrator tracks the locations of all the trains in the network, allowing the provision of better customer information as well as maintaining a model of the running system in readiness for degraded mode operation. This demonstrates that should a signalling fault occur the system would be able to respond appropriately. Tracking of locations is illustrated by displaying those positions on a map, annotated with the headcodes of the trains being tracked. A proposed extension to this scenario calls for the display of metadata for the train – projected arrival time for example – along with the headcode. The use of ontology in this project also makes easier displaying such things as live updated possible connections, in light of the arrival time.

The quiescent state demonstrator monitors vehicle movements across the entire British rail network, and as a result provides solid evidence of the ability of ontology-based systems to scale to the data volumes and rates needed in a nationwide industry deployment.

#### 2.2.2 Degraded mode

Degraded mode occurs when the underlying signalling system is not operating as intended, regardless of reason, and signaller chooses to use the COMPASS system to manage the train through the area, known within this project as the "area of interest". It is expected, though not required, that the area will be relatively small and when the end of the area is reached the train returns to the control of the primary signalling system.

## 3 COMPASS SYSTEM IMPLEMENTATION

For demonstration purposes all subsystems were located in the same room and the radio link in particular was considered beyond the scope of this project; it was assumed that an IP data link existed between the vehicle and the control centre. The high-level system architecture is shown in Fig. 1 Full System architecture.

The individual subsystems are described below, first those which would be housed within the ROC.

#### 3.1.1 VLS Centre Comms Virtual Line-side Signalling

The VLS system handles communications between the in-cab elements and the control centre elements. This is a pre-existing commercial product modified for this project.

#### 3.1.2 DMC – Diverse Monitor and Control

This component acts a central control block bringing all the other subsystems.

#### 3.1.3 RBC – Radio Block Control

This is a standard part of a moving block signalling system. It is a pre-existing commercial product which handles safety critical aspects of the system.

3.1.4 PCIDR Track based Point Control Inhibit, and Detection Repeat (PCIDR) and Control

This "isolates" the points (switches). Whilst the system is in operation the points do not move. The current position of the points is detected and fed back to the rest of the system, which will only signal trains to pass in directions allowed by the points.

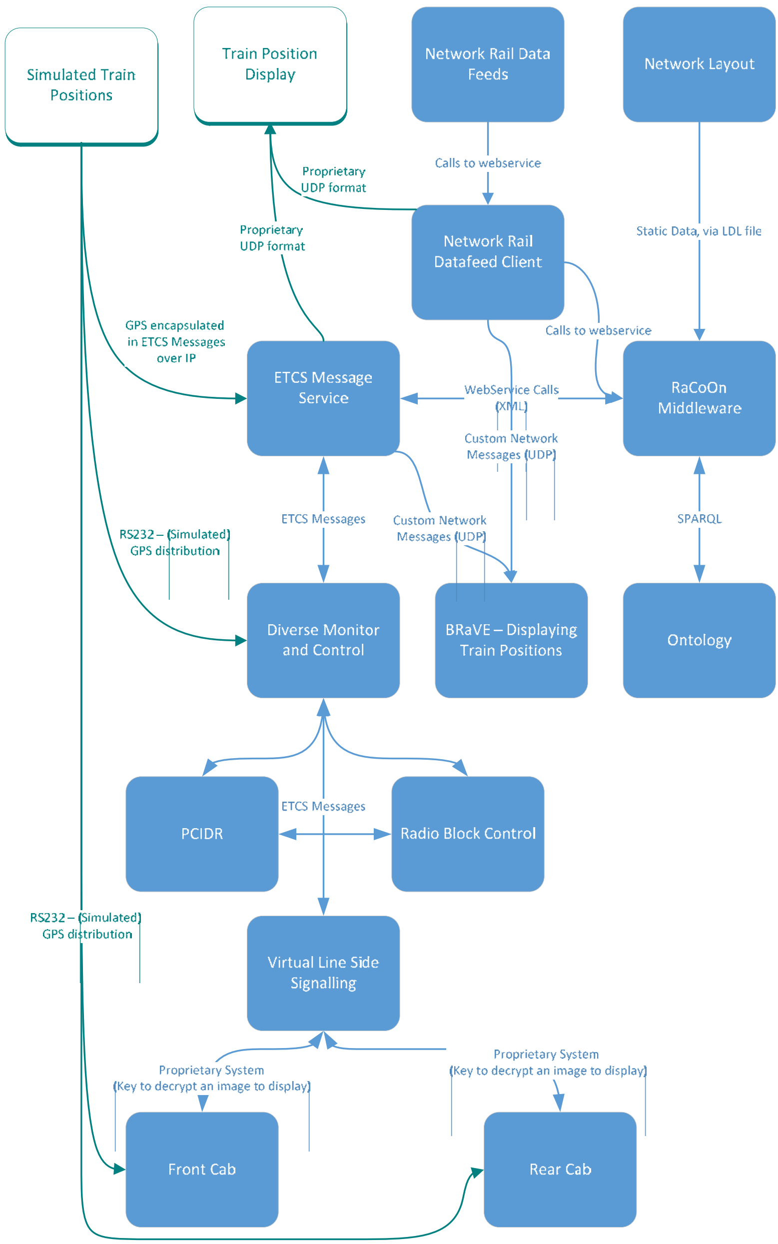


Figure 1: Full System architecture.

