Cross-industry Remote Condition Monitoring Programme: Phase 2

Review of Relevant RCM Developments and the use of Asset Condition Information

RESEARCH AND
DEVELOPMENT

Copyright

© RAIL SAFETY AND STANDARDS BOARD LTD. 2015 ALL RIGHTS RESERVED

This publication may be reproduced free of charge for research, private study or for internal circulation within an organisation. This is subject to it being reproduced and referenced accurately and not being used in a misleading context. The material must be acknowledged as the copyright of Rail Safety and Standards Board and the title of the publication specified accordingly. For any other use of the material please apply to RSSB's Head of Research and Development for permission. Any additional queries can be directed to enquirydesk@rssb.co.uk. This publication can be accessed by authorised audiences, via the SPARK website: www.sparkrail.org.

Written by: CH2M HILL

Published: April 2015

Review of relevant RCM developments and the use of asset condition information

Executive summary

Background

RSSB research project T1010 Cross-industry RCM programme, Phase 2, aims to deliver a number of enablers of greater integration of remote condition monitoring (RCM) data. It is structured in 4 parts:

T1010-01	Data architecture
T1010-02	Commercial architecture
T1010-03	Standards
T1010-04	Business case tool update.

In June 2013 Halcrow (now CH2M HILL) was awarded the contract to deliver T1010-01.

The T1010-01 project will generate 2 major deliverables: a review of RCM developments, which is a thorough survey of influences on the data architecture; and an architecture requirements document, which will define the RCM data architecture.

This document

This document is the Review of RCM Developments. It summarises the work done in:

- Reviewing relevant documents in the following areas:
  - The formal background to the architecture: binding standards and the broader rail industry agenda
  - Influences from developments in the IT industry: technologies and approaches which might impact upon the architecture and its implementation
  - International standards in rail data, RCM, common data, integration and geography
  - Earlier research in this area by RSSB and others
  - Best practice in the development of data architectures in other domains such as aerospace, chemical/process, geographical
- Consulting with relevant stakeholders about RCM systems, data, software, IP, implementation, hardware, impediments. Consultation has been done both with individual stakeholders and in two industry-wide workshops

- Surveying existing RCM systems to identify the types of data they use
- Analysing and consolidating the findings into a set of requirements which feed into the next stages of the work: to identify the architecture principles and define the architecture requirements

Project methodology and process flow

This figure shows diagrammatically the process flow for the project. This document covers the steps shown under the banner 'Review of relevant RCM developments'.

train monitors train	train monitors infrastructure
infrastructure monitors train	infrastructure monitors infrastructure

The results of the research have been captured as a set of initial findings. These were then consolidated into a set of findings which were placed in the following categories:

- Business goals. These are high-level aspirations for the rail industry and the whole cross-industry RCM initiative which the data architecture should support
- RCM activities. These represent the needs and future aspirations of the different RCM projects and initiatives identified in the documentation
- Recommendations. These are the guidance offered by earlier research, by consultation and by our consideration on the best form of the RCM data architecture
- Constraints. These are considerations which limit the possibilities for the architecture, such as requirements to interface with existing rail systems or conform to particular legal or technical restrictions
- Dependencies. These are external events or developments which will be needed to take place before the full benefit of the RCM data architecture

can be realised, such as the deployment of automatic vehicle identification or a train location service

- Use cases. These are a set of user-orientated tasks that the RCM data architecture will need to support or enable. They will be used to validate that the architecture requirements cover the necessary functionality

A diagramming and requirements tracking tool *Enterprise Architect* has been used to store the findings and requirements and to trace between them. The tracing will be used to ensure that all the relevant initial findings are traced right through the intermediate findings to the requirements; and that all the requirements are helping to address a previously-identified need.

Findings and themes

The work so far has identified 427 findings, categorised in Table 1.

Table 1 - Findings by Category

Category of Finding	Number of Findings
Business Goals	35
RCM Activities	93
Recommendations	181
Constraints	24
Dependencies	9
Use Cases	85
Total	427

Through the work, a number of clear themes has emerged. These are described briefly in the following sections.

Identification

The ability to quickly and accurately identify a vehicle or component is critical to realising the full benefits of an RCM data framework. In practice this "identification" means the matching of data gathered by one component (e.g. a trackside monitor) with the identity of the component to which the data corresponds (e.g. a passing vehicle); or accurate identification of location - being able to match the location of a fixed asset to a piece of data gathered by a moving vehicle.

Accurate identification of both vehicle or component and its location are thus stated as key dependencies in the report.

Improved analysis and information

A clear point arising from research indicates that simply making the data available will not deliver the sort of improvements to reliability and costs that are envisaged. It must be accompanied by improvements in the analytical capabilities of those using the system, realised through the improvement of algorithms upon which the analysis is based. Alongside this there is an additional dependency for a better information base on all aspects of asset management. For example, asset configuration data will be important in component identification, deterioration profiles are needed to support assessment of an asset's health or lifespan and so on.

Integration with existing systems

RCM-based data can enhance and improve the efficiency of existing business processes for train operations, maintenance and asset management. These processes are mediated by existing information systems operated by different stakeholders. It is clear that the RCM data architecture therefore needs to enable communication between these existing systems and those offering RCM data. There should be no wholesale replacement of asset management, work management or rail operations systems.

Reuse, standards and interoperability

The RCM data architecture will be comprise definitions of data items, data types, and their associated parameters. In many cases these definitions will already exist and may be standardised. For example, asset management concepts are already clearly defined in ISO standards and rail terminology in the RailML schemas. Although this will save work in the definition stage of the architecture the clear benefit is of course the ease and consistency with which RCM can integrate with any system which uses the same definitions - in other words, interoperability.

Modularity and extensibility

The overall scope of the RCM data architecture is broad, covering many different types of RCM data for many different types of asset, generated by different types of equipment and IT system and offered in different data formats, being consumed by different types of operations and asset management users and systems. The RCM data architecture will therefore comprise many parts and will need to be constructed and evolved over time as uses for it become clear. Modularity and extensibility are 2 architectural concepts which will help make this possible. Modularity means that the architecture should consist of parts which are as far as possible independent of each other. Any given RCM project should only need to know about the parts directly relevant to its needs and should not need to drag in a large number of unnecessary extras.

Extensibility means that the architecture can be augmented or enhanced over time or for particular purposes, without forcing users who do not need the extensions to upgrade unless they wish to.

Web ontologies

Many of the sources researched and several of the contributors to work so far on rail data integration have indicated the potential value of semantic metadata to achieving the higher-level goals of data and process integration in rail operations and asset management. This value has been demonstrated in prototype in, for example, InteGRail. Web ontologies and their attendant standards and technologies are seen as the most likely way in which semantic metadata will be implemented in future. However, the technology still has some way to go before being ready for full-scale rail operation.

The RCM data architecture will need to be defined in such a way that it can accommodate future web ontology developments. The impact will be in methods of structuring and querying data from existing systems in a compliant way, and in the use of common ontological methods for representing standard information types.

Next steps

The findings from this part of the work feed forward into the next stage which will comprise the tasks described in introduction to the section Project context - binding standards and industry agenda, all included in the process called 'Architecture modelling' in Figure 1. The results of the modelling work will be the outputs described in the Section 2.2 .

Modelling tasks

The modelling tasks will comprise:

- Creation of a conceptual data model. This will model all the core concepts for RCM Data: RCM-specific items such as sensors, loggers, configuration parameters; RCM activity-related items such as alerts and alarms, deterioration profiles, health assessments; asset-related items such as identifiers, history markers, structure and composition relationships, topology for both fixed and moving assets; basic notions of date, time, and location; rail-specific notions of track position, vehicle and train identification, train service; ownership, responsibility and intellectual property considerations; data quality characteristics.

- Creation of a set of datagrams and information schemas. These will indicate how RCM and related rail asset data can be expressed and packaged up for transfer. The datagrams will reflect the data transmission needs set out in the use cases. Standard ways of expressing data and standard file formats will be used where possible.
- Creation of a set of message protocols. These will indicate how the data can be requested and exchanged, again reflecting the usage requirements defined in the use cases. Standard data exchange techniques will be used, with care being taken to simplify interconnection of existing systems and the adoption of the architecture in new RCM projects.

Modelling outputs

The modelling outputs will be:

- A set of architecture principles which set out the basic shape of the architecture, the key dependencies it has on external developments and external IT systems, some sample usage scenarios which can be validated by RCM data users, system suppliers and IT specialists.
- A set of architecture requirements. These will be in 2 categories:
  - Requirements on the deliverers of the infrastructure which will enable the architecture to operate
  - Requirements on the users of the architecture for RCM data interchange, so that they are compatible with it
- A list of dependencies - external tasks and activities which need to occur for the architecture to work
- A set of recommendations for the next stages of the cross-industry RCM data initiative
- An outline implementation plan

Glossary and abbreviations

Term	Definition
Advisory Generation	The sixth stage in the 6-level processing model for condition monitoring defined in ISO13374. Work orders, recommendations etc are issued based on the prognostic assessment.
AIXM	Aeronautical Information Exchange Model - framework to enable management and distribution of Aeronautical Information Services data.
Algorithm	A method of calculating a result based on input data.
Architecture	From TOGAF®: a) A formal description of a system, or a detailed plan of the system at component level, to guide its implementation (source: ISO/IEC 42010:2007). b) The structure of components, their inter-relationships, and the principles and guidelines governing their design and evolution over time.
ATOC	Association of Train Operating Companies
AVI	Automatic vehicle identification
CCS	Control and command systems
Conceptual Model	A data model showing how the core concepts of an area of interest are defined and are related to each other, independently of any particular way of storing or exchanging the data.

Data Acquisition	The first stage in the 6-level processing model for condition monitoring defined in ISO13374. Data are captured from sensors.
Data Manipulation	The second stage in the 6-level processing model for condition monitoring defined in ISO13374. Data from a sensor undergoes basic summarising cleansing operations such as averaging or signal processing.
Data model	A way of expressing the data items associated with an area of interest and the relationships between them.
Deterioration profile	A graph describing the expected rate of deterioration of some characteristic of an asset from an assumed level of usage or passage of time.
DRACAS	Defect resolution and corrective actions system
EIA	Enterprise integration architectures - an approach to integrating IT systems to support cross organisational business processes.
Encapsulation	The property of a system whereby only the elements of any object or data item that need to be made accessible to the outside world are visible; all other elements are kept private. Systems exhibiting this characteristic are easier to extend and develop.
Enterprise Architect	Data modelling tool used in T1010-01 to model the various artefacts of the project (such as sources, recommendations, dependencies, use cases) and traceability between them.
ESB	Enterprise service bus - a component of enterprise integration architectures whereby all participating IT systems interact via a single 'bus' with standard data formats, rather than directly with each other.

Extensibility	The property of a system which enables its capability to be enhanced easily after initial construction by adding new functions without disrupting existing ones or their users.
HABD	Hot axle box detection system
Health Assessment	The fourth stage in the 6-level processing model for condition monitoring defined in ISO13374. The health of the asset or component asset is assessed using the state from the previous stage and any supplementary information.
INSPIRE	Infrastructure for Spatial Information in the European Community - European directive aimed at creation of data infrastructure for spatial data.
InteGRail	European research project into information integration and management across the rail industry
Intelligent Infrastructure	Network Rail strategy to deliver infrastructure management improvements through intelligence design and maintenance through prediction and prevention rather than fixing.
LINX TM	Network Rail initiative to enable open interchange of information about the control of moving trains using EIA components to handle interfaces between traffic management and other industry systems. See EIA.
Logger	A device which records, stores and passes on the data from one or more sensors.
Metadata	Data about data. Information about the format and type of data; possibly also including information about how it can be used.
MIMOSA	Operations and Maintenance Information Open System Alliance - association for development of open information standards for operations and maintenance in manufacturing, fleet, and

	facility. Enables collaborative asset lifecycle management.
Modularity	The property of a system in which the functions and data are organised in groupings (modules) by purpose, with a small number of clear relationships between the modules. This property makes the system more extensible and maintainable because it limits the spread of impact of any single change.
NETEX	Network Exchange - XML schema based standardisation initiative for data formats and interchange methods for railway data, used widely in Europe.
NR	Network Rail
OLE	Overhead line equipment
Ontology	A schema where objects are defined in relation to other objects to allow understanding of a new concept by a machine.
Ontology Engineering	A set of processes associated with constructing and maintaining an ontology.
ORBIS	Offering Rail Better Information Services - a Network Rail programme to improve its approach to acquisition, storage and usage of asset information
OWL	Web Ontology Language - a language used to create ontologies. OWL is recommended by the internet's governing body World Wide Web Consortium (W3C).
Prognostic Assessment	The fifth stage in the 6-level processing model for condition monitoring defined in ISO13374. An estimate of the future health of the asset is made using the health assessment and any supplementary information about environment, expected usage etc.

R2	An IT initiative intended to replace the Rolling Stock Library and the Rail Vehicle Record System (RAVERS)
railML	Railway Markup Language - data standard for the interchange of railway industry, in the form of XML schemas. See XML.
RCM	Remote condition monitoring -the activity of monitoring the condition of assets remotely ie via a sensor at or on the asset which sends data to another location.
RFA	Railway Functional Architecture
RFID	Radio frequency identification - a system whereby tags wirelessly transmit identification data to a sensor.
RIS	Rail Industry Standard - a definition of technical or functional requirements. In this document it is the Traction and Rolling Stock RIS which is relevant.
Schema	A formal definition of a set of data structures. A specific type of schema that will be used is an XML Schema (http://www.w3.org/XML/Schema.html) which enables the structure and valid content of an XML file to be defined formally in a way that can be checked automatically in software.
Semantic	Pertaining to the meaning of data - i.e. what it represents and how that affects its relationship with other data.
Semantic Metadata	Metadata which describes a way of defining the meaning of data. An ontology is a type of semantic metadata.
Sensor	A device which converts a physical property of the environment or of an asset into an electrical signal.

Server	A computer which hosts a software service which can be used by other connected computers.
State Detection	The third stage in the six level processing model for condition monitoring defined in ISO13374. Processed data from sensors is assessed in relation to a threshold or band of acceptability and alerts or alarms raised if exceeded.
TLS - OB and MT	Train Location Strategy - RSSB project to enable provision of train location information to application providers to an industry standards through collation, consolidation and distribution of information from trains and signalling. TLS OB refers to on-board systems which send location information to the central system. TLS MT (multi train). MT will enhance the accuracy of this information based on data from many trains and send it back to the train OB system.
TOGAF	The Open Group Architecture Framework (http://www.opengroup.org/togaf/) A standard methodology for defining software architectures.
TOPS	Total Operations System
TRUST	Train Running System on TOPS
TSI	Technical Specifications for Interoperability - European Standards for the interoperability of data including TAP (Telematics Applications for Passenger Services) and TAF (Telematics Applications for Freight)
UJG	Uninterrupted Journey Group
UOMS	Unattended Overhead Line Management System
XiRCM	Cross-industry remote condition monitoring (in the rail industry) - remote condition monitoring occurring in:

	a) any of the quadrants where train based sensors monitor train, or infrastructure, and infrastructure based sensors monitor train, or infrastructure b) across different parties and sections of the industry - different groups may be involved in the operation of an asset, provision of a sensor, monitoring of that sensor etc.
XiRCM Programme Phase 2	Cross-industry remote condition monitoring programme: Phase 2 - a research project to facilitate the introduction of more cross industry remote condition monitoring, comprising packages to deliver architecture (01), commercial recommendations (02), review of standards (03) and decision support tool extension (04).
XiRCMSG	Cross-Industry Remote Condition Monitoring Strategy Group - a subgroup of the Vehicle/ Vehicle System Interface Committee with the responsibility to deliver the XiRCM Programme.
XML	Extensible Markup Language - a widely used language used to define data structures (see Schema) whereby elements and attributes and their structure can be declared and assigned values.

Contents

Review of RCM developments

1 Introduction

1.1 Background

In 2013 the Cross-Industry Remote Condition Monitoring Strategy Group, a subgroup of the Vehicle/Vehicle System Interface Committee, launched the Cross-Industry Remote Condition Monitoring Programme Phase 2. The purpose of the programme is to support the industry's high-level objectives of improving reliability, capacity and value for money, detailed in reports such as the Rail Technical Strategy 2012: The Future Railway and Rail Value for Money 2011.

The RSSB-managed research programme T1010 is intended to set up the enabling framework for this initiative. T1010 is in four parts:

- T1010-01 – the current project – is to define a data architecture to support cross-industry RCM data sharing
- T1010-02 is to define a commercial and legal framework to be used by parties wishing to share RCM data
- T1010-03 will update industry standards and guidance to promote and simplify the use of the data sharing architecture
- T1010-04 will update the business case assessment tool for RCM data sharing initiatives.

In 2013 Halcrow (now known as CH2MHill) were awarded the contract T1010-01 to develop the Cross Industry Remote Condition Monitoring data architecture for the UK rail industry. The data architecture is intended to remove some of the technical barriers to cross-industry data sharing caused by incompatibilities of approach between stand-alone RCM projects. It will do this by defining standard methods for structuring, representing and interchanging RCM data.

Remote condition monitoring is an aspect of asset management where the condition of an asset is monitored using data output from a sensor, so it can if necessary be monitored remotely from the asset itself (such as, if a person or system in a base station receives data directly from an asset on the railway).

Railway remote condition monitoring activities can be partitioned into four quadrants, defined by which types of asset – trains or infrastructure- are being monitored; and where – on trains or on infrastructure – the monitoring is being done. In this view, the 2 cross-interface quadrants - trains monitoring infrastructure and infrastructure monitoring trains - are of particular focus in this research as they involve different industry parties.

Figure 1 - The remote condition monitoring quadrants

train monitors train	train monitors infrastructure
infrastructure monitors train	infrastructure monitors infrastructure

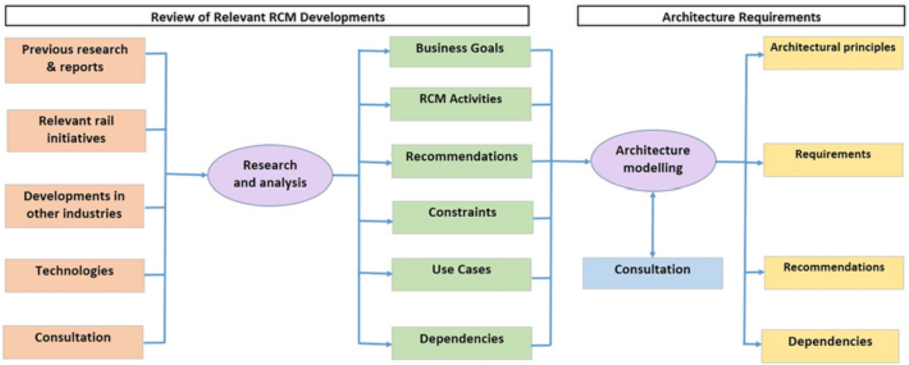
However, the 'cross-industry' element of RCM data sharing is not limited to this view. Sharing of RCM data on the same side of the rail boundary can deliver benefits by breaking down information silos. One of the visions for the rail industry discussed in the Rail Technical Strategy 2012 is the concept of a whole system approach, with aligned asset management practices. The greater availability of shared data through the industry will enable a greater level of condition monitoring to take place with resulting improvements to reliability. This is in contrast to the current situation where many RCM systems in use and being developed operate as stand-alone systems where the data types, communication protocols and architecture do not necessarily conform to any agreed standards and can be tied to the equipment vendor.

1.2 Project methodology

The project is required to deliver a Review of Relevant RCM Developments (this document); and a set of Architecture Requirements informed by the findings setting out the requirements of a data sharing architecture (to follow).

Figure 2 illustrates the process flow for the project. The output of this phase expresses the findings of the Review in a form which intends to inform the next stage of the project – to firm up the principles of the architecture and define its requirements. The following sections discuss the process steps briefly.

Figure 2 - Process diagram for T1010-01



1.2.1 Areas of research

The first part of the project was to carry out a review of recent developments in the field, to feed into the requirements for the architecture, The research covered a number of documents which have been grouped under the following headings described in the sections below:

- Context – the legal and regulatory framework and the broader rail industry agenda informing the project requirements (Section 2).
- Relevant information technology developments which may be put to use in the chosen solution (Section 3).
- Standards and sets of definitions likely to be relevant (Section 4). Amongst the requirements which emerged from the research, a key theme was interoperability and extensibility. In practice this would mean that an architecture initially set up to handle a particular set of systems, data and processes can easily adapt to handle others. This is achieved by establishing a single understanding of the world and the way elements in it are defined and this is supported by the application of standards. Relevant sets of standards in this case include ISO standards in areas such as asset management and condition monitoring; and existing rail data interchange standards.
- Precursors to T1010 in the form of previous RSSB research projects and papers on Remote Condition Monitoring (Section 5)
- Similar data architecture initiatives in other industries (Section 6), because they suggested useful methodologies or covered similar technical areas.

The team have also consulted with industry groups including potential users, maintenance managers, fleet engineers, infrastructure managers, asset engineers and IT managers (Section 7). This consultation, which is ongoing, will provide requirements on the asset management information programme ORBIS, infrastructure management programme Intelligent Infrastructure, MIMOSA open systems architecture and the AVI vehicle identification programme as well as on the RCM activities chosen as case studies: hot axle box detection and acoustic axle bearing monitoring.

A later stage of consultation will also validate the chosen architecture in terms of fit to business processes and ease of implementation.

1.2.2 Initial findings stage: collating findings and influencing factors

The first stage of the review was to collate all potentially relevant findings from all of the research material and capture them in our requirements tracking tool, Enterprise Architect. These findings were of many different types, for example, information on:

- current industry environment recent background and industry goals
- current and future RCM activities
- technologies and activities upon which RCM development is dependent
- recommendations for suitable technologies, languages, toolsets and methodologies

These initial findings were recorded as 'Influencing Factors'. There were a large number of influencing factors and a significant amount of duplication amongst them, indicating common themes in the research.