#### 3.1.5 STiR interlocking

The STiR interlocking is a simplified version of a conventional signal interlocking, since the points are isolated.

#### 3.1.6 RaCoOn the railway core ontologies

The RaCoOn comprises a number of subsystems and is used to integrate data from a number of different external and internal sources, alerting the other components when a train is approaching. This in turn comprises a number of sub-systems:

- The ETC Message service; This receives ETCS messages and triggers appropriate changes to the ontology;
- The RaCoOn Middleware, which acts as a buffer between the others systems and the datastores;
- The datastores;
- A system to display train locations on a map. In this case an existing rail simulator, BRaVE which discussed further by [12].

### 3.2 Simulator

In order as to demonstrate the system it was necessary to use a rail simulator to recreate the effect of the train being in motion. The simulator generated coordinates representing the position of the front of the simulated train. These were sent to the in-cab signalling equipment in the same format as would have been used were the data coming from a real GPS receiver mounted on a train. The same position data was then sent to RaCoOn, to simulate the train sending position data.

### 3.3 In-cab elements

The in-cab elements of this system is a cab mounted display which provides signal aspect information to the driver. This system is part of the virtual line side signalling system. The GPS simulator located in the cab was simulated for the demonstrators. Where a train is fitted with two cabs (front and rear) the project proposed (but did not implement) fitting a second GPS transmitter and VLS system to provide train integrity information.

### 3.4 The role of ontology

In the compass demonstrators the primary role of the ontologies is as the basis for data integration across the diverse data sources. There are many data sources for this system and it is one of this project's objectives to show that the project partners could work with data from a multitude of heterogeneous sources, not limited to those included in the initial design. Ontology is also used for classification in this project, for example for the classification of nodes, which are classified by sub-type. Nodes can be any of the many types of object located on the track such as: simple nodes, points nodes or signals. Most significantly they can be the signals that mark the start of the area of interest; this information is used to determine when to trigger degraded mode operation.

Some decision making was also abstracted to the ontology, in particular, whether or not a train was approaching the start of the COMPASS area of interest. To be considered as approaching the train should be on the same line as the COMPASS start signal and within a certain physical distance of the signal. Where it desired to make the logic more complex, then the ontology could be changed without changing the software. It would be beneficial, and possible, to abstract further functionality to the ontology.

## 4 OUTCOMES

The project produced two demonstrators as per Section 2.2, both of which functioned as intended.

### 4.1 Demonstrator one

Demonstrator one proved its success in linking real world train positions to values from the Network Rail train describer feed to real world locations, as shown in Fig. 2.

The Network Rail train describer feed provides an update every time there a train moves from one signalling “Berth” to another. This limits the resolution of the data to that of the conventional signalling system and puts the velocity of the data at tens to hundreds of messages per minute. The ontology was capable of running queries every time such a message arrived, with the datastores hosted on a standard desktop PC; by the standards of sensor data (often many readings per second) this is not however high frequency data.