1.2.3 Traceability, validation and verification

The influencing factors are all linked to their sources, including the page number. In fact, all of the artefacts which will be discussed in this section are linked back to their source material, and to each other, to provide full traceability from every principle or component of the architecture back to its justification. This is enabled by Enterprise Architect: for more detail on the method, see Appendix D,

Traceability will enable the processes of validation and verification, which will be carried out through the various project stages to ensure that every component of the architecture is supported by a requirement gathered in the research or consultation stage, and conversely, that every requirement and use case is met by the architecture and every dependency or constraint is acknowledged in the project documentation.

1.2.4 Analysis of findings stage

To make sense of the large volume of influencing factors, analysis was carried out to categorise and de-duplicate the findings. The resulting consolidated list was categorised under the following headings:

- Business goals – high-level goals of the rail industry or of the cross-industry RCM initiative
- RCM activities – existing or planned RCM projects which may fall within the scope of the cross-industry data sharing architecture
- Recommendations – recommendations made by earlier RCM reports; guidance picked up from other similar work; input from stakeholder consultation
- Dependencies – factors in the outside world that impact upon the architecture. These may be enabling works in other projects, constraints already established by the industry structure, or the regulatory and legal framework within which the architecture must work

These consolidated findings, which can be used as the basis for guiding and validating the architectural design, are all linked to the influencing factors and hence to their source in the research, providing traceability from the eventual design components to the initial justification. The project requirements are contained within the consolidated findings. More detail on this is provided in the analysis of research in Section 9.

The main body of this document details the research that has been undertaken. Each of the elements of the research is covered in a separate section, with findings applicable to T1010 discussed. Analysed findings are shown in bold and referenced with a footnote to an item held in our Enterprise Architect software suite which is being used to handle the project artefacts. See Appendix D.

1.2.5 Consultation stage

The consultation stage is shown in Figure 2 in 2 places – as a part of the initial research stage and as a validation activity occurring prior to the creation of use cases and architectural principles.

In practice consultation has occurred throughout the project and is an iterative process – for example, influencing factors and requirements are fed back into the process from the validation of later stages of work which may inform or guide them; and some new influencing factors or requirements may be generated as part of the consultation.

A variety of relevant individuals and organisations have been consulted:

- those who use RCM systems
- those who develop or supply RCM system components – software, analysis, hardware
- those involved in other rail systems which may align with RCM systems
- experts in certain technologies or subject matter

The consultations aimed to:

- provide further information on the research stages
- provide opinion on importance of certain activities and technologies
- provide information on details of these (data requirements for example)

The findings of the consultation exercise are in Section 9 of this document.

1.2.6 Use case development stage

Further to this, a number of use cases were developed, based mainly on the outputs of the review of relevant RCM developments and the industry consultation. The use cases detail an actor, an action and a reason: 'An XX wants to do YY so that ZZ'. Use cases will be used to directly determine elements and attributes of the architecture, and all use cases will be traced to one or more of the consolidated findings, either directly or indirectly, thus playing an important role in establishing the traceability of the project. Further detail on the use cases is available in Section 9.6.

2 Project context - binding standards and industry agenda

2.1 Introduction

The documents reviewed in this section set the context for the current T1010-01 activity in terms of the binding standards, broader industry agenda and higher level architectures that the data architecture fits into. They are:

- The Rail Technical Strategy 2012 (Section 2.2)
- Network Rail's Technical Strategy 2013 (Section 2.3)
- The Railway Functional Architecture defined in RSSB research project T912 (Section 2.4)
- Industry constraints: relevant European TSIs and UK Railway Group Standards (Section 2.5).

2.2 Rail Technical Strategy 2012

2.2.1 Overview

This document details the vision of UK rail in 30 years' time and the strategies to achieve it. Reliability is a prime component of the vision, impacting as it does on customer experience and cost. Best practice asset management is thus a critical enabler for this vision, as is the realisation of the principles of a whole-system approach, itself an enabler of required improvements to asset management.

The document details the vision, objectives, strategy and enablers for various aspects of the industry, including, relevant to RCM, Infrastructure and Information. It provides essential background to the context that the proposed RCM data architecture will be operating in, in particular in relation to the management of infrastructure and associated information.

2.2.2 Relevant findings

This strategy provides a clear justification of the case for an RCM data architecture:

1. RCM is listed as an enabler of infrastructure-based improvements in cost and performance, both train monitoring infrastructure and vice versa.
2. A cross-industry commercial and technical framework is declared as an enabler in allowing asset condition data to be shared and exploited.
3. It lists as an objective a whole-system approach stating that it will deliver improved resilience and reliability of infrastructure and rolling stock through optimising interfaces with rolling stock and infrastructure.

Whilst not providing detailed information about any aspect, the document points to a number of activities and practices which are deemed key to achieving the strategy, which will thus inform the scope of the architecture.

1. Asset management practices – streamlined and short cycle maintenance operations between service trains, and automated routine maintenance will mean that track is available for more time. This is dependent on access to, and analysis of, condition based data.[1]
2. Condition-based intervention from train-borne (including service train) inspection with remote defect verification – effective maintenance is dependent on the accuracy and timeliness of data from this.

The document lists a number of pre-requisites to achieving this vision. These are considered to be dependencies for an RCM data architecture:

1. Greater and more documented asset knowledge, through a single, industry-wide asset register and improved understanding of life cycle cost and degradation. The latter will become dependent upon the data coming out of the architecture as it is used to fine-tune asset information and generate prognostic and advisory information.[2] So whilst beneficial use of RCM is dependent on a library of asset knowledge, the knowledge will itself become dependent on RCM to stay up to date. [3]

---

1 Business Goal 018 and RCM Activity 140a

2 Author's note

3 Dependency D01

2 Apply in-depth theoretical modelling to predict minimum life and asset failures [4]

3 Improve location accuracy[5]

4 Improve identification of problem sources.[6]

Many points in the document have been taken forward as requirements of the RCM data architecture:

1 Adoption of international RCM, data processing and asset management standards[7]

2 Consolidate existing asset management systems such as ORBIS. [8]

3 Rail businesses can integrate a variety of internally and externally generated information.[9]

4 Information systems are built for expansion and easy replacement of outmoded parts and based on 'commercial off the shelf' systems to minimise costs. [10]

5 Components of a dynamic architecture include ontologies, security classifications, storage requirements, agreed data types. [11]

6 Follow guidelines from InteGRail and TSI standards. [12]

7 Resilience to cyber-attacks.[13]

2.3 Network Rail Technical Strategy 2013

2.3.1 Overview

This document sets out Network Rail's response to the Rail Technical Strategy discussed in Section 2.2. It develops the themes mentioned in that document and applies them to Network Rail's own responsibilities and circumstances.

---

4 Dependency D02

5 Dependency D03

6 RCM Activity 022

7 Recommendation 036a

8 Recommendation 013a and 214

9 RCM Activity 035

10 Recommendation 038 and 039

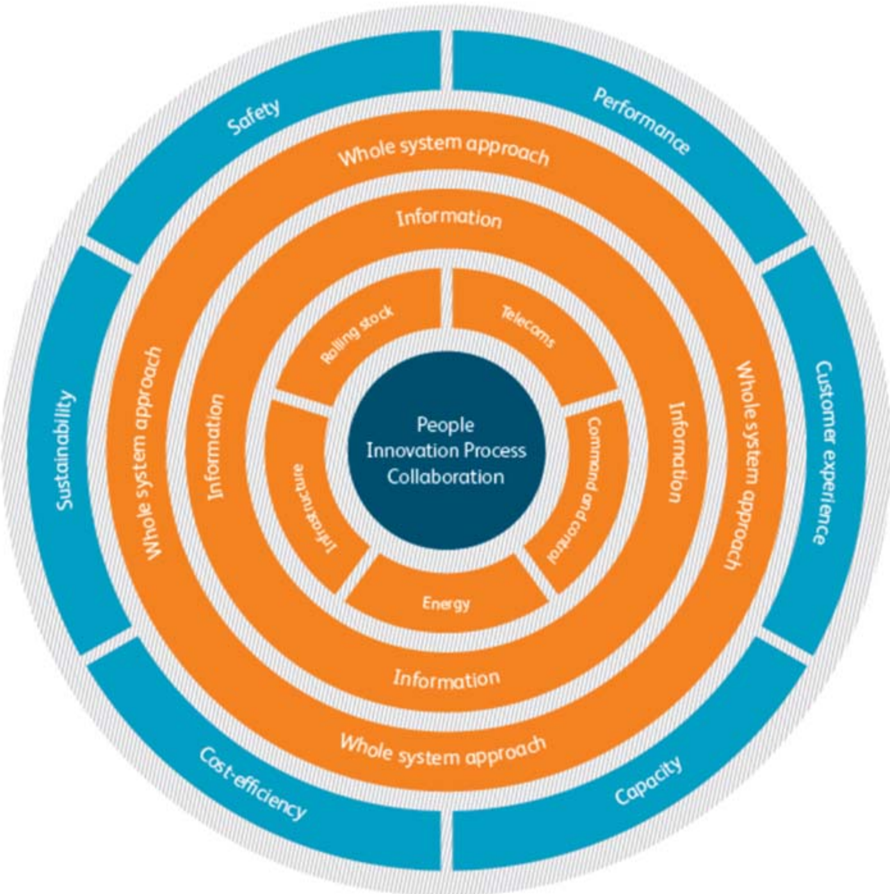
11 Recommendation 044 - 047

12 Recommendation 76, 76a, 76b, 37

13 Recommendation 051

The document considers the Technical Strategy under a number of outcomes, themes and enablers, summarised in Figure 2.

Figure 3 - Network Rail Technical Strategy – outcomes, themes, enablers



The outcomes, which the strategy seeks to realise benefits in, are shown in light blue. These represent the axes on which the performance of Network Rail is measured.

The enablers, which support the research process, are shown in dark blue. These represent the resources and attitudes which will help bring about the changes necessary to deliver the results of innovation.

The themes of research and development, which are the heart of the strategy, are shown in yellow. These match those used in the Rail Technical Strategy. Innovation in each of these areas is planned to be driven by a portfolio of research tasks, not shown in this Figure. There is a theme for each of the major asset types and activities managed by Network Rail: Infrastructure, Rolling Stock, Telecommunications, Command and Control and Energy. Two themes cut across all of these: the Whole System Approach tackling the need to see the entire railway as an integrated system rather than a set of independent disciplines; and Information which provides the means to do this.

The research and development portfolio comprises initiatives in these areas:

- Monitoring, Assessment and Optimised Maintenance. This covers the whole area of Remote Condition Monitoring and the processes built upon the data so gathered.
- Modelling and Decision Support. This involves improving the theoretical understanding of how the network behaves and the presentation of information to human operators to help them take better decisions.
- Traffic Management. This concerns optimising the movement of trains to maximise the use of capacity, performance, and passenger satisfaction; and to minimise the use of energy.
- Seamless End-to-End Journeys. Improved retail systems, information for passengers in disruption and across transport modes, station designs; optimised freight and passenger services.
- Increased Data and Information Flow. Improving security, creating common information architectures, opening information processes to external suppliers, enhanced communications bandwidth.
- Better Energy Use. Electricification, smart power distribution, reducing waste.
- Enhanced Design. Improving the design of assets and systems to improve standardisation, favour commercial off-the-shelf products, improve performance and reduce cost
- Transformation. Improving and re-designing processes and procedures to enhance collaboration and add focus to R&D efforts.

2.3.2 Relevant findings

Several of the themes and research portfolios overlap directly with the scope of the Cross-industry RCM data architecture; and the architecture will generate indirect benefits for several others.

Under the whole system approach theme, the data architecture directly impacts the short-term goals to develop modelling capability to understand trade-offs in operation and to improve whole-life costing[14]; to turn data into information for engineers[15]; to adopt worldwide best practice[16]; and to improve safety of trackside workers by increasing the use of remote monitoring of fixed assets[17]. In the longer run the architecture will support the need to double passenger capacity in 30 years by improving network availability and performance[18] and improving the precision and level of integration of railway operations and maintenance processes.[19]

Under the Information theme, the data architecture directly supports the goals to automate the collection of asset information[20], the interchange of messages across the industry and international borders[21], associated processes such as analytics, reporting and augmentation[22]. It will also indirectly support the improvement of data management and governance[23] and the application of integrity standards by demanding high-quality reference data to support cross-industry processes.[24]

Under command and control, the data architecture supports the goal of improved recovery from perturbation. [25] It could also drive adoption of the Compass train positioning programme. [26]

Under the Energy theme, the Data Architecture will support moves to improve the reliability of the overhead line equipment by automating the assessment of its condition and by reducing time to fix.[27]

14 Business Goal 000, RCM Activity 011

15 RCM Activity 029

16 Business Goal 012

17 Business Goal 248

18 Business Goal 003 and 016

19 Business Goal 003 and 013

20 Business Goal 006 and 007

21 Business Goal 006a, 013, Recommendation 013a, 040a, 066-070, RCM Activity 035

22 RCM Activity 011a, 029

23 Dependency D04

24 Dependency D05

25 Business Goal 242

26 Business Goal 243

27 Business Goal 244

Under Infrastructure the Data Architecture contributes to remote condition monitoring of earthworks[28], to the real-time monitoring of assets, building on the Intelligent Infrastructure programme.[29]

Under Rolling Stock, the Data Architecture supports the remote collection of condition data, particularly at the system interfaces: wheel / rail, pantograph / OLE, shoe / 3rd rail, train to ground telecommunications[30]. It will also indirectly impact the standardisation of on-board equipment and the move to reliability-centred maintenance of rolling stock.[31]

2.4 T912 Railway Functional Architecture

2.4.1 Overview

The Railway Functional Architecture, published in January 2011 as the output from RSSB research project T912, describes, amongst other things, the logical functions, entities and systems which describe the operations of a rail network, via a series of views or 'perspectives' including the concept perspective (the 'what') and the solutions perspective (the 'how'). There is also a 'why' aspect included in the enterprise perspective views: this is not of immediate interest in defining the RCM data architecture.

The architecture described is based on The Railway Architectural frameworK (TRAK), a method of describing the high-level functions of a railway organisation originally developed for London Underground.[32]

2.4.2 Relevant findings

Amongst the views contained in the Railway Functional Architecture there are a number which are of relevance to RCM activities and the dependencies of an RCM data architecture, which are listed below. **The RCM data architecture should be aligned with the relationships in the views in Table 1.**

28 RCM Activity 174

29 RCM Activity 020, Recommendation 078

30 RCM Activity 195 and 197

31 Business Goal 012, 018, 245

32 http://trak.sourceforge.net/

Table 1 - Railway Functional Architecture views relevant to RCM

Ref	Title	Elements incorporated in view:
CV03	Condition monitoring	asset, condition monitoring agents and maintainers
CV01	Maintenance	asset, condition monitoring agents and maintainers
SV04	Rolling stock condition monitoring	functions such as HABD and systems and standards
SV04	Signalling and control infrastructure maintenance	condition monitoring and analysis, scheduled maintenance and corrective maintenance
SV04	Track condition monitoring	various automated track inspection functions
SV04	Train service location	various functions involved in determining velocity and location of trains
SV01	Condition monitoring	the various physical elements and systems involved in train-borne and trackside monitoring
SV01	Hot axle monitoring	track-mounted and train-borne
SV01	Track Inspection	automated track inspection
SV01	Automatic train location service	train and track based systems

The RFA also suggests some best-practice architectural guidelines that data integration initiatives should follow, that the RCM data architecture should respect: modularity (defined as the degree to which a system's components may be separated and recombined, the tightness of coupling between components, and the degree to which component relationships may or may not be allowed); and the importance of eliminating 'sub-optimal modular interfaces': cases where the same item or interaction element exists in multiple relationships. [33]

2.5 Constraints and binding standards

Any RCM data sharing initiative needs to respect the mandated processes and their guiding documents already in place. These are the European Technical Specifications for Interoperability (Section 2.5.1) and UK Railway Group Standards (Section 2.5.2).

2.5.1 European standards – Technical Specifications for Interoperability

2.5.1.1 TAP & TAF TSIs

The Technical Specifications for Interoperability for Telematics Applications for Passengers Services (TAP) and Freight (TAF) were adopted by the European Commission in 2011 and are now in the process of implementation. These are mandatory standards for railway undertakings (train operators) and infrastructure managers who participate in cross-border train operations (though not just cross-border trains are impacted – the ultimate scope of the TSIs is all trains).

The 2 mentioned TSIs (there are several others) define coding standards for railway vehicles and railway locations, plus a set of message types – a 'Common Interface' - intended to support standard business processes associated with cross-border train operation. Each EU country's nominated data manager has to maintain standard reference lists of rail vehicles, locations and companies, with a history for each item in each list.

Recently, the EU has funded a reference implementation of a messaging architecture capable of managing the TSI-mandated Common Interface. It is recommended that this implementation be used as a basis for each country's own effort to conform to the TSI requirements. The reference implementation uses an open-source Enterprise Service Bus tool and is based on Enterprise Integration Architecture principles. (These principles, as they would apply to the RCM data architecture, are described in Section 3.1).

Whilst the TSIs do not directly contain any reference to RCM data, they are relevant to the extent that RCM data are attached to rolling or fixed assets that may need to be referred to by TSI messages. For example, TSI messages refer to timetabled and proposed routes which involve sections of infrastructure; they involve interoperability issues such as OLE and CCS standards; and they refer to train consists. It is possible also to envisage that

---

33 Requirements 207

future data interchange on rolling stock health or on network availability may need to contain some derived information from an RCM analysis.

We therefore take forward as requirements that **Representations of trains, route sections and asset health statuses should be compatible with those defined in the TAF/TAP TSIs.**[34]

2.5.2 Railway Group Standards and guidance

Railway Group Standards define technical or safety standards that are to be followed by participating bodies in the UK rail industry. RSSB is the statutory body responsible for managing the Standards and their lifecycle.

As well as standards, RSSB manages other types of less formal document which have an advisory rather than mandatory force. Guidance Notes and Rail Industry Standards fall into this category.

The following Railway Group Standards and Rail Industry Standards have been reviewed as part of this work. Those which are relevant to the cross-industry remote condition monitoring data architecture have been indicated, for the reasons stated.

Table 2 - Relevant Railway Group Standards

Document	Relevance to RCM Data Architecture
GC/RT5021 Issue 5 12/2011: Track System Requirements.	Important elements of data required to describe track.
GE/RT8054 Issue 2 09/2011: Management of Shared Information Systems.	The primary requirement for safety in the management of IT systems; the need to set up a system management group, jointly between railway undertakings and infrastructure manager, for shared systems.
GE/GN8565 Issue 1 06/2004: Guidance on the Retention of Design Information for the Validation of Technical Change and Configuration Management	The need to maintain a history of asset configuration to enable assets to be tracked over time.
GE/RT8047 Issue 6 12/2013: Reporting of Safety Related Information	Key safety-related systems: SMIS, FMS, Control Log; the duty to report safety infringements and near misses; information to be recorded.

34 Recommendation 037aa

Table 2 - Relevant Railway Group Standards

Document	Relevance to RCM Data Architecture
GE/RT8250 Issue 2 06/2007: Reporting High Risk Defects	The role of the National Incident Register Online; the duty to report vehicle faults immediately if serious fault found; the need to share information industry-wide.
GE/RT8014 Issue 2 06/2011: Axlebox Condition Monitoring – Hot Axlebox Detection; GE/GN8614 Issue 1 06/2011: Guidance on GE/RT 8014; BS EN 15437-1:2009 Railway applications - Axlebox condition monitoring - Interface and design requirements.	Types of Hot Axlebox Detector equipment; the need to communicate expected temperature ranges for known stock classes and bearings; functions, thresholds and alarms for each type of HABD; temperature measurement standards; correction for ambient temperature.
GM/RT2466 Issue 3 02/2010: Railway Wheelsets; GM/RC2496 Issue 2 02/2010 Recommendations for Railway Wheelset Maintenance	Data requirements for identifying and tracking wheelsets and their components; thresholds / exceedences for different types of problem; times to remove from service if damaged.
GM/RT2472 Issue 1 06/2002: Data Recorders on Trains – Design Requirements; GO/RT3272 Issue 3 06/2002: Data Recorders on Trains – Operating Requirements.	On-train Data Recorders: data items stored; data extraction requirements; data storage requirements.
GE/RT8270 Issue 2 10/2007: Assessment of Compatibility of Rolling Stock and Infrastructure.	The requirement that information regarding rolling stock compatibility with infrastructure needs to be kept on the Rolling Stock Library; how that should be done.
GE/RT8106 Issue 2 12/2011: Management of Safety Related Control, Command and Signalling (CCS) System Failures	The data required to be transferred between safety duty holders for cross-industry faults.

Table 2 - Relevant Railway Group Standards

Document	Relevance to RCM Data Architecture
GM/RT2453 Issue 2: 09/2011: Registration, Identification and Data to be Displayed on Rail Vehicles	The data items and formats of rolling stock data to be shown on the vehicles themselves.
RIS-2706-RST Issue 1 03/2013: Rail Industry Standard for Recording of Rolling Stock Data	Data items mandated and advised to be made available for sharing about rail vehicles.

Requirements on the RCM data architecture have been created where the reviewed document has an impact[35].

35 Recommendation 037a

3 Information technology influences

This section describes some developments in Information Technology which have the potential to help in the design and implementation of the RCM Data Architecture.

3.1 Enterprise integration architectures

3.1.1 Introduction – the challenge

In recent times a key challenge for software architects has been to extend the automation of business processes. Most existing IT systems work at the level of a corporate or technical function or a department, and tend to be focussed on applications that facilitate a task or small number of tasks. It is now necessary to connect together these IT systems to enable them to support higher-level, cross-functional or cross-organisational business processes.