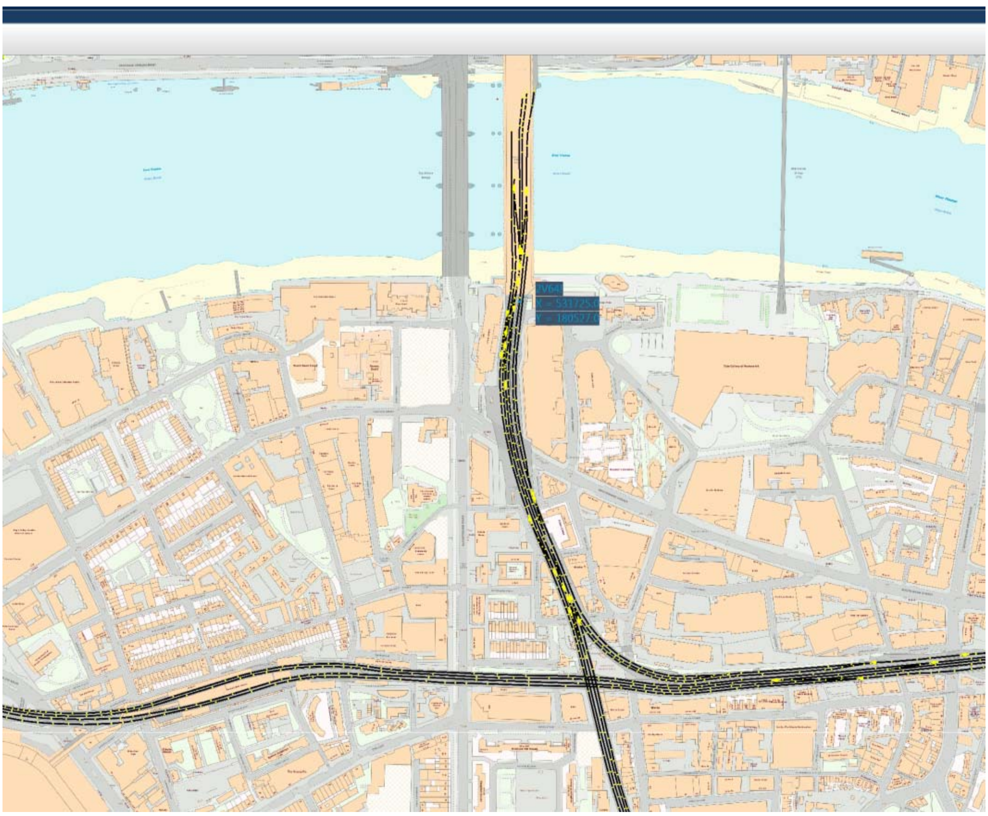


Figure 2: Demonstrator one train position display.

### 4.2 Demonstrator two

The second demonstration showed three scenarios:

- First a train moves across the simulator network area under normal signalling and its progress is displayed on a map. This demonstrates that the system can communicate internally, from the simulator to the ontology then onto the display. It further shows that the system can track the approaching train.
- In the next scenario the train drives into an area, then the signalling fails and the alternative system, known as STiR is activated. The driver is instructed, via the in cab signalling, to drive out of the area of failed signalling then to obey normal signalling once the train is clear.
- In the final scenario an area of signal has failed, a train approaches, is switched to STiR control and leaves. This was repeated with a second train.

This scenario made use of GPS data from the simulator, in lieu of GPS data from the train, however this was only simulating the movements of a small number of trains, not a complete national rail network. Processing GPS updates, to ascertain position relative to any signals of interest, at this frequency was well within the capacity of the ontology when hosted on a moderately specified desktop machine.

## 5 CONCLUSIONS

The project produced functioning demonstrators which illustrated a possible architecture for a degraded mode signalling system. This would allow for train movements in the event of a signalling failure, which was the projects objective.

The use of ontology in this project made integration simpler as data previously made available to the ontology could be reused. Mapping schedule data to the ontology had already been partially completed, so it was possible to simply complete the mapping and use the existing work. Had this been done to a proprietary format that work would almost certainly have been of no value to this or other future projects. Going forward the mapping from the train describer feed will be available for use in other projects, as will the schedule mapping. The use of ontology for storing rules in this project could be expanded upon when the project is fully implemented. This would bring significant maintainability benefits; as surrounding systems changed and the system is deployed in new locations, it is not necessary to alter the software in any way. All that would be required is changes to the rules hosted within the ontology (for example the siting distance of signals), not within the software. This can be carried out without software development experience, thus by a far larger pool of people, in this case engineers with knowledge of signalling, but not software.

## ACKNOWLEDGEMENT

This project was carried out in conjunction with Siemens AG and Park Signalling Ltd.

## REFERENCES

[1] Network Rail Infrastructure Ltd, Our information and data, 2018, Online. https://www.networkrail.co.uk/who-we-are/transparency-and-ethics/transparency/datasets/. Accessed on: 7 Jun. 2018.
[2] Köpf, H., InteGrail – Publishable Final Activity Report, 2010.
[3] Verstichel, S. et al., LimeDS and the TraPIST project: A case study – An OSGi-based ontology-enabled framework targeted at developers in need of an agile solution for

building REST/JSON-based server applications. *IC3K 2015 – Proc. 7th Int. Jt. Conf. Knowl. Discov. Knowl. Eng. Knowl. Manag.*, **2**, 2015.
[4] Gogos, S. & Letellier, X., IT2Rail: Information Technologies for Shift to Rail. *Transp. Res. Procedia*, **14**, pp. 3218–3227, 2016.
[5] Genesereth, M.R. & Nilsson, N.J., Logical foundations of artificial intelligence. *The Journal of Symbolic Logic*, **55**(3), p. 405, 1987.
[6] Technical Strategy Leadership Group, The industry's rail technical strategy 2012, 2012.
[7] Siegel, M. & Madnick, S.E., Context interchange: sharing the meaning of data. *ACM SIGMOD Rec.*, **20**(4), pp. 77–78, 1991.
[8] Lehmann, J. et al., DBpedia – A large-scale, multilingual knowledge base extracted from Wikipedia. *Semant. Web*, **6**(2), pp. 167–195, 2015.
[9] Groß, A., Pruski, C. & Rahm, E., Evolution of biomedical ontologies and mappings: Overview of recent approaches. *Comput. Struct. Biotechnol. J.*, **14**, pp. 333–340, 2016.
[10] Kobilarov, G. et al., Media Meets Semantic Web – How the BBC Uses DBpedia and Linked Data to Make Connections, 2009.
[11] POSC Caesar Association, A Bit of History, 2011, Online. https://www.posccaesar.org/wiki/ISO15926Primer_ShortHistory. Accessed on: 22 Aug. 2017.
[12] Wen, T., Lyu, X., Kirkwood, D., Chen, L., Constantinou, C. & Roberts, C., Co-simulation testing of data communication system supporting CBTC. *IEEE Conf. Intell. Transp. Syst. Proceedings, ITSC*, 2015–Octob, pp. 2665–2670, 2015.