When automating business processes by connecting together existing separate IT systems several problems emerge:

- Connections across a network are unreliable. Wide area networks, virtual private networks or the Internet are subject to disconnection, mis-routing of data, bandwidth limits, hardware failure and other impediments
- Connections across a network are slow. Web-based interactions are several orders of magnitude slower than communications between programs running on the same computer
- The systems are different from each other. They may be based on different technologies, use different languages, be different ages, hold data items in different formats. Connecting them together needs translation of data between the different styles and formats
- The systems change at different rates and for different reasons. Each individual IT system has its own rationale for change, upgrade and development, driven by its local business needs and constraints. Such

changes may now demand cascading changes in other connected systems to maintain compatibility

These problems are made exponentially worse as the number of connected systems increases, such as where the goal is an open architecture which new participants and their systems may join up to at any time.

3.1.2 The response

In response to these challenges, an architectural approach and set of best practices has emerged in recent years. The approach, techniques and practices are collectively known as Enterprise Integration Architectures (EIA). They have these key characteristics:

- They are bus-based. Rather than each participating IT system interfacing directly with others it needs to talk to, it interfaces with a 'data bus' which links all the systems and which internally has a known standard data format. Data coming from a source IT system are translated from their specific format to the bus format using an 'adaptor'; and another adaptor is used to translate data from the bus format to the destination system's specific format. This means that the many possible pair-wise data format mappings and translations are replaced by a much smaller number of system-to-bus and bus-to-system translations
- They are service-based. Each participating IT system interacts with others through a series of basic scripted interactions involving requests for data or action, and responses which provide the data or indicate the success or otherwise of the action. This means that business-level processes can be built up from orchestrated sequences of these actions; and each system's responsibilities to the whole environment are clearly defined. Data responses can be provided in a number of formats, from structured text in the form of XML or JSON, to standard file formats such as CSV for text, JPG for photos or MP3 for sound and video
- They are message-based. The requests and responses associated with the services are interchanged using messages. This has many benefits, primary among them being that the requestor and responder can operate asynchronously (they do not have to be working at the same time for the interaction to succeed) and that communications can be guaranteed (messages can be re-sent periodically until received; and proof of transmission and receipt can be provided). There are challenges to the approach: the business process must be defined in such a way that it is can

work asynchronously and that it can cope with responses coming back in an unexpected sequence

- They are standards-based. Published open standards exist for many aspects of the technologies used; and a set of de-facto best practices and shared behavioural patterns is evolving. Example standards are SOAP (originally Simple Object Access Protocol), which uses the same basic technology as the World Wide Web to handle message requests and responses; and XML (Extensible Markup Language), which enables arbitrarily complex data structures to be encoded and decoded in a textual form which can be validated automatically and read by humans if necessary

Along with these Enterprise Integration Architecture concepts has come a number of key technologies which can be used and combined to implement them. The main ones are:

- Enterprise Service Bus (ESB) tools. These supply the data interchange backbone of an architecture and the tools to create, operate and manage data adapters to get data into and out of it
- Web Service design tools. These enable web-based request and response messages and their data formats to be designed and managed effectively
- Message Queues. These enable a number of types of message-based interaction to be defined to meet a number of common business needs. 'Store-and-Forward' and 'Publish-and-Subscribe' are typical patterns

The approach is supported by software suppliers with products that implement these technologies. Examples are:

- Enterprise Service Bus: Commercial tools such as TIBCO BusinessWorks, IBM Websphere ESB, Microsoft BizTalk; open source tools such as JBOSS ESB, Mule ESB
- Web Service Design Tools: Stylus Studio; Eclipse
- Message Queues: IBM Websphere MQ; Microsoft MSMQ; Open MQ

3.1.3 Use in the rail industry

Some recent developments in the UK rail industry have adopted EIA principles as a means to address their particular challenges of integration with existing IT systems and business processes (described further in Section 5.4):

- LINX TM uses a commercial ESB product and a message-based architecture
- R2 will use a message-based architecture based on a data bus for integration with mainframe systems
- ORBIS is conceived as a web services based application suite

3.1.4 Conclusions

The Enterprise Integration Architecture approach is relevant to the RCM data architecture because it addresses a number of its key concerns:

- Integration with existing legacy IT systems. RCM data will need to be shared between existing IT systems and RCM data capture systems. Reference data needed to support the RCM data will need to be sourced from existing rail industry systems using their own formats
- Integration with systems under development or consideration that use the EIA approach
- Openness and ability to connect new systems to existing ones. The use of standards and a bus-based approach will facilitate this [36]

Therefore, Enterprise Integration Architecture concepts should be strongly considered as the underlying principles for an RCM data architecture. Resulting recommendations for the architecture will include that it is bus-based, service-based, message-based and standards-based.

3.2 InteGRail

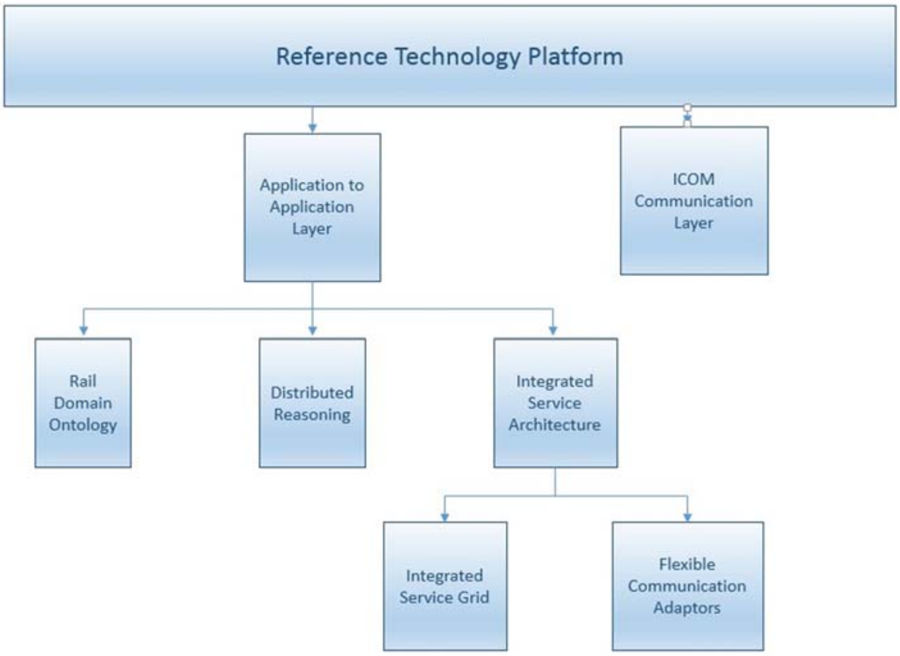
3.2.1 Overview

InteGRail was a European research project which investigated the possibilities for information integration and management platform that can be used across the rail industry. Its deliverables can be classed into two groups – a description of the technology platform and a set of prototype applications which would use it.

The components of the platform are shown in Figure 4.

36 Recommendation 318

Figure 4 - InteGRail technology platform



The platform is composed of an application to application layer and a communications layer. The application layer contains:

- a basic rail domain ontology – a set of rail ontologies forming a standard interchange format allowing a machine-interpretable conceptual model of physical and data objects
- distributed reasoning capabilities - the leveraging of the definitions of the rail domain ontology to create views from distributed information
- the platform's service-orientated architecture which is based on an 'integrated service grid' (by which is meant an enterprise service bus, as described in Section 3.1.2) and 'flexible communication adaptors' (the same as ESB data adaptors) through which the various applications and systems connect

In addition to the platform, the creation of the application prototypes in a set of development scenarios led to proofs of concept for the integration of the service grid with existing information systems and a recommendation that it should be reused across different scenarios.

3.2.2 Relevant aspects

3.2.2.1 Objectives

There is a strong correlation between the objectives of the Xi RCM initiative and those of InteGRail, being stated as enabling information and its context to be shared within the railway and optimising decision-making based on this information to improve performance[37]. In particular, improvements to monitoring systems and optimisation of maintenance are mentioned in InteGRail's Vision document. **For this reason the findings of the project are of particular interest in the development of the RCM data architecture and the application of the findings of InteGRail as standards should be strongly considered.**

3.2.2.2 Targets

The targets of the InteGRail architecture were integration, flexibility, decision support and possibly evolution. This aligns very closely with the other requirements which have been gathered for an RCM data architecture, in which extensibility for wider and future use is critical to its value, as is the enabling of decision support functions, through integrated data from distributed sources. **Again, this supports the use of InteGRail findings and approach as standard in this project.**

3.2.2.3 Service oriented architecture

Described as a key architectural notion in the integration of different existing systems in InteGRail, the service-oriented architecture has the following characteristics: [38]

- A web-browsable front end. This frees users from needing to install specific client tools and so broadens adoption
- A modular architecture. Functional elements of the system are largely independent of each other and loosely-coupled to each other. This supports future extension and enhancement
- Use of a service bus. A 'bus' architecture reduces the number of point-to-point interfaces between systems and therefore simplifies the task of connecting systems together. The use of web services provides a standard mechanism for systems of all types to interact with each other via the bus, independently of their own structure, architecture and processing platform

37 InteGRail Final Activity Report section 2.3.1 pp9-10

38 Recommendation 093

- Support for integration of legacy systems. Most existing business functions are supported by existing systems. The risk and cost of replacing them wholesale to support new cross-industry processes are prohibitive
- Clear functional layers. It has been well-proven in IT systems that a layered architecture, with clear interfaces between layers, promotes interoperability and standardisation. (An excellent example is the OSI 7-layer model for communications systems, where the lowest level deals with the physical methods of interconnection such as Ethernet; the middle layers deal with networking protocols such as TCP/IP and the top-level represents an application protocol such as the http web protocol)
- Orchestrated processes. Rather than build monolithic applications to support business processes, this approach uses a method of combining a set of services, which may be provided by different IT systems, in a particular order. The advantages are that arbitrarily complex processes can be supported by the same kit of parts; and development and improvement of the services themselves can proceed independently of the processes using them

For an RCM data architecture, this approach has much to recommend it, speaking as it does to the environment in which different existing IT systems need to support new methods of collaboration as yet undefined, in a way which enables extension and enhancement.

Ensure that the RCM data architecture is planned with consideration to all of the aspects of service-oriented architecture as they are discussed above. [39]

The discussion in Section 3.1 on enterprise integration architectures develops these themes further. The service oriented approach is a key element of an enterprise integration architecture.

3.2.2.4 Rail domain ontology

The concept of a single rail domain ontology is an important part of the proposed InteGRail framework. It is proposed on the basis of successes in other industries, particularly appropriate where there are numerous heterogeneous data sources, and it is anticipated that it can be used to exchange data, develop ‘plug in’ software, and capture implicit domain data. The rail domain ontology created by the project used Web Ontology Language (OWL) **The importance placed on the rail domain ontology should direct the design of the RCM data architecture in these ways:** [40]

39 Recommendation 093b

- in the adoption of a single set of domain definitions (no overloaded terms)
- in the intention to develop ontological metadata to allow machine interpretability
- in the decision to use OWL as the schema language which should be noted when the decision is made on what language to use for the schemas in the RCM data architecture

3.2.2.5 Processing levels

The main processes involved in InteGRail are defined as: recognise, define, measure, analyse, improve, and control. These map well on to the usages of the data architecture described in the Use Cases, Section 9.6.

3.2.2.6 Prototype applications

Prototype applications were developed as a series of 'development scenarios' DS1 to DS4, each of which covered a particular set of information integration challenges. The prototype applications developed as part of development scenario DS2 are of particular relevance to an RCM data architecture. DS2 illustrated a scenario where wheel, track and event analysis applications and rolling stock, axle and wheel data are linked via the InteGRail service grid**. This provides a pointer for areas to focus on implementing, with a view to reusing the definitions created by InteGRail to reduce development time and achieving consistency with any definitions in current use.** [41]

3.2.2.7 Methodology - requirements

Requirements for InteGRail were arrived at in two ways, being defined through both top-down and bottom-up approaches. Top-down needs were those related to achieving the key performance indicators which were identified as part of the project work (the *required* functionality). Bottom-up requirements reflect the current practices. **It may be useful to categorise requirements for an RCM data architecture in this way (existing application and potential future application) in order to provide additional rigour to the documentation of and elicitation of requirements – it is clear what is required to support current applications, and an exercise to envision a future set of applications is not forgotten.**[42]

40 Recommendation 076a

41 Recommendation 076c

42 Recommendation 222

3.3 Web ontologies

3.3.1 Introduction

The concept of 'ontology' has been around in computer science from the 1980s when it was first used to refer to a theory of a modelled world by those working on the development of artificial intelligence. It was later formally defined as a technical term in the 1990s thus: 'An ontology is a description (like a formal specification of a program) of the concepts and relationships that can formally exist for an agent or a community of agents.'[43]

In practice, however, it is a concept that is still in its infancy – its use is developing in fields related to browser search and advertising however it has enormous potential in almost every aspect of the digital world. The use of ontology developed with the ontological language OWL is recommended by the Internet's governing body World Wide Web Consortium (W3C).

3.3.2 What does ontology mean

An ontology is a specification, rather like the XML specifications seen in Section 4.2.2; it defines an element, or class of object, in relation to other items. However, whereas a standard definition (for example in XML) would contain just the information that 'wheel' is a component of 'vehicle', which is understandable to a human reader who is familiar with the concept of 'wheel', an ontological definition would allow a machine to understand more about what the tag 'wheel' actually means. For example, it is a moving part, it has a relationship with the track, it has a relationship with axle, it is found on all land vehicles and there are x many of them on a particular sort of vehicle.

This extra layer of meaning is known as semantic metadata, and creates the possibility of machine reasoning with information and creating inferred knowledge rather than just storing data.

Figure 5 shows an example of the definition of a rail vehicle from SUMO's set of domain ontologies (SUMO is one of a number of 'upper level ontologies' and has a set of associated 'mid-level ontologies'. If someone wanted to create their own ontologies in a new area they would use an existing upper or mid-level ontology to connect their own classes to those already existing which represent higher level, more generic things. For example, the main mid-

43 Tom Gruber 1995 'Toward Principles for the Design of Ontologies Used for Knowledge Sharing'. *International Journal of Human-Computer Studies* **43** (5-6): 907–928.

level ontology MILO contains definitions of person, human, organism, death and so on).

Figure 5 - Extract from SUMO rail vehicle ontology

```
;;      (1) b. Rail Vehicles

(subclass RailVehicle LandVehicle)
(documentation RailVehicle EnglishLanguage "A Vehicle designed to move on &%Railways.")

(=>
  (instance ?X RailVehicle)
  (hasPurpose ?X
    (exists (?EV ?RAIL)
      (and
        (instance ?RAIL Railway)
        (instance ?EV Transportation)
        (holdsDuring
          (WhenFn ?EV)
          (meetsSpatially ?X ?RAIL))))))

(subclass Train RailVehicle)
(subclass Train PoweredVehicle)
(subclass Train Collection)
(documentation Train EnglishLanguage "&%Train is the subclass of
&%TransportationDevice whose instances are linked sequences
of &%RollingStock.")

(=>
  (instance ?TRAIN Train)
  (exists (?X ?Y)
    (and
      (instance ?X RollingStock)
      (instance ?Y RollingStock)
      (not
        (equal ?X ?Y)))))

(subclass TrainStation TerminalBuilding)
(documentation TrainStation EnglishLanguage "TrainStation is the subclass of
&%Buildings that are located at a &%RailwayTerminal and used in support
of its functions, especially for the handling of passengers and freight.")

(subclass RailTransportationSystem TransitSystem)
(documentation RailTransportationSystem EnglishLanguage "RailTransportationSystem
is the subclass of &%TransitSystems whose routes are &%Railways.")
```

It can be seen that the rail vehicle is defined logically and in relation to other elements.

3.3.3 Applicability

In 2011 Dr Clive Roberts of the University of Birmingham presented a paper titled *The Specification of a System-wide Data Framework for the Railway Industry* in which a strong case was made for the suitability of ontologies as the basis of a common data model for the rail industry. It is also clear to see the potential benefits to an RCM data architecture where elements from one system are defined in a way that is understandable by machines in other systems, as the importance of interoperability and extensibility is evident from the research summarised in this document.

However, the creation of ontological metadata for the schema being reused or created for use in an RCM data architecture will be a time consuming process, as it is recommended that this is carried out collaboratively and iteratively. For this reason the architecture will not depend on the creation of these specifications, but will allow for a layer of semantic metadata to be added to the architecture in future.

4 Standards-based precursors

This section describes international standards which have been identified as having potential value in the RCM data architecture. They cover areas such as condition monitoring, IT systems integration and formats of various types.

4.1 ISO Standards

4.1.1 Asset management and condition monitoring

4.1.1.1 ISO13374 Condition monitoring and diagnostics of machines – Data processing, communication and presentation

Compliance with this standard is intended to allow condition monitoring data to be processed and shared by different heterogeneous systems and applications. The standard provides the recommended requirements whose implementation in a data architecture will enable this.

ISO 13374 specifies the six processing levels of condition monitoring which it is likely the RCM data architecture will be built on, and includes the MIMOSA specification on open information standards – see Section 6.1 for more on this.

It has been published in three parts, 1: General, 2: Data Processing and 3: Communication

Part 2, published in 2007 and reviewed in 2010, includes data processing requirements on the six process levels, archiving and display, and information requirements, such as semantic definitions, and data model and data library requirements. It also includes the specifications for MIMOSA's Open Systems Architecture in Condition Based Monitoring (see Section 6.1 and references to the relevant modelling languages UML and markup language XML).

Part 3, published in 2012, contains requirements on interfaces, message initiation and message content requirements.

4.1.1.2 ISO55000 series and PAS55

PAS55 is a physical asset management standard initially produced through the Institute of Asset Management and most recently updated in 2008. It is widely used in the UK and Australia. It is concerned with optimal practice and the impact of assets upon each other, and the importance of information (amongst other things) in this practice.

In 2014 the full standard ISO55000 is expected to replace PAS55 in 2014. It is based on the same 28 elements as the current PAS and will be published in three parts – 55000: Overview, principles and terminology, 55001: Requirements, and 55002: Application.

The architecture should ensure that asset management concepts and processes align with those defined in PAS55 / ISO 55000. [44]

4.1.2 Data processing

4.1.2.1 ISO 15926 Industrial automation systems and integration—Integration of life-cycle data for process plants including oil and gas production facilities

ISO 15926 is the data integration standard which originated in the process plant field but is now recognised to have a much wider scope. Its objective is to make data exchange easier and more reliable and is based on the concept of the two systems which are exchanging data not knowing anything about each other. This concept of encapsulation is widely adhered to in software to produce robust and extensible code and the same applies to enterprise architectures. By building an architecture with interfaces that do not depend on the specifics of any system the architecture can be extended to other usages relatively easily.

The standard recommends that data has the same structure and definitions to allow this. It also covers the communication of these (the use of OWL (Web Ontology Language)) and transmission. Also recommended is the separation of data model and data and the use of semantic metadata.

This standard is used by MIMOSA in their work to develop best practice in open data systems (see Section 6.1 for more on MIMOSA).

It can be seen that ISO 15926 is highly relevant to a data architecture for RCM, so application of the concepts of no-knowledge data transfer and encapsulation of data and architecture should be carefully considered in the design. [45]

44 Recommendation 037f
45 Recommendation 037d

4.1.2.2 ISO 18876 Industrial automation systems and integration - Integration of industrial data for exchange, access and sharing

ISO18876 provides a model for extensible data integration in the industrial area in the form of architecture, specification and methodology. It is suitable for integrating data from different sources, contexts, models and languages to share with applications. The standard includes integration models and methods for mapping, encoding and consolidating data sets. It specifies the selection, creation and extension of integration models, creation of application models to support the requirements of particular applications, and the creation of mappings between the integration and application models. These methods are independent of any particular language.

ISO 18876 contains useful methodologies, at a lower level than found in ISO 15926, and compliance should therefore be considered. [46]

4.1.2.3 Geo-spatial

The ISO19100 series of standards cover the representation and interchange of geo-spatial data. They have been included in other geographical standards described below so are not considered in detail here.

4.2 railML

4.2.1 Overview

railML is a data standard for the interchange of railway industry data. It is expressed as an XML Schema: a structured set of definitions, understandable by any person or machine reading the schema documentation and able to validate automatically whether any given example of data is correctly specified. If 2 systems apply the same schema then they will have a shared understanding of what things mean.

railML was started in 2002 with the objective of improving the ability of different railway IT applications to communicate and exchange data in a simplified way. Development is ongoing, with version 3 being due for release during 2014. Development is collaborative, involving any parties wishing to improve or extend the schemas in order to implement them with their own applications. Current usage of the schemas is restricted to Germany and Switzerland.

46 Recommendation 037c.

The current scope of railML is rail infrastructure, rolling stock, timetable and metadata schemas and the subschemas containing all of the elements which these contain. The images below show an example. Version 3 will incorporate the UIC's (International Union of Railways) RailTopoModel, a 'systematic model describing the railway infrastructure per its nature (and no longer for such or such usage)'. This data model aims at supporting all railways identified business needs, including concepts such as topology, track geometry, object and event location, referencing system and multi-level aggregation.

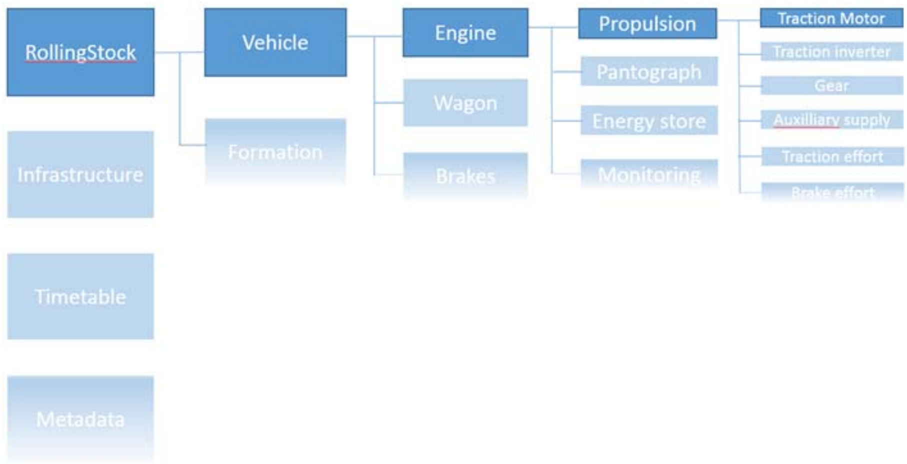
4.2.2 railML schemas

railML schemas are XML documents. XML is a widely used format for such schemas, the structure of the model being defined by the structure of the document itself and 'tags' used within it.

An element is declared with a heading and may be further defined by a type, by references or by descriptive data. Other elements can be nested inside that definition thereby associating one element with another.

Figure 6 shows a visual representation of an example from the railML rolling stock schema, in which it can be seen that traction motor is an element of propulsion, which is an element of vehicle, which is an element of rolling stock. In this way it is possible to define the structure and relationships between different elements in a simple text file which can be read and understood by a person or used to implement these definitions in an application which implements the schema.

Figure 6 - Diagrammatic representation of elements of the railML rolling stock schema



The following images show how the above elements are presented within an XML schema.

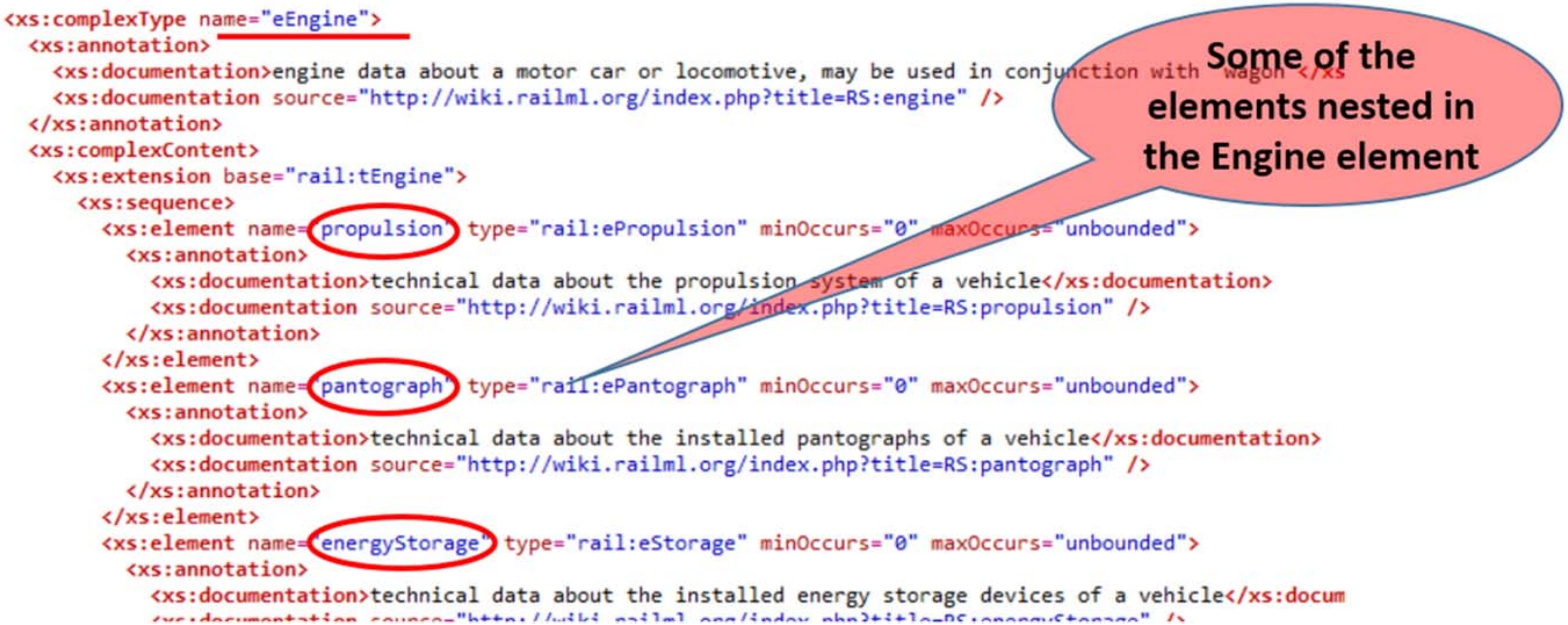
Figure 7 - XML representation of elements of the railML rolling stock schema

1 Rolling stock schema

2 Vehicle schema

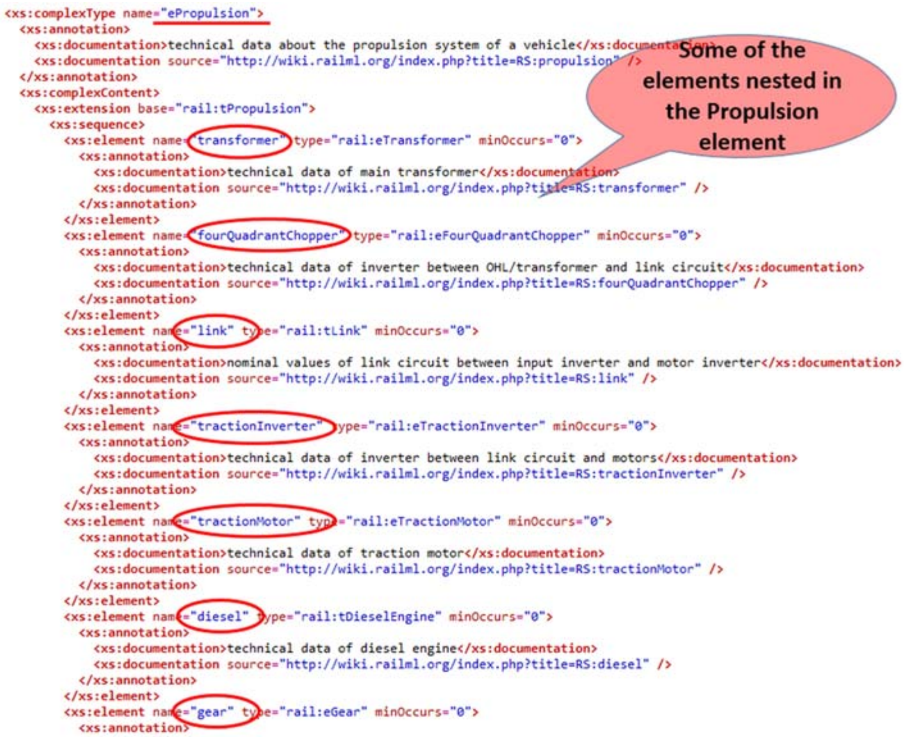

```
<xs:complexType name="eVehicle">
  <xs:complexContent>
    <xs:extension base="rail:tVehicle">
      <xs:sequence>
        <xs:element name="classification" type="rail:eClassification" minOccurs="0">
          <xs:annotation>
            <xs:documentation>general management and classification data of vehicle</xs:documentation>
            <xs:documentation source="http://wiki.railml.org/index.php?title=RS:classification" />
          </xs:annotation>
        </xs:element>
        <xs:element name="engine" type="rail:eEngine" minOccurs="0">
          <xs:annotation>
            <xs:documentation>engine data about a motor car or locomotive</xs:documentation>
            <xs:documentation source="http://wiki.railml.org/index.php?title=RS:engine" />
          </xs:annotation>
        </xs:element>
        <xs:element name="wagon" type="rail:eWagon" minOccurs="0">
          <xs:annotation>
            <xs:documentation>technical and payload data about a (motor) car or locomotive</xs:documentation>
            <xs:documentation source="http://wiki.railml.org/index.php?title=RS:wagon" />
          </xs:annotation>
        </xs:element>
        <xs:element name="vehicleBrakes" type="rail:eVehicleBrakes" minOccurs="0">
          <xs:annotation>
            <xs:documentation>data about brake systems independent from propulsion</xs:documentation>
            <xs:documentation source="http://wiki.railml.org/index.php?title=RS:vehicleBrakes" />
          </xs:annotation>
        </xs:element>
        <xs:element name="loadLimitMatrix" type="rail:eLoadLimitMatrix" minOccurs="0">
          <xs:annotation>
            <xs:documentation>matrix of permissible speed, line classification and payload</xs:documentation>
```

3 Engine schema




```
<xs:complexType name="eEngine">
  <xs:annotation>
    <xs:documentation>engine data about a motor car or locomotive, may be used in conjunction with "wagon" </xs
    <xs:documentation source="http://wiki.railml.org/index.php?title=RS:engine" />
  </xs:annotation>
  <xs:complexContent>
    <xs:extension base="rail:tEngine">
      <xs:sequence>
        <xs:element name="propulsion" type="rail:ePropulsion" minOccurs="0" maxOccurs="unbounded">
          <xs:annotation>
            <xs:documentation>technical data about the propulsion system of a vehicle</xs:documentation>
            <xs:documentation source="http://wiki.railml.org/index.php?title=RS:propulsion" />
          </xs:annotation>
        </xs:element>
        <xs:element name="pantograph" type="rail:ePantograph" minOccurs="0" maxOccurs="unbounded">
          <xs:annotation>
            <xs:documentation>technical data about the installed pantographs of a vehicle</xs:documentation>
            <xs:documentation source="http://wiki.railml.org/index.php?title=RS:pantograph" />
          </xs:annotation>
        </xs:element>
        <xs:element name="energyStorage" type="rail:eStorage" minOccurs="0" maxOccurs="unbounded">
          <xs:annotation>
            <xs:documentation>technical data about the installed energy storage devices of a vehicle</xs:docum
```

4 Propulsion schema



```xml
<xs:complexType name="ePropulsion">
  <xs:annotation>
    <xs:documentation>technical data about the propulsion system of a vehicle</xs:documentation>
    <xs:documentation source="http://wiki.railml.org/index.php?title=RS:propulsion" />
  </xs:annotation>
  <xs:complexContent>
    <xs:extension base="rail:tPropulsion">
      <xs:sequence>
        <xs:element name="transformer" type="rail:eTransformer" minOccurs="0">
          <xs:annotation>
            <xs:documentation>technical data of main transformer</xs:documentation>
            <xs:documentation source="http://wiki.railml.org/index.php?title=RS:transformer" />
          </xs:annotation>
        </xs:element>
        <xs:element name="fourQuadrantChopper" type="rail:eFourQuadrantChopper" minOccurs="0">
          <xs:annotation>
            <xs:documentation>technical data of inverter between OHL/transformer and link circuit</xs:documentation>
            <xs:documentation source="http://wiki.railml.org/index.php?title=RS:fourQuadrantChopper" />
          </xs:annotation>
        </xs:element>
        <xs:element name="link" type="rail:tLink" minOccurs="0">
          <xs:annotation>
            <xs:documentation>nominal values of link circuit between input inverter and motor inverter</xs:documentation>
            <xs:documentation source="http://wiki.railml.org/index.php?title=RS:link" />
          </xs:annotation>
        </xs:element>
        <xs:element name="tractionInverter" type="rail:eTractionInverter" minOccurs="0">
          <xs:annotation>
            <xs:documentation>technical data of inverter between link circuit and motors</xs:documentation>
            <xs:documentation source="http://wiki.railml.org/index.php?title=RS:tractionInverter" />
          </xs:annotation>
        </xs:element>
        <xs:element name="tractionMotor" type="rail:eTractionMotor" minOccurs="0">
          <xs:annotation>
            <xs:documentation>technical data of traction motor</xs:documentation>
            <xs:documentation source="http://wiki.railml.org/index.php?title=RS:tractionMotor" />
          </xs:annotation>
        </xs:element>
        <xs:element name="diesel" type="rail:tDieselEngine" minOccurs="0">
          <xs:annotation>
            <xs:documentation>technical data of diesel engine</xs:documentation>
            <xs:documentation source="http://wiki.railml.org/index.php?title=RS:diesel" />
          </xs:annotation>
        </xs:element>
        <xs:element name="gear" type="rail:eGear" minOccurs="0">
          <xs:annotation>
```

4.2.3 Netex

RailML is not the only XML-based standardisation initiative for sharing of railway data. Another longer-established one is NETEX/SIRI which has a considerable following in continental Europe where integrated transport systems with train, light rail and bus elements need to be coordinated. NETEX is a comprehensive standard which covers not just the data formats but also the interchange methods that can be used to transfer data, dealing with infrastructure, planning, and timetabling data. SIRI is a related set of standards that covers real-time operational and customer service data.

The RCM data architecture should take account of the careful thinking and depth of experience that have gone into NETEX/SIRI.[47]

4.2.4 Strategy

railML is of great interest in the design of an architecture for RCM because it is going to be supported by Network Rail's ORBIS programme, so there will be clear gains in shareability of data if railML is used to express rail concepts.[48]

4.3 T990 Train location strategy 2013

4.3.1 Overview

The train location strategy is an RSSB initiative, currently at the functional specification stage, for a system to improve the position information of trains through an iterative process of computing, sending, receiving and adjusting position information between a number of location services.

Position information sourced at the train includes on-board GNSS (GPS) signals, wheel revolution counters, inertial devices measuring acceleration, and Doppler radar devices measuring speed.

Along with computational functionality, this collated information enables the on-board train location service (TLS-OB) to compute the most likely position. TLS-OB can operate as a stand-alone system providing location information to on-board systems, or via a base station as TLS multi-train (MT). In its standalone capacity, TLS-MT can combine location data from multiple trains and make it available through a common interface.

47 Recommendation 235a

48 Recommendation 235

However the full functionality of the train location service would be realised through interfaces to TLS-MT from positional information in other systems such as the Train Running System TOPS (TRUST) which monitors train progress, and other positioning and train control systems, including trackside data sources such as balises, Radio Frequency Identification (RFID) devices and Wi-Fi networks. TLS-MT would use this additional information to adjust the estimation of the train's position, and feed this back to the train's TLS-OB. In turn, TLS-OB would take this adjusted estimate and use it along with current data coming from the train to make further adjustments to the estimated position.

4.3.2 Relevant findings

Whilst this initiative is still at the research stage, its findings are relevant to the RCM data architecture in two main ways – firstly it proposes a vision of a future system, which should be taken account of when planning for future extensibility of the data architecture, and secondly it provides a set of data and interface descriptions and definitions which will be useful in ensuring the data architecture is capable of dealing with positional data in as open and flexible a way as possible.

These definitions described in Table 4 should be considered in the design of the data architecture and the content of the schemas which will support it. Reuse of descriptions can save development time and contribute to consistency across industry systems[49].

49 Requirement 217

Table 3 - Definitions from Train Location Service Strategy 2013 which are relevant to RCM

Type	Detail	Reference
1 Data items	A description of data items in the train record, which is the set of data held by TLS-OB and TLS-MT about the location and identity of a particular train.	See Appendix A Section 1.
2 Reference information	A description of the reference information – background data that would be required in order to associate trains with timetable, map latitude and longitude to track positions and identify timetabled trains	See Appendix A Section 2.
3 Train identification	A description of the various train identifiers in use, and how they are currently used.	See Appendix A Section 3.
4 Standard interfaces	A description of the interfaces required to operate a TLS system as specified, of which some would be to current standards and some would be newly specified, to standard internet protocol. It should be noted that detection equipment must be registered as compliant with relevant standards	See Appendix A Section 4.
5 Interfaces to existing systems	A description of the interfaces that would be required with NR systems (some in development), ATOC, other infrastructure owners, and rolling stock providers.	See Appendix A Section 5.
6 Message formats	A description of the message formats required to enable the communications for a TLS system	See Appendix A Section 6.

In addition it would make sense to draw on other findings from this research project – more general considerations on a system of this nature, where it is sensible to draw on work done already.

1 Integrity - The computed position of a train can be classed as high, medium or low integrity, based on the consistency of signal, time since last update, and how many sources agree. High integrity information can be used for

safety critical applications, medium integrity information must have additional measures carried out before it can be used thus, and low integrity can never be used for safety critical applications. **Consider the application of different levels of integrity, how each level would be defined, and which sort of RCM activities each may be suitable for.** [50]

2 Communication Reliability – note that the effectiveness of the computation and subsequent integrity of position will be affected by intermittent reception, and where this is the case any application reliant on accurate train position will not be possible. **Consider whether the RCM data architecture needs to define acceptable levels of reception for the levels of integrity stated above.** [51]

3 Security – to prevent accidental or malicious damage it is recommended that hard wired connections are used wherever possible with access to ports strictly controlled and MPLS protocol used with all industry connections. Security protocols should be used for all message transfers. **Consider the security recommendations of the TLS train location strategy.**[52]

4 History – positional data will be held for a period at a specified degree of accuracy and then be archived off-line. **Ensure that the RCM data architecture defines for each sort of data attributes such as how long it needs to be stored, in what format, and at what level of granularity.** [53]

4.4 DHS Cyber Security Procurement Language for Control Systems

4.4.1 Overview

This document, produced by the US Department of Homeland Security, outlines a set of principles which should be followed by suppliers of Supervisory Control & Data Acquisition (SCADA) systems for controlling and monitoring important assets in order to make them secure against external cyber attack.

The security concerns associated with SCADA systems are grouped under the headings of confidentiality, integrity, and availability; with availability and integrity being seen as the most important in this context.

50 Recommendation 218

51 Recommendation 219

52 Recommendation 220

53 Recommendation 221

The approaches used to improve security of SCADA systems are described under the following headings, with interventions at each stage in the project lifecycle listed for each one:

- System hardening
- Perimeter protection
- Account management
- Coding practices
- Flaw remediation
- Malware detection and protection
- Host name resolution
- End devices
- Remote access
- Physical security
- Network partitioning

4.4.2 Relevance to the architecture

Most of the security concerns covered by this document relate to the design and configuration of the hardware and the networking infrastructure and so do not impinge directly on the data architecture. However, the following should be considered:

- Sensors and loggers must be able to be configured with the minimum number of access accounts and privileges. The architecture should support the verification and implementation of this [54]
- Web user interfaces and web services interfaces should use best practice to avoid common attacks on such types of interface, including authentication, 'man in the middle', command injection, SQL injection, spoofing [55]
- The guidance in this document should be seen as contributing to RSSB's overall guidance on cyber security [56]

54 Recommendation 246

55 Recommendation 247

56 Recommendation 320

5 Guidance from previous RSSB RCM projects

This section describes the work done in previous research projects carried out by RSSB in the field of remote condition monitoring. These represent the direct precursors of the current project and the influences on its scope and direction.

5.1 T844 Mapping RCM activities to the System Reliability Framework

5.1.1 Overview

In 2009 the RSSB research project T844 was commissioned by the Uninterrupted Journey Group (UJG), a cross-industry group facilitated by RSSB, to discover how existing and planned RCM systems were improving or could improve reliability, and what areas of RCM had most potential for further development.

It is thus of relevance to T1010-01 in that it highlights those established or important systems for whose data we should consider supporting access with the architecture, and those areas with most potential for future development and reliability savings which the architecture must allow to be supported in future.

5.1.2 Systems Reliability Framework

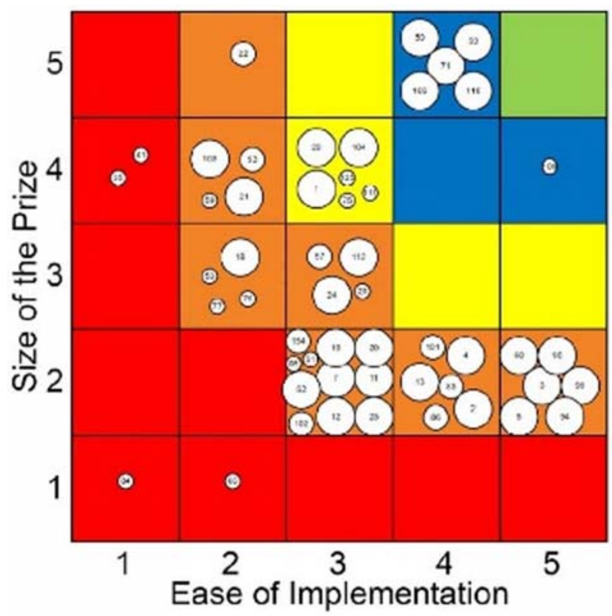
The Systems Reliability Framework is a framework representing the functional aspects of the rail industry in categories such as infrastructure, train, rules, processes, operations, and so on, which are associated with a percentage showing the impact of that category on reliability (expressed as delay minutes). In Figure 7 it can be seen that infrastructure- based delays cause 29 % of reliability failures and train-based delays 15 %.

Each category is then broken down to show attributable systems or events and their share of the failings due for that category. For example, in this case, battery/control is attributable for 25 % of train reliability failings, doors 20 % and underframe and bogies 5 %.

5.1.3 RCM activities identified and mapped

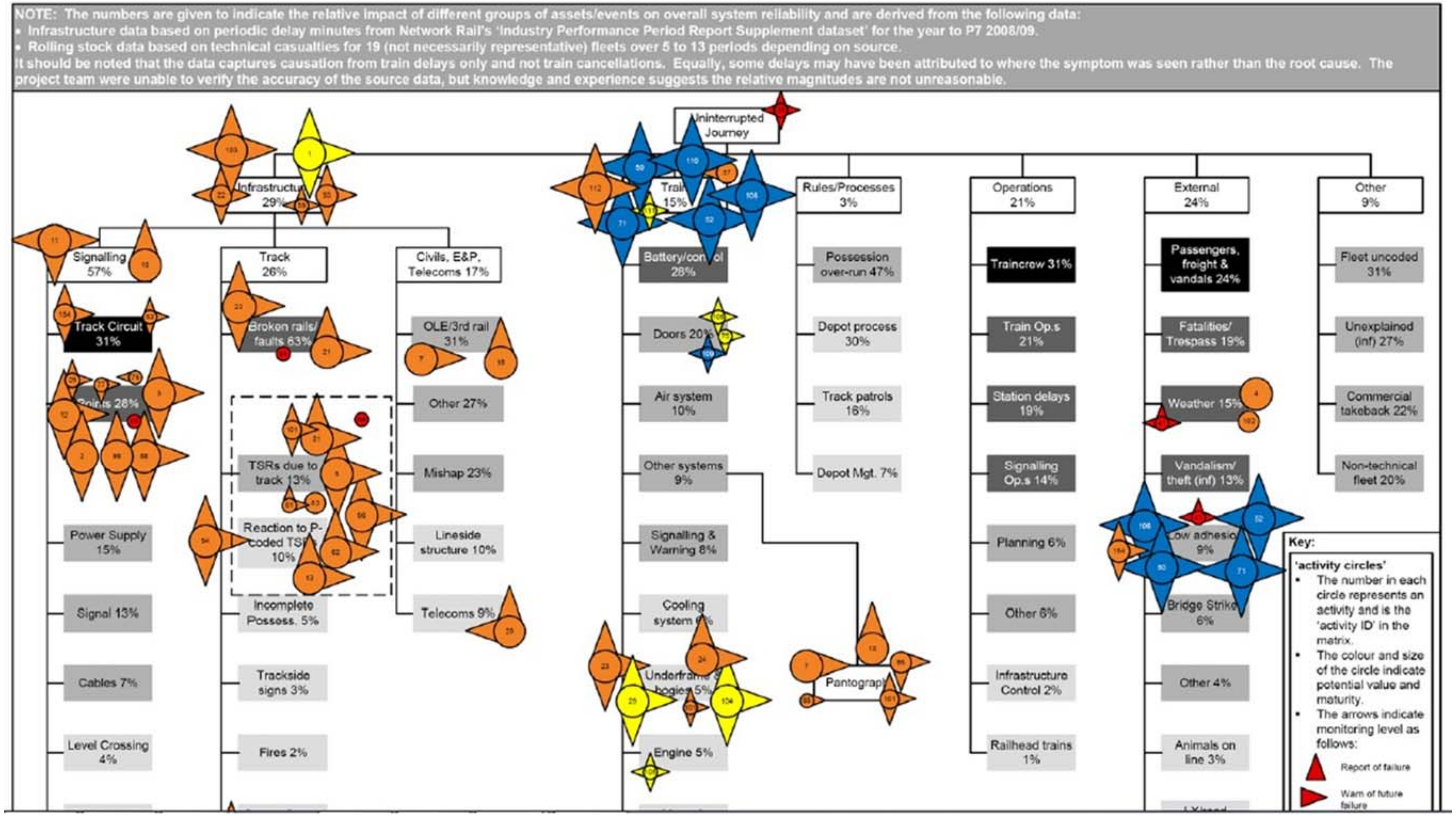
A shortlist of 52 RCM activities with the potential to assist in improving reliability was created. Based on a series of interviews with industry experts, each RCM activity was scored for the 'size of the prize' and 'ease of implementation'. Every activity was placed in a band according to its value, the activities in blue being highest value and those in red being lowest. Maturity of the system is also assessed, with the more mature activities being represented by larger circles, indicating immediate rather than potential value.

Figure 8 - RCM activities graded



These activities were then mapped to the Systems Reliability Framework to illustrate what areas of reliability they would address. The resulting clustering gives an indication of which areas were being most addressed by RCM developments and, based on the level of delay caused by that area, the resulting reliability improvements were estimated.

Figure 9 - RCM activities mapped to Systems Reliability Framework



5.1.4 Relevant findings

Firstly it should be noted that the report does not recommend low-scoring activities should not be supported. For example, InteGRail is classed in the lowest band due to scoring badly for ease of implementation, but it does represent a long-term future state. **So the banding and analysis in this report should be used as a guide and not treated as prescriptive**

Train-related RCM activities (meaning those measuring an aspect of an asset of a train either from the train itself – in the Train Monitors Train quadrant - or from a fixed sensor – in the Train Monitors Infrastructure quadrant[57]) scored the highest in value along with those concerned with adhesion monitoring. These activities were also found to be the most mature, and are all being carried out in the UK. Train-borne RCM equipment , (of which some but not all will monitor train assets), is also being rolled out to new and existing fleets so it can be expected that there will be more of this activity in the future. **Train asset-related RCM activity is thus of most interest to those developing an RCM data architecture. They should be investigated with the intention that access to their data will be supported by the architecture. Refer to Appendix B Section 1.**[58]

Hot Axle Box detection is a well-established form of infrastructure-based monitoring of trains with some equipment in second or third generations. The most recent generations are capable of reporting sub-alarm-threshold temperatures; and if integrated with an AVI solution will be able to alert about degrading bearings prior to their actual failure. **Therefore it ought to be ensured that HABD data is fully defined in the architecture and possibly used as a case study.** [59]

AVI is an essential enabling system. **We must acknowledge this recommendation of the author when making our own recommendations on vehicle identification.** [60]

Real-time downloads are a necessity. **The architecture needs to take into account the requirement for a degree of real-time downloads, however the extent of this is to be investigated and defined. It should be noted that the**

57 For further information on the cross-industry monitoring quadrants refer to the Introduction to this document,

58 Recommendation 321

59 RCM Activity 205

60 Raw Dependency 094

subsequent report T853, Mapping the Remote Condition Monitoring Architecture, found that this was only a requirement for key events[61].

Infrastructure monitoring, that is, activities where the infrastructure is monitored by fixed or mobile sensors in the Train Monitors Infrastructure and Infrastructure Monitors Infrastructure quadrants[62], whilst generally being of less value than train monitoring, having a smaller ''size of the prize'', is stated to have the greatest potential to become valuable. There are a much larger number of activities relating to infrastructure. Most are concerned with signalling (track circuits and points), track (broken rails and TSRs), electrification and telecoms. Of these, track circuits, points and broken rails or rolling contact fatigue account for the greatest reliability issues**. To maximise realisation of the benefits of RCM in the medium to long term it should be ensured that these activities are taken account of when agreeing scope of the architecture. See Appendix B Section 2.** [63]

Application of remote condition monitoring across the industry is not enough to deliver the benefits generally expected. That will depend upon the analytical stages, as improving predictive algorithms and the way they are used is necessary to deliver significant improvements in reliability, 'to better link condition to incidence of failure'**.** [64]

More sophisticated analysis through improved algorithms is thus a major dependency for the full realisation of the benefits of RCM and must be noted as such in the architecture documentation. [65]

Accordingly, a key requirement for the architecture will be to allow access to the data required by the algorithms from internal and external sources. [66]

A further requirement will then be to ensure that the output of the analytical stages can be stored or transmitted to allow it to be acted upon. [67]

The report defines RCM according to 4 levels of activity: reporting a failure, predicting a failure, predicting a failure and identifying a cause, and enabling

61 RCM Activity 072

62 For further information on the cross-industry monitoring quadrants refer to the Introduction to this document,

63 Recommendation 322

64 Mapping Current RCM Activities to Systems Reliability Framework p42

65 Dependency D02

66 RCM Activity 035

67 Recommendation 323

failure to be designed out. It found that most of the 52 activities included all apart from the final level. Care should be taken to take account of existing systems and processes in the design of the architecture. [68]

5.2 T853 Mapping the RCM architecture 2010

5.2.1 Overview

This RSSB research project followed on from T844, Mapping RCM activities to the System Reliability Framework. It maps the architecture of RCM activity in the UK in 2010, focussing on the effectiveness of the activities. The project makes recommendations for implementing improvements including architecture, interfaces and information standards. As such it represents the main predecessor to the current work.

Key to this report is the 6-level processing model for RCM which originates from the ISO standard 13374 Condition Monitoring. This model covers every sort of activity which may be carried out as part of a condition monitoring process via the 6 levels in Figure 10.

68 Recommendation 013a and 117a

Figure 10 - Six-level condition monitoring processing model and associated activities

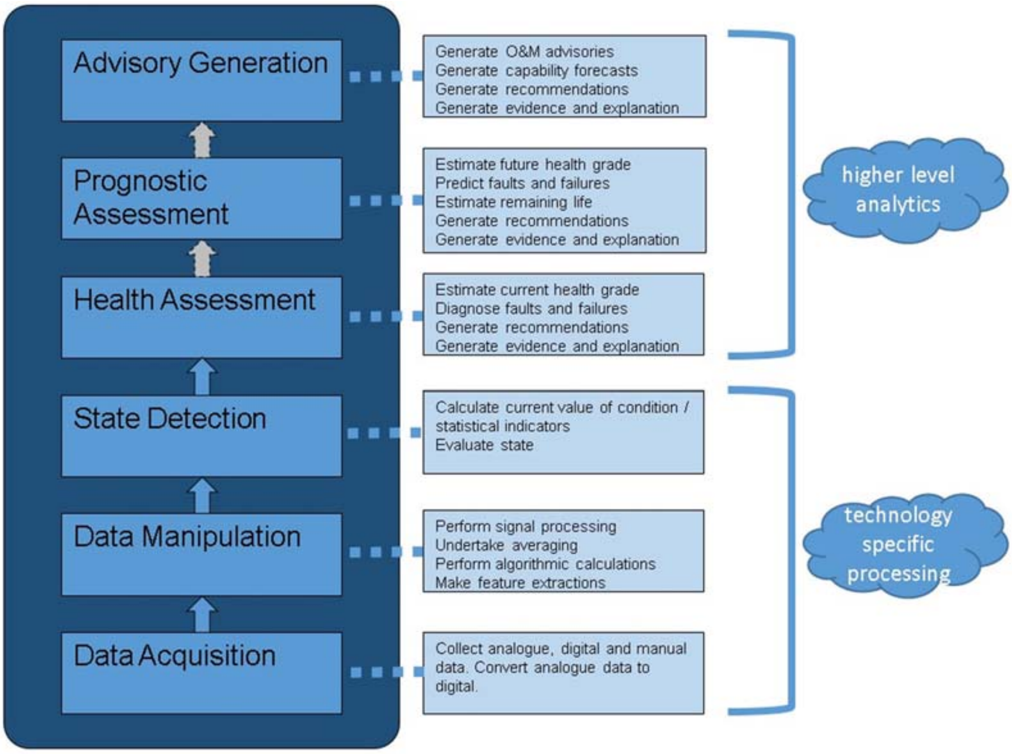


Figure 10 is based on that presented in T853 however it has been inverted to fit with the way that the model is presented elsewhere in T1010 documentation.

The 6 levels of this hierarchy have a relationship with the 4-level RCM activity stack (Report Failure / Warn of Impending Failure / Warn with Cause Identified / Facilitate Design-Out of Future Failure) as shown in Table 5.

Table 4 - T844 vs ISO 13374 mapping

T844 Level	Report failure	Warn of impending failure	Warn with cause identified	Facilitate design-out of future failure
ISO 13374 Level				
Advisory generation				Yes
Prognostic assessment		Yes	Yes	
Health assessment			Yes	
State detection	Yes			Yes
Data manipulation	Yes			Yes
Data acquisition	Yes			Yes

The report concludes that the majority of existing condition monitoring projects meet the definitions of state detection or health assessment, with train monitoring activities being at a more mature level of development than infrastructure monitoring activities.

An architecture which allows the transmission and sharing of condition data more easily will allow more intelligent use of that data and therefore allow an increase in the RCM maturity levels towards ‘prognostic assessment’ and ‘advisory generation’.

The model must be fully explored to inform the development of an architecture that encourages increased RCM maturity levels. [69]

Relevant information from this document can be classified into two groups – ‘as is’ findings and ‘to be’ recommendations.

69 RCM Activity 057 - 062

5.2.2 Relevant findings: 'as is'

1 Most current systems conform to a generic architecture of sensor, logger, server, and workstation, with a limited number of interfaces, developed on a case-by-case basis.

2 Suppliers have proprietary black box algorithms which can be hard to analyse and compare, and proprietary data formats. This restricts sharing. Many have embedded firmware-based algorithms, well suited for large scale mature systems. Intelligent Infrastructure is developing an improved set of algorithms to assess state and health. A risk is stated of loss of the experience embedded in the existing algorithms.

3 Current infrastructure-based train monitoring systems are restricted by the lack of a consistent vehicle identification system, and train-based infrastructure monitoring systems are restricted by the lack of a consistent location identification system. These gaps will be key dependencies for the future success of RCM. [70]

5.2.3 Relevant recommendations: 'to be'

5.2.3.1 Standards

A number of relevant standards are suggested in the report:

- ISO 13372 for terminology of condition monitoring
- ISO 13374 for processing model
- ISO 18876 and ISO 15926 for data integration architecture
- InteGRail as foundation for developing ontologies, standard interface adaptor, common service grid, distributed storage and processing

The standards listed above should be reviewed with a view to building compliance to them into the architecture, so unless there is a reason to decide otherwise for any element, these standards will be reflected in the requirements and ultimately, in the architecture. [71]

5.2.3.2 Six-level processing model

The RCM systems reviewed by the project were found to typically support RCM activity up to the 'Health Assessment' level.

70 Dependency D06

71 Recommendation 037b, 037c, 037d

The report's Vision section proposed the development of RCM systems to the 'Prognostic' level, mid-level for track and high-level for fleet. A number of things were recommended to support this:

- More sophisticated algorithms (which are being developed by the Intelligent Infrastructure project)
- The possible use of rule-based technologies in place of fixed algorithms
- Libraries of deterioration profiles, issues and utilisation, with asset data held consistently and possibly centrally
- Access to data from other RCM and other railway operational systems

To support delivery of the 'advisory generation' level, the report recommends better information on asset configurations, and implementing links to planning, work management and logistics systems.

1 This processing model must be considered as a strong contender on which to base a cross-industry RCM data architecture, in which case providing interfaces between each of the six levels would be an important requirement. [72]

2 In particular, attention will be paid to the functions of the Prognostic and Advisory levels and ensuring that the architecture supports activities at these levels, such as enabling algorithms and use of library data, which will deliver the most benefits to the industry. [73]

3 Dependencies will be created to record the need for library data such as deterioration profiles. [74]

5.2.3.3 The Intelligent Infrastructure programme

Implementation of the full potential of this initiative (to deliver improvements to infrastructure management through the application of intelligence in design and maintenance) is recommended by the report, as it is considered an enabler of the required asset data libraries and more sophisticated algorithms. Intelligent Infrastructure implements centralised storage and processing of asset condition data and the transmission to support this and is considered to support the attainment of health assessment/prognostic assessment processing levels for infrastructure systems. For more on Intelligent Infrastructure see Section 5.5.

The components of Intelligent Infrastructure must be considered in the development of the architecture. [75]

---

72 RCM Activity 057 - 062

73 RCM Activity 061-062

74 Dependency D01

75 Recommendation 078

5.2.3.4 InteGRail

The report recommends implementing the following ‘products’ developed by the InteGRail research project to provide a sound platform to support the vision of improved data sharing, for the reasons detailed below.

- The Intelligent Communication Infrastructure (ICOM) and Rail Domain Ontology (RDO) components have proved data can be effectively shared whilst not residing in same system
- The InteGRail Service Grid (ISG) and Flexible Communication Adaptors (FCAs) together allow a single interface to operate with multiple systems thus reducing number of interfaces and associated costs. A common interface is ‘essential’ to realise amount of integration required as otherwise there would be a large number of point to point interfaces with high costs

The architectural design must align with the architectural concepts developed by the InteGRail project. These are explained in Section 3.2. [76]

5.2.3.5 OWL-based data standard transmissions

The research project identified the Web Ontology Language (OWL) as a suitable language for the Rail Domain Ontology, for communication of asset and condition information across different organisations and potentially different industries, without the need for complex interfaces. More is included in Section 3.2 on InteGRail and Section 3.3 on ontologies.

The use of OWL should be strongly considered in the further development of the architecture to assure consistency with data and objects in other industry initiatives such as ORBIS which may use elements of the InteGRail approach. [77]

5.2.3.6 Location and vehicle identification

Both these current gaps are identified by the report as being key dependencies in the realisation of the full benefits of many RCM systems in improving reliability.

The architecture must allow for the integration of location and vehicle identification data into the processes it supports. The project must flag

76 Recommendation 091

77 Recommendation 097

these gaps as key dependencies to the full benefits of investment in RCM being realised. [78]

5.2.3.6.1 Other points to note

- Most data will not have to be real time however some key events will. [79]
- Data may be analysed on train or shore. [80]
- Various transmission methods are in use to get data from sensors to receiving computers. The specific method used in any given project is determined case-by-case. There is no evidence that RCM data interchange is impeded by transmission limitations such as cost or bandwidth. [81]

5.3 Data framework feasibility study 2011

5.3.1 Overview

This is the final report for a research project carried out by the Universities of Birmingham and Nottingham into the feasibility of a 'cross-industry, semantic' data framework for the UK rail industry. The original purpose was to investigate the benefits of such a framework towards the reduction in carbon from a shift to rail, however the study's conclusions support a framework for applications across the whole spectrum of industry activities. At a requirements workshop, stakeholders from across the industry identified 153 different applications for a data framework, of which 40 related to asset management.

Following this a prototype domain model was developed, for a 'train monitoring track' scenario.

The report recommends a 'conceptual data model' as the basis of the framework, i.e. one that includes semantic information amongst its metadata. Specifically the use of Web Ontology Language (OWL) is proposed. The study also included substantial research into the use and development of ontologies, particularly into the methods used to elicit domain knowledge.

78 Dependencies D03 & D06, Recommendation 319

79 Requirement 072

80 Constraint / Context 083

81 Constraint / Context 186

5.3.2 Relevant findings

5.3.2.1 Support for a common data framework

The report strongly recommends a common data framework as an alternative to a series of stand-alone interfaces, quicker and cheaper to set up in the short term, but with reduced flexibility and additional maintenance and extension costs and difficulties. The report specifies that existing data models should be integrated with this and not replaced.

Take account of recommendation to integrate and not replace existing data models. [82]

5.3.2.2 Support for a conceptual data model

The report makes the case for a conceptual data model as the basis of a common data model for the railway industry; one which includes semantic data with its metadata and thus allows machine to reason with information. It specifies the Web Ontology Language OWL as a suitable language with which to implement this.

Include the recommendation for ontological data and the use of OWL in RCM data architecture requirements. [83]

5.3.2.3 Suitable ontology engineering method

Further to the above recommendation for a semantic data model, a review was conducted of the different approaches to building this sort of data model and developing its content, in other words, 'ontology engineering'. In the review the collaborative DILIGENT methodology is suggested as the most appropriate for the task of a common data framework for the rail industry, as it allows for expertise from different sources to be amalgamated. The DILIGENT[84] methodology is based on the understanding that a full and accurate definition of a subject will depend on input from a variety of different subject matter experts from different geographic or organisational areas, and the process involves argument with the aim of building a consensus.

In practice, for the prototype development work, the more monolithic METHONTOLOGY[85] approach was used, being better suited to the lack of time

82 Recommendation 013a and 117a

83 Recommendation 044 and 097

84 Pinto, Staab, Sure, & Tempich, 2004; Pinto, Staab, & Tempich, 2004; Tempich *et al.*, 2007)

and resources. Monolithic methodologies develop ontologies based on the input of a more limited group: an individual or a single team. The use of ontology editing software such as Protégé is recommended.

Take account of these recommendations for a collaborative approach, where time and resources allow, when planning the development work. [86]

5.3.2.4 Eliciting domain knowledge

A thorough explanation is given of the difference between explicit and tacit domain knowledge, and the natural and artificial techniques used to elicit this from the processes and individual experience wherein it resides. Upon this knowledge, ontologies are built. **In planning the consultation, thought should be given to what sort of knowledge is required and suitable ways to elicit it.** [87]

5.3.2.5 Eliciting requirements

Although distinct activities, the requirements process was included in the knowledge elicitation referred to above to save time. The process was in three stages – a workshop to generate potential uses of the proposed initiative (breadth), a knowledge elicitation session to gain depth in one area, and then a later session with stakeholders to validate the developed concepts. **In the progress of the RCM project, it ought to be ensured that there are activities to provide the appropriate breadth and depth of knowledge and to validate findings with the steering group.** [88]

5.3.2.6 List of 40 asset management scenarios

A list of 40 asset management scenarios were selected by the project stakeholders as being able to benefit from a cross-industry data framework, and are therefore of great relevance to this project. Within this 40 there is likely to be a significant amount of overlap or duplication **however this list can be used as an initial guide to priorities and scope of the architecture, and a starting point for the consultation process.** [89] **Refer to Appendix G where this list is shown, with those relevant to an RCM data architecture highlighted.**

85 Fernández *et al.*, 1997

86 Recommendation 208

87 Recommendation 210

88 Recommendation 210

89 Recommendation 211

5.3.2.7 Prototype domain model – train monitors track

Within the workshop, analysis was carried on track to train condition monitoring, with participants identifying benefits, barriers and data types and sources. **The data types and their sources, which are of particular relevance to the design of an RCM data architecture, are shown in Table 5.**

Table 5 - Data types and their sources - train monitors track

Data	Notes	Source
Train formation	Unit and components, vehicle number, orientation	
Headcode	Vehicle ID, automatic?, down to level of wheels	Ops
Location of measuring equipment	To within 10-20m	
Properties measured, units		
Vehicle normal parameters		
Trended from historical data		
Appropriate operator and maintenance group		
Maintenance schedule and records		
Measurement time	Close enough to timetable time	CCF

Further to the workshop this scenario was developed into a prototype domain model, comprising a description of the ontologies that would be required to support this case study, and then the creation of the vocabulary lists which would form each ontology.

The existence of these prototype ontologies will be of great use in the development of the architecture and should be expanded upon to ensure other monitoring activities deemed to be in scope are supported. [90] See the example vocabulary lists from the Data Framework report in Appendix H to this report.

5.3.2.8 Prototype domain model – what can be learnt?

As well as the obvious value of the content of the prototype domain model, there is a lot of value in learning from the methods the project team used to create it.

Drawing on the content of the Rail Domain wiki saved much development time – though at the time of writing further development was required to the wiki.

Re-use of existing ontologies such as OWL-time and location, possibly through the implementation of upper level ontologies such as SUMO or ISO 15926 should be considered. [91]

It is evident that the same data can support different kinds of business process, either the generation of knowledge, the proving of known or partially known facts and the diagnosis of known issues. Algorithms and user interfaces need to be tailored according to the process they support rather than the source from which the data originates. [92]

The role of time must be considered: for example, is the data needed in real-time?[93] [94]

5.3.2.9 Issues

Issues raised in the report must be captured within the RCM project documentation as either constraints, contexts or dependencies. For example, data quality issues such as mileage and data entry errors ought to be addressed in order for the information made available by the architecture to be trustworthy.

---

90 Recommendation 212

91 Recommendation 115,116a and 076b

92 Recommendation 090

93 Note the different uses for real-time data (operational use) and the use of historical data and planning systems for tactical and strategic uses.

94 RCM Activity 072

5.4 RSSB cross-industry information systems workshop 2012

5.4.1 Overview

This workshop, held at the end of 2012, was attended by a number of industry experts involved in some key ICT initiatives within UK rail; the workshop was an exercise in assessing the current state of data initiatives and their potential usefulness, with a view to development of an industry-wide information architecture involving reuse of technology between the initiatives.

The output included an analysis of the commonalities between the initiatives (overlaps in objective, data source and data provision, and information gaps) and illustration of how these initiatives could be used together to address a number of information usage scenarios.

The initiatives considered are shown in Table 6:

Table 6 - Initiatives considered in cross-industry information systems workshop

Initiative	Description
Offering Rail Better Information Services (ORBIS)	Undertaken by Network Rail with the aim of improving its approach to acquisition, storage and usage of asset information
	Improved asset information capabilities through information driven asset management. Work to date includes putting in place master data management (including asset location using GPS amongst other methods) and validating and cataloguing the asset information base. Future work will be to use this along with existing information systems to allow 'predict and prevent' asset management policies, work planning and asset location.
Cross-Industry Remote Condition Monitoring (XiRCM) Group	A subgroup of theVehicle/Vehicle SIC, facilitated by RSSB, with the aim of increasing the use of cross-industry RCM.
	Of which T1010-01 is a part, with the objective of implementing the enablers of data sharing.

Table 6 - Initiatives considered in cross-industry information systems workshop

Initiative	Description
Defect Recording Analysis & Corrective Action System for Control & Command Systems (DRACAS CCS)	RSSB research project, T754, T957 and T960 commissioned by the Vehicle/Train Control & Communications SIC to support greater convergence of existing industry systems for the development of DRACAS within the railway industry.
	An asset management and knowledge repository system enabling decision making on operation and management of assets. A single common process for CCS DRACAS is proposed for which an architectural design has been produced.
Traction and Rolling Stock Systems	Managed by the Traction and Rolling Stock Systems Management Group. Currently considering 3 initiatives – Railway Industry Standard for shared rolling stock data (RS RIS), an update of RAVERS and RSL rolling stock information systems (R2) and a rolling stock systems architecture (RSSA) to integrate systems and enable whole-system approach.
Layered Information Exchange for Traffic Management (LINX TM)	Undertaken by Network Rail
	Providing interfaces between traffic management and other industry systems via a data bus to create a 'single view of truth' to inform users of real-time timetable changes.
Network Rail GPS Gateway	GPS data gathered by train operators into a single source for the benefit of systems across the industry.

5.4.2 Relevant findings

5.4.2.1 Overlap between DRACAS and LINX TM

The workshop report concludes that both the DRACAS and LINX TM systems had considerable overlap with RCM systems (amongst other initiatives). RCM data would feed historical asset information to DRACAS and current advisory condition information to traffic management systems. **This should inform the requirements gathering process for the architecture, in order to ensure that the greatest potential for RCM data across the industry may be realised through prioritisation.** [95]

5.4.2.2 Information gaps: vehicle identification, vehicle location

In addition these initiatives share key information gaps, being the availability of reliable vehicle identification and asset location data. GPS data, whilst being found to have relatively few applications amongst the other initiatives, was considered to be a key dependency for train based systems that are used to monitor infrastructure. **These findings validate the understanding of the dependencies which exist for the full realisation of the benefits of RCM.** [96]

5.4.2.3 Role for ontologies

A need for common data architecture(s) with ontologies was identified for the benefit of all initiatives, and the difficulty in defining data across different processes was highlighted. At least four of the initiatives were found to require a data bus and the potential to reuse technologies was discussed. Also discussed was the vision for an integrated future industry, with ideas such as completely open rail data, demand-driven train operation, and European interoperability. **These concepts should inform the architecture requirements and the way that extensibility is incorporated.** [97]

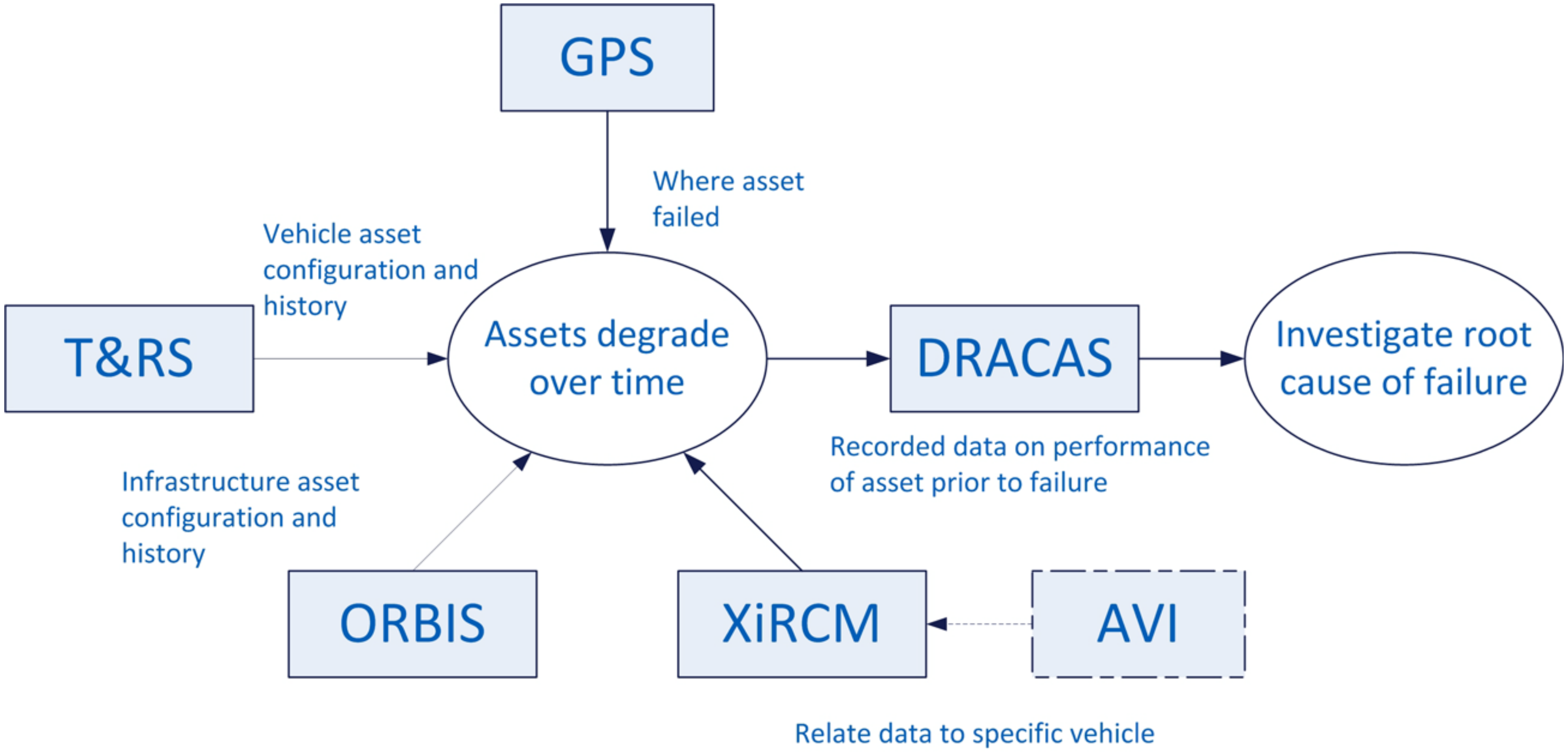
Based on the descriptions of the scenarios discussed in the workshop, diagrams of those relevant to RCM data usage have been drawn up by the author of this report, to illustrate the involvement of different initiatives in each scenario. These are shown in Figures 11 to 14. The data initiative is shown in boxes, with the arrows indicating the sort of data that the initiatives are concerned with the provision of. The information scenario is shown in an ellipse.[98] Information gaps are shown with dotted lines.

95 Recommendation 215

96 Dependencies D03 & D06

97 Recommendation 044 and 216

Figure 11 - Data initiatives involved in the scenario of 'Assets degrade over time'



98 Produced by the author of this document and not the report from the workshop

Figure 12 - Data initiatives involved in the scenario of 'Establish infrastructure condition in real time'

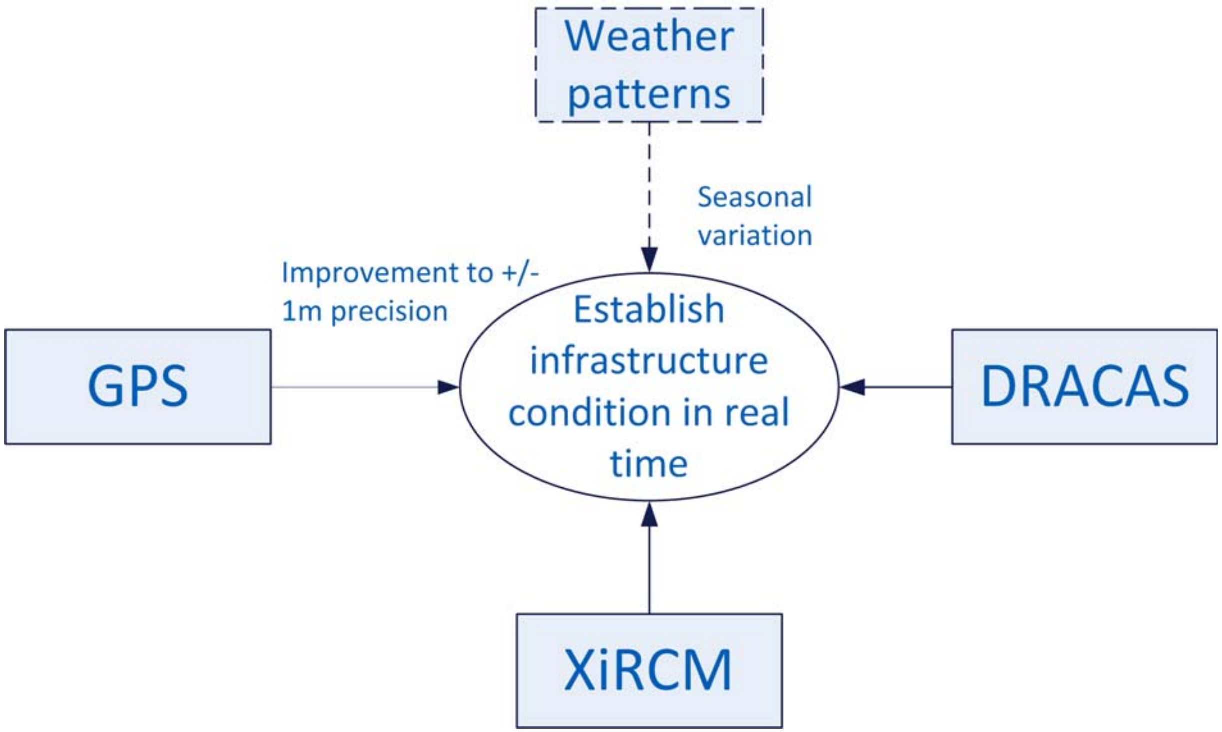


Figure 13 - Data initiatives involved in the scenario of 'Determination of asset condition'

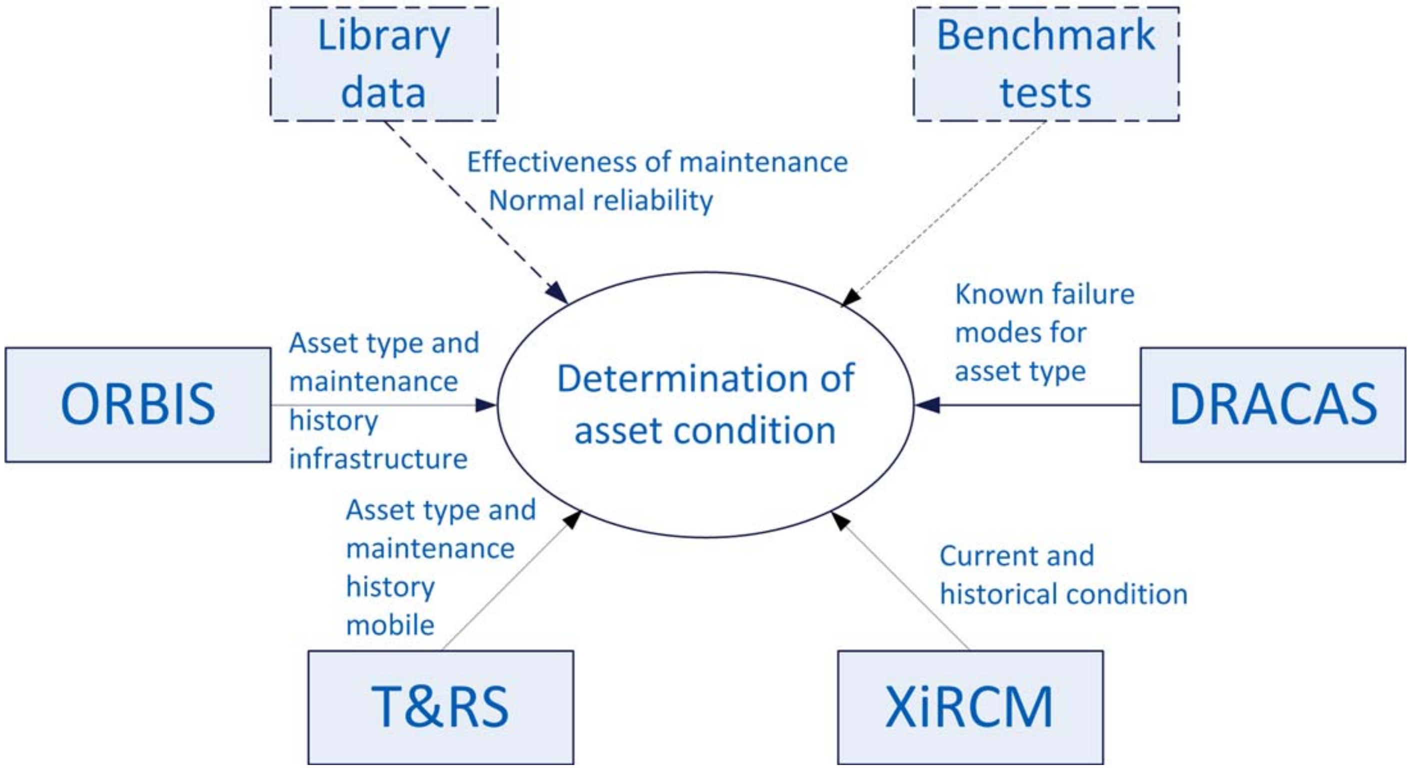
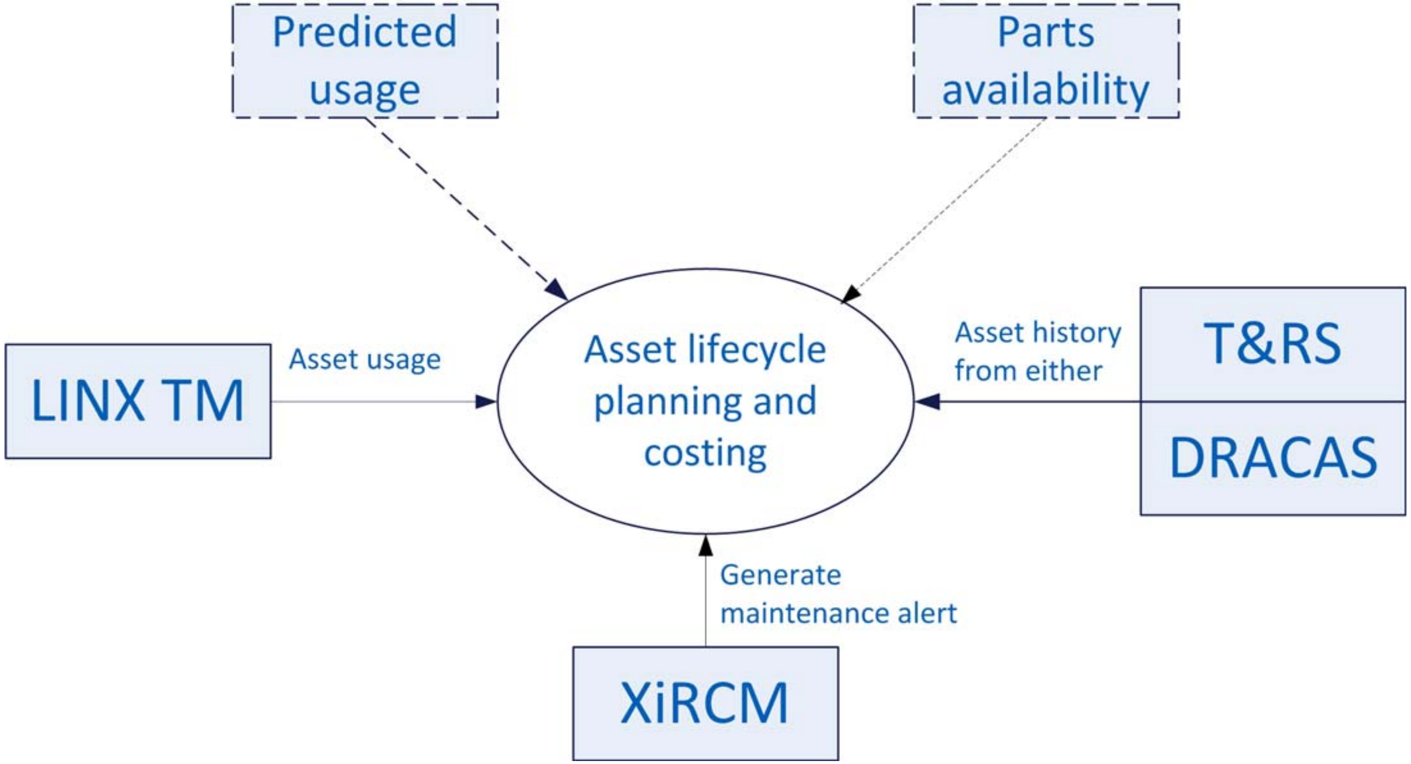


Figure 14 - Data initiatives involved in the scenario of 'Asset lifecycle planning and costing'



5.5 The Intelligent Infrastructure strategy

5.5.1 Overview

Intelligent Infrastructure is a Network Rail strategy with the objective of delivering improvements to infrastructure management through the application of intelligence in design and maintenance. Intelligent design is intended to design out failure, and intelligent maintenance is required to support a move from a 'Find and Fix' approach to 'Predict and Prevent'. For example, the benefits are expected to include a reduction or elimination of manual inspection, fixing assets before failure, self-repairing or adjusting assets, automatic deterioration detection and understanding of root cause.

These benefits would be realised via self-monitoring assets, the existence or provision of more data, with centralised storage and transmission via XML, and open and consistent algorithms with which to process it.

The Intelligent Infrastructure Good Practice Guide has been reviewed in this section. It should also be noted that the related Intelligent Infrastructure project, with the remit of establishing the technical capacity to implement this strategy, is of interest to the development of the RCM data architecture, and consultations have been on the technical elements of this aspect: see Section 7.2.4.

5.5.2 Relevant aspects

Clearly the objectives of the Intelligent Infrastructure strategy are closely related to those of the cross-industry RCM initiative, both being concerned with the enabling of intelligent asset management through the ability to access, process and act on data from diverse suppliers. **As this is proposed to be implemented fully (to enable aspects of RCM such as improved algorithms and deterioration profiles as stated in T853) the design of the RCM data architecture should consider applying the requirements and functionality of the Intelligent Infrastructure strategy, in particular:**

Ability to drill down data to geographic levels (national/corporate, route/line/area, station/junction, asset) and equipment levels (all assets and asset groups, systems, subsystems, components) and to allow reporting on a matrix of these. [99]

99 Recommendation 223

Analysis may take place either locally at the asset or centrally, with the analysis software capable of operating in either. [100]

The strong focus on standardised central storage and transmission. [101]

100 Recommendation 224

101 Recommendation 078

6 Best practice examples

The documents described in this section relate to data integration initiatives from other industries which may have value to the RCM Data Architecture initiative. The value may come from the data content or the process followed.

6.1 MIMOSA and OSA-CBM

6.1.1 Overview

MIMOSA is an association dedicated to the development of open information standards for manufacturing, fleet and facilities. It has developed the Open Systems Architecture for Condition Based Monitoring (OSA-CBM) which is currently used to enable collaborative asset lifecycle management in military and commercial applications.

The OSA-CBM comprises a set of specifications required for the operation of an open data system in the condition monitoring field, including interfaces, message formats and the information and data types that make up the architecture. These are provided in the form of UML diagrams, Enterprise Architecture packages and diagrams and XML schema definitions.

It provides an implementation for the concepts of ISO 13374, the standard for condition monitoring, which is of particular interest in the specification of an RCM data architecture and about which more can be read in Section 5.2.3.1 of this report.

Included in these specifications is each of the six processing levels of RCM identified in ISO 13374 – Data Acquisition, Data Manipulation, State Detection, Health Assessment, Prognostic Assessment and Advisory Generation. As discussed elsewhere in this report, these are likely to be the foundations of the architecture. In addition there are other relevant definitions such as alerts, site, data types, and data event sets. The screenshot in Figure 15 shows part of the XML schema of the Advisory Generation specification, including ItemRequestForWork which has elements such as 'start_after', 'end_before', 'auto_approve' and 'work_order' nested inside.

Figure 15 - Extract from OSA-CBM Advisory Generation XML schema

```xml
            </xs:sequence>
        </xs:extension>
    </xs:complexContent>
</xs:complexType>
<xs:element name="ItemRequestForWork" type="ItemRequestForWork"/>
<xs:complexType name="ItemRequestForWork">
    <xs:annotation>
        <xs:documentation>This class contains information related to work requests for items.</xs:documentation>
    </xs:annotation>
    <xs:sequence>
        <xs:element name="auto_approve" type="xs:string" minOccurs="0" maxOccurs="1">
            <xs:annotation>
                <xs:documentation>'Y' or 'N'</xs:documentation>
            </xs:annotation>
        </xs:element>
        <xs:element name="start_after" type="OsacbmTime" minOccurs="0" maxOccurs="1">
            <xs:annotation>
                <xs:documentation>Request for action to start after this time.</xs:documentation>
            </xs:annotation>
        </xs:element>
        <xs:element name="start_before" type="OsacbmTime" minOccurs="0" maxOccurs="1">
            <xs:annotation>
                <xs:documentation>Request for action to begin before this time.</xs:documentation>
            </xs:annotation>
        </xs:element>
        <xs:element name="end_after" type="OsacbmTime" minOccurs="0" maxOccurs="1">
            <xs:annotation>
                <xs:documentation>Request for action to end after this time.</xs:documentation>
            </xs:annotation>
        </xs:element>
        <xs:element name="end_before" type="OsacbmTime" minOccurs="0" maxOccurs="1">
            <xs:annotation>
                <xs:documentation>Request for action to end before this time.</xs:documentation>
            </xs:annotation>
        </xs:element>
        <xs:element name="interval_eu" type="EngUnit" minOccurs="0" maxOccurs="1">
            <xs:annotation>
                <xs:documentation>Specifies the engineering units for the interval (hours, days, etc.).</xs:documentation>
            </xs:annotation>
        </xs:element>
        <xs:element name="work_order" type="MIMKey3" minOccurs="0" maxOccurs="1">
            <xs:annotation>
                <xs:documentation>Identifies the work task type.</xs:documentation>
```

6.1.2 Relevant aspects

The content of the OSA-CBM will clearly be of great importance in the development of an architecture for RCM in the rail industry. [102]

6.2 INSPIRE

6.2.1 Overview

INSPIRE is a European spatial data infrastructure focussed on the interoperability of network services such as transport, utilities and geographic features, with the objective of combining spatial data without specific computer or human action, to enable sharing and open access to data by organisations and the public.

INSPIRE was implemented via the publication of a number of ‘implementing rules’ including metadata, interoperability of spatial data, network services,

102 Recommendation 234

technologies, data and service sharing, and monitoring and reporting procedures. It is accessible via an open access 'geoportal'.

There are a number of ways in which the implementation of INSPIRE is of particular interest in the consideration of an RCM data architecture.

6.2.2 Relevant aspects

6.2.2.1 Principles

Due to the parallels between the objectives of INSPIRE and those of the Xi RCM initiative, the principles developed by the INSPIRE project should be of particular interest.

Table 7 - INSPIRE principles and their relevance to RCM data architecture

INSPIRE principle	Benefit	Application for RCM data architecture requirements
Capture data once and store it where it can be maintained most effectively	Efficiency and consistency	The data sharing architecture should allow data collected for one purpose or by one agency to be used for other purposes. A key requirement of the RCM data architecture
Provide access to data at all levels of granularity	Summarised or raw data available for different purposes and efficient transmission and storage.	Ensure that raw data is available to be shared as well as summary data.
Data readily available	Data is available when it is required by the process, of particular importance for real-time, operational requirements.	Ensure that the time-based requirements for data items are defined.

Table 7 - INSPIRE principles and their relevance to RCM data architecture

INSPIRE principle	Benefit	Application for RCM data architecture requirements
Data transparently available, including how data can be used	Transparency allows the user an understanding of the data they are using, for example: what is the raw data, how was it produced, how was it processed	Raw data and algorithms are available to users or processes.
Users can extend the schema and add their own objects to support an application; schema can be used for a number of different applications	A longer lifespan and reduced upgrades and maintenance. Greater value for money through wider utilisation. Reuse reduces development investment.	Ensure that the architecture is not tied to any supplier or process. Ensure it is designed in such a way that additional inputs, processing layers and outputs can be added. [a]
Choose minimum profile of standard to fulfil requirements	Alignment with framework is as easy as possible while maintaining the relevant standards.	Ensure that the specified parts of the framework meet the specific requirements of the agreed scope and no more.

a. Recommendation 038 and 232

6.2.2.2 Methodologies

6.2.2.2.1 Development of specification

Due to the similar nature of the deliverable of the INSPIRE and RCM data architecture initiatives, it is of particular interest to note the way that the INSPIRE developers approached their task, and their methods should be considered in the work to develop an RCM data architecture. [103]

Use cases were developed based on reference material, user survey, studies, previous initiatives etc and then consolidated into a set of user requirements

103 Recommendation 226

and object types. In particular, requirements were identified **on data content, data quality, timing, identifiers, relationships between objects and metadata for evaluation purposes. When capturing RCM data architecture requirements from consultees it must be ensured that all of these aspects of data are covered.** [104]

A first cut data specification was produced and along with an analysis of existing data specifications that could be reused, and a gap analysis, the full data specification was developed. Different parts of the specification had different starting points – in some cases there were already agreed data specifications, in some there was sufficient raw data but no agreed specification and in some the data was insufficient or missing. **When documenting the data requirements it must be noted what the current state of data is: – what data is existing, how it is specified, what data is missing and when will it be provided?** [105]

- Choice of tools and languages

The INSPIRE documentation contains a thorough report on the process of choosing suitable tools and languages for the framework, for example schema description and model mapping languages; to analyse strengths and weaknesses each tool was graded for expressiveness, implementation, web compatibility, technology independence and intuitiveness. The result of the INSPIRE analysis[106] was that 2 groups of schema description and model mapping languages were identified which provided the highest level of compatibility with requirements: UML/QVT and OWL2/RIF. RIF, which is compatible with OWL, is the candidate, not OWL itself. However, RIF was developed by the same standards organisation (W3C) as OWL and is a partner with it in the Semantic Web domain. RIF is able to interchange with RDF and OWL. **As part of the development of the RCM data architecture a similar process should be carried out when making decisions of tools and languages, and a review done of the choices and rationale made in the INSPIRE case.** [107]

104 Recommendation 226

105 Recommendation 227

106 http://inspire.jrc.ec.europa.eu/documents/Network_Services/JRC_INSPIRE-TransformService_SAA_v2.pdf p54

107 Recommendation 228

6.2.2.3 Rail definitions

Some of the schema and associated ontologies which were developed as part of the framework are of interest in the field of rail RCM. The *Common Transport Application Schema*, the *Rail Transport Application Schema* and the *Transport Network Data Product Specification* are all relevant. The Rail schema uses a link and node structure to represent rail tracks as a linear network, and also includes classes such as ownership and track section restrictions. The network is modelled based on three aspects: space, time and theme.

Figure 16 - Diagram showing elements of INSPIRE rail schema

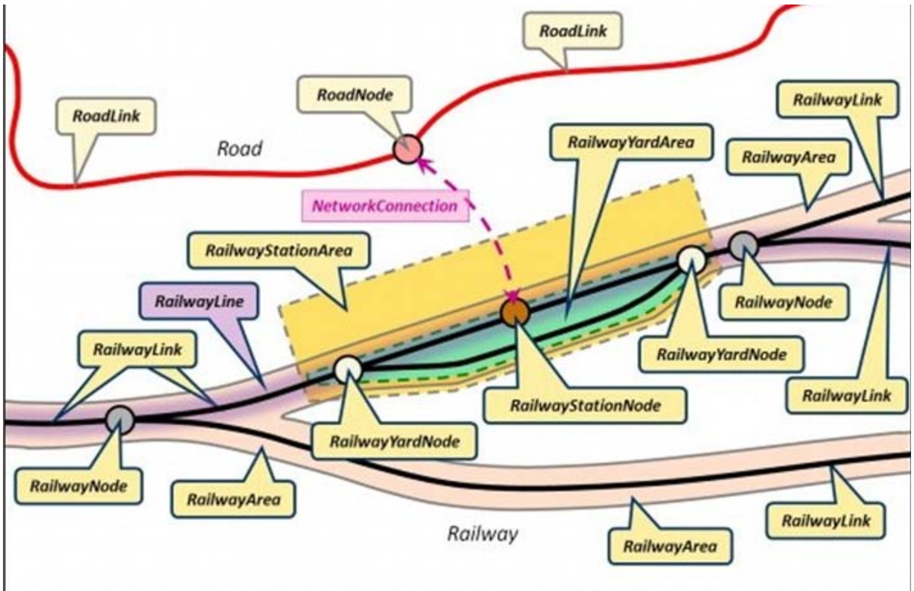


Illustration – Example of use of elements forming the Rail Transport Network

It is important that the development of definitions for an RCM data architecture takes account of the relevant definitions from these documents. This may save development time, prevent inconsistencies between schemas in use for RCM in the UK and European systems and leave an open path to future European interoperability. However, the rail definitions are currently fairly limited for the purposes of RCM, and consistency should also be required with other rail-related schemas. [108]

108 Recommendation 229

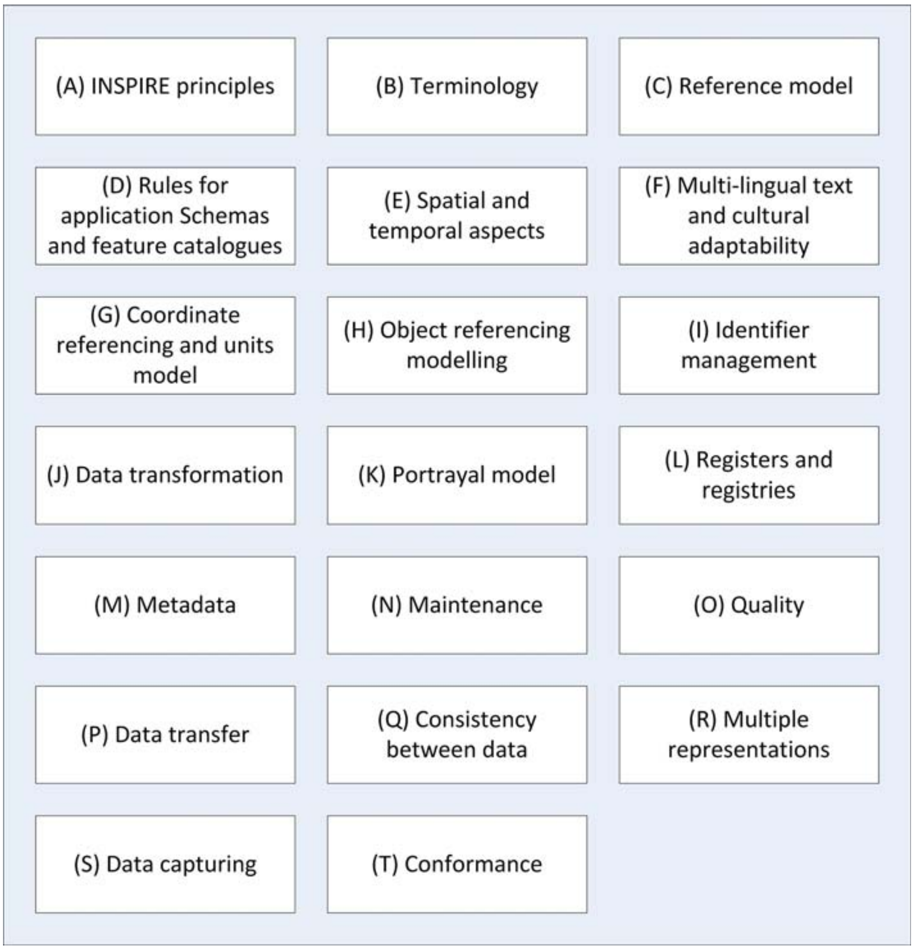
6.2.2.4 Interoperability definitions

The artefacts of INSPIRE include a generic conceptual model, defining the elements necessary for interoperability and data harmonisation including cross-theme issues, and incorporating a generic network model with specific mechanisms for interoperability of spatial data, such as object referencing, linear referencing, and linking network elements across different networks.

It will be beneficial for the designers of an RCM architecture to be familiar with the content of these models and to consider inclusion of all of the parts of the model listed in Figure 17 in the documentation architecture. [109]

109 Recommendation 230

Figure 17 - Elements of the INSPIRE generic conceptual model required to ensure interoperability



6.2.2.5 Spatial standards and specifications

The standard upon which the INSPIRE data specifications are based is ISO 19131: Geographic information - Data product specifications. The specifications themselves include the technical documentation of the application schema, the spatial object types with their properties, and other specifics of the spatial data themes using natural language as well as a formal conceptual schema language. They are intended to provide a basic framework with the focus on reuse and only including objects that will be widely referenced, allowing extension by users from the local level to wider geographic levels.

The data specification defines geometric objects in 3 ways:

- (topographic) area objects
- centreline objects
- point objects

Topology is handled in the data specification implicitly rather than explicitly, with the main reason to keep the model as simple as possible but expecting that most applications will use the network data within a topological environment. [110]

As identified by the review and consultation stage of the project, there is a need for geographic data to be available as part of the RCM architecture – for identifying the location of fixed or mobile assets. It will be important for the development of such schemas for RCM to draw on the work already done in this field by INSPIRE and potentially to make reuse of specifications where possible. [111]

6.3 Aeronautical Information Exchange Model

6.3.1 Overview

The Aeronautical Information Exchange Model (AIXM) is an information exchange model used in the aeronautical industry to enable machine-to-machine or service-to-service management and distribution of data. It is based on a data schema primarily concerned with airspace definitions and communications of those definitions, rather than on assets, however there are a number of ways in which it is comparable to the objectives that are being worked towards in the rail industry. These will be explored in the next section.

6.3.2 Relevant aspects and application to RCM

6.3.2.1 Broad community of users

The model addresses requirement from a number of different user communities from the same industry to access shared data from different sources. **It can therefore be seen that there are key similarities in the**

110 Reference INSPIRE Data Specification on *Transport Networks pVII*

111 Recommendation 231

context of this and of an RCM (or any sort of) architecture for the rail industry.

6.3.2.2 Value of a modular design

Modularity of design allows easy re-use of parts without having to deal with complexity of the whole standard. This means that supplier software and specific processes can easily integrate with the framework without a great overhead of time, cost or resource, which may deter integration. **The architecture design must address this aspect, and the requirements reflect consideration of how easily software suppliers, other initiatives or operational processes can integrate with an RCM architecture.** [112]

6.3.2.3 Extensibility

Third parties can expand the AIXM model adding features, properties, or domain values, for local applications. Concepts of local interest can be handled in a standard way without affecting the global interoperability. Place names in a local language are a typical example of local extension. **The architecture design must address this aspect, and the requirements reflect consideration of how easily software suppliers, other initiatives or operational processes can integrate with an RCM architecture.** [113]

6.3.2.4 Standards-based

Interoperability through being standards based –the model aligns with geographic standards from the ISO 19100 series, including spatial, temporal and metadata schemas, and the use of Geography Markup Language (GML 3.2 bundled with XLink) for the definition of geospatial concepts. **The design of the RCM data architecture must take account of the way standards are implemented to achieve interoperability. For RCM interoperability across the rail industry, with Europe and temporally (i.e. for future extension) are all areas where the application of relevant standards should be considered. The geo-spatial aspect is also of specific interest for RCM as integrating location information will be a key role of the architecture therefore the use of GML and the geographic standards by AIXM should be closely examined.**

112 Recommendation 206

113 Requirement 232

6.3.2.5 Asset unique identifiers

Universal Unique Identifiers (UUID) are used as artificial identifiers for features, representing not the feature itself but the data that represents that feature. **This is a concept which should be investigated for potential use in the schemas required for the RCM data architecture.** [114]

6.3.2.6 Importance of time

The concept of temporal (changes with time) or dynamic data, used in AIXM for temporal elements such as changes to air space, is of relevance in the field or condition monitoring as many aspects which will need to be represented by the architecture, for example, asset state, configuration and assignment can all change as can health thresholds. **Therefore an understanding of the way that AIXM models temporality would be of great use in the development of an RCM data architecture.**

114 Recommendation 233

7 Stakeholder consultation

7.1 Introduction

7.1.1 Purpose of the consultation

Stakeholders have been consulted initially to gain a more thorough understanding of the requirements of the RCM data architecture; to understand other overlapping initiatives such as AVI, LINX-TM and ORBIS; and to learn from the experience of previous and current RCM initiatives in data acquisition, processing and analysis.

Later in the project, we will consult further on the proposed architecture, with particular reference to its interfaces with other key initiatives and the issues that might arise for RCM data suppliers in conforming to it.

7.1.2 Method

The consultation involved a mixture of face-to-face discussions, telephone consultations and email exchanges.

Consultees were chosen initially based on suggestions by the project steering group, starting from a list shared with the team carrying out project T1010-02 on the commercial and legal framework for cross-industry RCM. The 2 detailed case studies identified for that project were also of interest for this project, since they exhibited important characteristics of most, if not all, cross-industry RCM initiatives.

Further consultees were identified during the earliest consultations as suitable people to provide specific information.

For each consultation, a record of the discussion was made by the project team and provided to consultees for agreement. The agreed text of each such record is contained in Appendix J.

Section 7.1.3 contains a summary table of all the consultations carried out and planned at the date of writing. The consultation exercise is still proceeding, so the following paragraphs are subject to change.

Section 7.2 contains a brief summary of the key points which emerged from the consultation and their impact on the RCM Data Architecture. The points are captured as findings and traced similarly to the findings from the background research exercise.

7.1.3 Summary of consultations undertaken

Table 8 contains a list of all the consultations carried out so far, scheduled or intended.

Table 8 - Summary of consultations undertaken or planned

Consultee(s)	Purpose of discussion	Date	Type of consultation
Michael Jacks, Virgin Trains; Mark Allen, Network Rail	UOMS – overview, commercial arrangements, data availability and processing	14/10/2013	Face-to-face, with T1010-02 team
Nicholas Kaye, Siemens	RailBAM – overview, commercial arrangements, data handling and availability	10/10/2013	Face-to-face, with T1010-01 team
Firdausi Irani, TTCI	TADS data structure and operation; InterRRIS principles and architecture	14/11/2013	Email exchange
Kevin Knott, Ryan Wickens, Network Rail	Intelligent Infrastructure – overview, use of MIMOSA, use of Wonderware, issues and lessons learned, future plans	12/02/2014	Face-to-Face
Tom Deacon, Network Rail	ORBIS location translation service	06/02/14	Email exchange

Table 8 - Summary of consultations undertaken or planned

Consultee(s)	Purpose of discussion	Date	Type of consultation
Jason D'Arcy, Network Rail	LINX-TM – architecture, messaging structure and message catalogue	13/02/2014	Phone conversation, email exchange
Dr John Easton, University of Birmingham	Web ontologies and their application to rail data integration. UB work with industry stakeholders in this area.	14/02/2014	Face-to-face
Dave Burbridge, Network Rail	Automatic vehicle identification – Network Rail's plans, standards in use.	18/02/2014	Face-to-face. Simon Tonks, Porterbrook also present
James Blaize-Smith, Network Rail	Intelligent infrastructure – plans for adding reporting and analysis functionality, use of higher MIMOSA stack levels	07/03/2014	Phone conversation, email exchange
Ashleigh Frost, Harinder Dhillon, Network Rail	UOMS – detailed method of operation, data formats, data processing, issues, business processes associated with it.	21/03/2014	Face-to-face
Kenneth Woods, Network Rail	UOMS on the East Coast Main Line	24/03/2014	Phone conversation; ongoing
Nicholas Kaye, Siemens; Mirek Vasely, Track IQ	RailBAM data formats and sharing; integration with a cross-industry architecture; Siemens on-train systems as providers of RCM data	12/05/2014	Face-to-face

Table 8 - Summary of consultations undertaken or planned

Consultee(s)	Purpose of discussion	Date	Type of consultatio n
Davin Crowley-Sweet, Network Rail	Network Rail data integration and data quality standards; ORBIS	27/05/ 2014	Face-to-face
Other ORBIS representatives, Network Rail	Data and message standards for fixed assets; location translation service; topological model of network	TBC	TBC
RSL representative, Network Rail	Rolling stock library; integration of this or successor with a messaging architecture for stock asset reference.	TBC	TBC
High Speed Measuring Train data processing team, Network Rail	Methods of merging different types of data about the infrastructure gathered on a train; location referencing	TBC	TBC

7.2 Findings of the consultation

These findings are grouped according to the type of insight they gave the project team rather than the people or organisations concerned.

7.2.1 Location referencing

The UOMS case study showed the amount of work that was generated by the absence of a standard way of representing GPS latitude and longitude co-ordinates. The data analysis team had to use an online lookup to convert the provided co-ordinates into a form usable to look up the location of a data point on a network map; then a further lookup using a home-grown asset list to identify the OLE assets the data point referred to.[115]

115 Recommendation 295a

The time and effort to do the manual lookups necessary were a significant limiter to the amount of data analysis that could be done by the team.

On the other hand, it is clear that relatively simple stipulations of a data architecture can greatly improve this situation: standard data formats and simple access to reference sources are the key.

From these observations we can identify the following data architecture requirements:

- Standard method of representing geographic locations[116]
- Easily available reference service to map geographical location to track location[117]
- Easily available reference service to identify fixed assets of a given type (Track / OLE / S&T) near a given track location or geographic location.[118]

7.2.2 Vehicle and train identification

Discussion with the AVI development team at Network Rail revealed a number of needs of a future national AVI system which would impact upon the RCM data architecture:

- Lookup mechanism to translate tag number (which may be the 12-digit EVN but may not) to a recognised vehicle number and orientation[119]
- Lookup mechanism to convert vehicle numbers into multiple units (for passenger trains) and their orientation and formation. This would have the capability of filling in missing data (for example for missing tags or missed readings) or correcting incorrect readings based on known asset configuration data [120]
- Lookup mechanism to associate identified vehicles / units with operational trains [121]
- Lookup mechanism to identify the individual key components (such as bearings, wheelsets) associated with each vehicle and longitudinal position [122]

116 Dependency D03A

117 Dependency D03B

118 Dependency D03C

119 Dependency D06A

120 Dependency D06B

121 Dependency D06C

- Need to reflect the different methods of AVI already present, associated with existing schemes such as TADS or RailBAM

The RCM data architecture itself will need to represent the data items involved in the requests and responses associated with vehicle number lookup methods.[123]

7.2.3 Commercial sensitivity, intellectual property and closed data

Discussions with RCM system suppliers and their users at Network Rail have revealed a tendency of RCM system suppliers to keep raw data confidential or to limit its accessibility for what they believe to be commercial reasons or to protect their intellectual property. It is clear that many benefits of RCM data sharing will only be realised if access is given to this raw data; and if competition can be brought to bear on the algorithms used to process it.

Data can be obscured by only making output data available and keeping the input data confidential; or by allowing the data to be viewed and extracted only using proprietary software or on payment of a license fee.

There is no question that suppliers do have intellectual property in their systems and in the algorithms they use to extract useful information from the incoming data. The ideal state of affairs is that suppliers have a way to claim and protect this intellectual property, whilst giving access to the data. It may take some effort on the part of the Cross-industry RCM Strategy Group to persuade suppliers that can be done.

From the point of view of the data architecture, the following principles will help support the goal.

- Formal adoption of the ISO 13374 layered model which separates out data from processing and identifies the different types of processing which are done on RCM data. [124]The intellectual property of suppliers will reside in the process elements, not the data [125]
- The use of open formats for data [126]

122 Dependency D06D

123 RCM Activity 086

124 Recommendation 037b

125 Recommendation 236

126 Recommendation 237

- The ability for data suppliers to identify themselves as owners of aspects of the data[127] and to specify alongside the data any restrictions or caveats on its use [128]

7.2.4 The Intelligent Infrastructure project

This project, having started as a pilot programme on a single line in Scotland, has now expanded in scope to have the goal of including nearly all switches and crossings, and signalling assets on the UK railway.

The concept behind the project is that raw RCM data are collected from sensors on fixed assets such as points and signals, assembled on site by local loggers, then transmitted via a variety of telecommunication methods but mostly by GPRS wireless back to a centralised server. At the server, the data are stored for historic reference and basic alert/alarm functionality is applied to generate warnings for maintenance engineers and train service controllers.

The project is of interest to the Cross-industry RCM initiative, although not itself a cross-industry project, for 2 main reasons:

- Its use of the MIMOSA OSA-CBM data architecture to provide a vendor-neutral way for different types of sensor provided by different manufacturers to be connected together using common data formats
- Its use of the Wonderware software and hardware platform for data collection, analysis and storage. This aspect of the project is described in Section 7.2.5.1.

MIMOSA OSA-CBM (See Section 6.1) is used in Intelligent Infrastructure to define a set of data interchange standards for sensor data moving between the data loggers in the field and the central computer. The field equipment is supplied by three different suppliers, but all are able to pass data to the central computer system. The data standards apply at the lowest-level (but highest volume) 'data acquisition' level only; the other functions of the system corresponding to the 'data manipulation' and 'state detection' levels are carried out by application code running in the central computer.

Presently the project has succeeded in getting large numbers of assets instrumented and gathering and storing data for them all. However, little has been done so far in tapping the potential of this data other than by the most basic generation of alerts and alarms for measurements falling outside pre-set

127 Recommendation 049

128 Recommendation 049, 085 and 085a

bounds. An extension project is under way to provide a data warehouse solution to enable more sophisticated time-based reporting and analysis.

Observations from the project relevant to the RCM data architecture are:

- It is important to include within the data standards ways to handle configuration and control information for field-based equipment. These aspects had to be developed after the initial deployment of Intelligent Infrastructure and are still a source of much support and maintenance work for the system [129]
- Although the strategy of capturing all sensor data centrally gives the best potential for preserving the data for analysis later, it is an expensive option. It may on reflection be more cost-effective to have some data processing done in the field and to pass aggregated data back to the central server. The data architecture should be able to handle both these cases and to allow the 'level' of data to be altered over time [130]
- The project has been conceived as a single application stack rather than as a service-based architecture; and it is managed and supported by a single organisation on Network Rail's behalf. This may limit its potential to be developed using similar standards to the RCM data architecture. The RCM data architecture should allow multi-party participation [131]

7.2.5 Hardware platforms for RCM data interchange

An important consideration for the RCM data architecture is how it could be implemented: what hardware/software platform would be necessary to achieve all the goals set out for the cross-industry RCM data strategy.

There are aspects of the data architecture which will generate value independently of the hardware used: the requirements to conform to standard data formats and datagram structures, for example, will significantly improve the shareability of data irrespective of how that sharing takes place; however, other aspects of the architecture will need a software infrastructure to implement them on.

There is obvious reluctance in the rail industry to build a new set of infrastructure specifically for RCM data sharing. Options for co-opting existing software platforms to do the job should therefore be considered.

129 RCM Activity 238 and Recommendation 262

130 Recommendation 083, 224 and 257

131 Recommendation 078

In the consultations so far undertaken, two candidate platforms have been encountered: the Wonderware platform used by the Intelligent Infrastructure project and the message-based architecture of LINX-TM.

7.2.5.1 Wonderware

Wonderware is a software platform for the collection, processing and archiving of sensor-based data. It is typically used for SCADA applications, so comes with a wide range of tools for managing connected sensors and data loggers. It is used as the software platform for the Intelligent Infrastructure project described in Section 7.2.4.

The Wonderware implementation used for Intelligent Infrastructure is a centralised server-based one, with a single large data store ('Historian') currently occupying about 20Tb of storage and 100 server instances divided into route-based groups managing data capture from 15,000 data loggers in the field, with 30,000 connected assets.

The Historian data store keeps raw data logged from sensors for up to 2 years at full detail, with summarised data able to be kept for longer. This aspect of the design, though expensive in bandwidth and data storage costs, does mean that all data are available for investigation and analysis and so opportunities for productive new uses are not lost.

Wonderware has the capability for data processing modules to be added. Under the present arrangement for the Intelligent Infrastructure project, these are created and supported by Thales who manage the system for Network Rail. There does not appear to be a way for third-party data processing modules to be integrated. All applications using the data need to be built in this way.

Whilst it is clear that Wonderware has the capability to manage and store large quantities of low-level sensor-based data, it is not yet clear how easy it would be to extend it to add in support for message-based interactions, the use of web services for reference lookup, or the creation of an open services-based architecture for data sharing.

Consultation is continuing on the possible future use of the Intelligent Infrastructure Wonderware platform to hold data about trains captured using trackside sensors such as RailBAM.

7.2.5.2 The LINX-TM enterprise service bus

LINX-TM has been conceived from the outset as a real-time message-based architecture whose prime function is to allow the open interchange of information about the control of moving trains. A catalogue of services has been defined, which at this stage relate to timetabled trains and their rolling stock.

The architecture of LINX-TM corresponds closely with the service-based model considered most suitable for the open interchange of RCM data in the terms required. However, there may be concerns about its suitability for handling the large volumes of data generated by some RCM activities, particularly if the presence of such data might threaten LINX-TM's performance in time-critical areas such as real-time train control.

7.2.5.3 Proprietary RCM data warehouses

Several vendors have produced data warehouse solutions to consolidate asset condition information gathered from rail vehicles and trackside detectors. Typically these systems have started life as home-grown systems intended to capture data from the supplier's own train management systems or trackside detectors; they have then been extended to include data from other types of detector, to interface with asset management systems, to provide rule-based analytic functions and to offer web user interfaces.

Examples of this type of system are listed below. Recommendations associated with them are to ensure that the data architecture supports data interchange with them:

- Bombardier's ORBITA (http://www.bombardier.com/en/transportation/products-services/services/innovation-technology.html)[132]
- Siemens' Remote Rail Service Desk [133](http://www.bahnindustrie.info/uploads/media/Vortrag_1_Emmelheinz.pdf)
- TTCI's InteRRIS (http://www.aar.com/pdfs/VehicleMonitoring_OnePager.pdf), [134]
- TrackIQ's FleetONE (http://www.trackiq.com.au/fleet-one.html)[135]

These systems all store their data in a standard relational database system such as Oracle™ or SQL Server™, all have a data loading process which can be configured to take data from different sources, and all have a web-based user

132 Recommendations 285 and 288

133 Recommendations 273, 274 and 279

134 Recommendations 249 - 253

135 Recommendation 275

interface which enables end users to browse data, create reports and set up alerting conditions.

The basic model of these systems is a 'client-server' one which involves the consolidation of data into a single server, with multiple clients connecting to it over a network. Each tends to specialise in the type of data for which it was designed – rolling stock or infrastructure.

In considering whether systems of this type form a suitable basis for hosting a cross-industry RCM data sharing architecture, we should recognise the following plus points:

- They are established systems specifically architected to handle RCM data. They already have the facilities to build applications involving
- The use of standard SQL databases at the heart and the presence of data processing engines means that extending the systems to get data in our out in standard ways should be straightforward.

The following potential drawbacks should also be noted:

- These systems have not been built around open standards or agreed data models
- No single system will cover all the types of asset or types of enquiry likely to be needed
- The client-server model with centralised database will not support the higher levels of process integration mentioned in ISO 13374.
- It may be difficult to persuade the system vendors to open up the systems to the degree necessary, since their current business model is based around paid licenses to use the systems.

7.2.6 Web ontologies

The technology of web ontologies (described in Section 3.3) has for several years now been promoted by academics in that field (see Section 5.3) and has been used in the successful InteGRail trial (Section 3.2). There is much interest in how this technology can help address some of the constraints impeding effective cross-industry information sharing.

There are developments and prototypes being tried out; the University of Birmingham is engaged in some work with Network Rail and others to set up a 'core rail ontology' which would define a structure for the main types of railway data. However, it remains true that effective tools which can place the building and maintenance of ontologies into the hands of rail industry professionals rather than academics are still some way off; and the technologies may face challenges from the scale of the railway network and the volume of information that RCM activities generate.

From the point of view of the cross-industry RCM data architecture, the consultation on web ontologies has led to some guidance on ensuring that the architecture is not inimical to the use of ontological approaches where they may later add value. Likely areas where value will be added will be in reasoning and deduction from presented evidence to expected outcomes – for example:

- State detection for a composite asset based on alerts, alarms and states of the component assets. An ontological approach may make it easier to capture and code the structures and interrelationships between the component assets and the whole composite asset
- Health assessment likewise: understanding the way in which, for example, failure of a set of points in conjunction with a nearby broken rail will impact upon the capacity and capability of a section of route

To facilitate an ontological approach, the data architecture should:

- Use the approaches and road map set out in the ISO 15926 standard for data integration (described in the review of standards in Section 4.1.2.1) to ensure that vocabularies are shared and there is a consistent approach to versioning and validity of data. The key elements for our immediate purposes are a shared conceptual data model and a common set of reference data [136]
- Use the methods and concepts set out in the Vocabulary of Interlinked Datasets (VOID) to identify where data of a given type are stored and in what format[137]
- Use standard basic ontologies to express shared concepts such as responsibility, ownership, copyright, rights and permissions. The Dublin Core ontology, for example, covers this area in a widely-understood way.[138]
- Use common standards for commonly-shared data items. These can be global standards, such as ISO 8601[139] for dates and times or ISO 17000[140] for geographical locations; or rail industry standards such as the data types

136 Recommendations 037d and 266

137 Recommendation 239 and 264

138 Recommendation 116b

139 Recommendation 37j

140 Recommendation 37k

defined in railway group standards or the rolling stock data integration standard RIS-2706-RST. [141]

These stipulations make excellent sense in their own right for standard data integration reasons, so there need be little concern that the data architecture would be gold-plated to suit a technology that may never be required.

7.2.7 Algorithmic sophistication

With the possible exception of the acoustic bearing monitoring systems TADS and RailBAM, there is little evidence in the consultation done so far of the use of sophisticated data analysis techniques to extract valuable information from raw RCM data. UOMS, for example, uses a very simple acceleration threshold to identify events of interest; and the job of distinguishing actual events of interest from spurious ones generated by electrical noise is a manual one done by visual inspection of a recorded voltage vs time trace. The threshold used is the same for all locations and vehicle speeds.

Similarly for the Intelligent Infrastructure project, alerts and alarms are generated based on straightforward rules for time, current or pressure exceeding standard thresholds. The thresholds are set by class of asset rather than by individual asset; and there is no systematic way to calibrate these based on historic expected readings.

In all cases, the logic associated with the rules and thresholds is not made explicit in the output data and tends to be kept well hidden by the owners of the system or the programmers of the rules.

The reason for this lack of sophistication is mostly to be found in the difficulty of taking a time-based view for a given asset because of the way in which data are stored or made inaccessible. For example, in the West Coast UOMS, there is no database into which earlier readings can be stored and called back for comparison; and therefore each event is treated as if it were brand new.

Even for the acoustic bearing monitoring systems, the details of the data processing approaches used to get useful readings from raw acoustic data are deemed commercially valuable and so not made public. This means that independent assessment of the data quality is possible; and there is no scope for benchmarking against other algorithms and a harder task of continuous improvement through fine-tuning of the algorithms.

141 Recommendation 37h

There are several ways in which the Data Architecture can help to remove some of the impediments to improving the quality of data analysis:

- Enable data elements to be separated from processing elements of any RCM project, such as by the use of the ISO 13374 layered stack. This will enable best-in-class algorithms to be brought to bear on the data and to be improved independently of the data itself [142]
- Support the storage of time-based data about the same asset or segment by requiring standard nomenclature for assets and segments, standard location referencing techniques and standard methods of referencing dates and times [143]
- Support the encapsulation of processing algorithms so that the intellectual property they contain can be protected [144]
- Provide a way for algorithms and processing steps to indicate their performance in standard terms such as precision, accuracy, tolerance, signal to noise ratio, likelihood of false-positive and false-negative errors so that data consumers know what to expect from the data [145]

142 Recommendation 240

143 Recommendation 241

144 Recommendation 240

145 RCM Activity 057

8 RCM systems and data

We have surveyed a number of rail industry RCM systems to consider the types of RCM data they use or generate and to assign them to a category. The categories are used in the actors and use cases analysis in Section 9.6.1.2 as a list of generic system actors.

The RCM data involved in each system are categorised in the following ways:

- What formats of data are involved
- Whether the data are from moving or fixed detectors
- What asset classes are being monitored
- What levels of the ISO 13374 RCM data processing stack are covered.

Other rail industry systems are considered as well, particularly if they play a role in supplying important contextual information for RCM data users or data processors.

The results of the survey are shown in Appendix F. The results of this analysis will be used to inform thinking in latter stages of the project about the prioritisation and order of delivery of elements of the RCM data architecture.

9 Outputs of the review

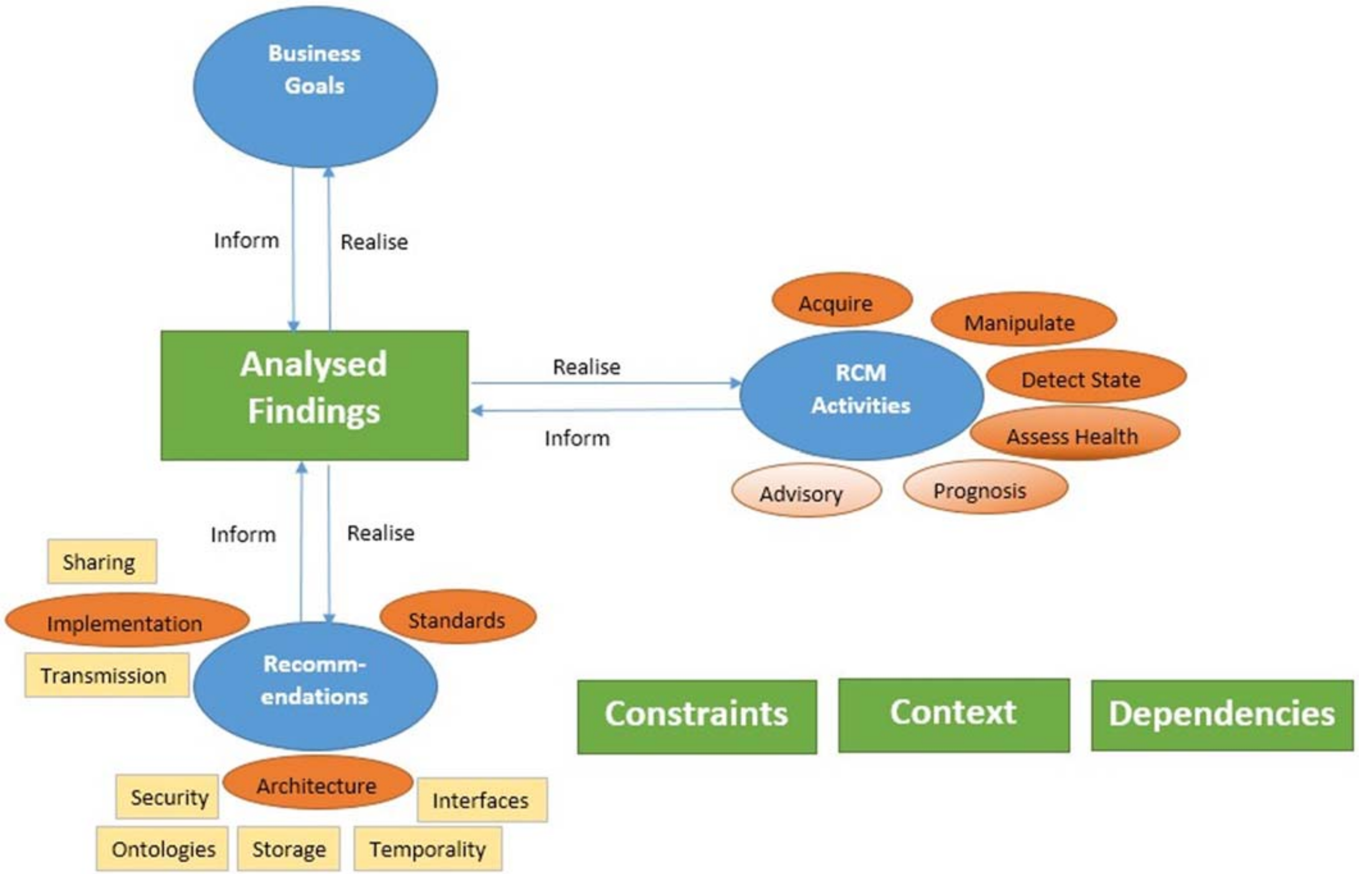
This section describes the process by which all the initial findings identified in the sections above have been analysed to generate a set of outputs which will pass forward to the next stage of the work to define the Architecture Principles and the Architecture Requirements.

The influencing factors acquired from the research were consolidated, de-duplicated and categorised into a number of groups, namely business goals, RCM activities, recommendations, constraints, context and dependencies. These consolidated findings all cross refer to one or more of the influencing factors. In this way every consolidated finding can be traced, via the raw requirement, back to a source.

These findings were then able to be used as the basis for guiding and validating the architecture requirements through the formulation a set of use cases and architectural principles, each linked to one or more of the findings, either directly or indirectly, again allowing traceability back through the process. Each of these is discussed in the following sections and then a summary of the findings is provided.

Figure 18 shows the composition and relationships between these sets of findings.

Figure 18 - Findings of research



Based on the findings, a set of use cases was defined which will provide an additional means of verifying that the architecture meets the requirements of its users.

9.1 Management of findings

The findings discussed in this section are held in the data modelling tool Enterprise Architect, as elements of various packages. Appendix D on Enterprise Architect contains information on this and on how the packages operate and enable traceability. In this section we indicate which Enterprise Architect package holds each type of finding.

9.2 Business goals

These range from the very high-level, industry-wide objectives defined in the McNulty Report or the Rail Technical Strategy (for example 'The transformed railway is available seven days a week and is reliable, resilient, safe and sustainable.') to those concerned with best practice of asset management or with availability of decision support data.

Business goals are found in the Enterprise Architect package 'Consolidated Requirements' with the stereotype 'Business Goal' assigned.

The business goals inform the background to the requirement for an RCM data architecture, providing the necessary context that grounds all of the project's requirements so that the development of the architecture supports the intended direction of the industry as a whole. For this reason, it will be ensured that every requirement taken forward to inform the architectural design can be linked to a business goal.

Business goals at the highest level include the following themes:

- improved reliability
- improved safety
- improved cost performance

At a more detailed level they include the ways that have been envisaged to achieve those highest level goals:

- improved systems
  - integration and implementation
  - costs
  - reliability
- improved knowledge
  - knowledge generation

  - support of operational and strategic decision making
  - data used effectively and targetted
- improved asset management practices
  - intelligent maintenance techniques
  - lifecycle costing tools
- reduction of failure and unplanned maintenance

For all these selected goals, a requirements statement has been formulated, stating that '*The architecture must enable……*' followed by the detail of the goal.

9.3 RCM activities

These include any information acquired on RCM activities, current or planned, and the actions, processes, systems, technologies involved in these activities

The findings in this group are all titled 'The architecture must…..' and so have the status of requirements: these are the activities which will or may be integrated through an RCM data architecture.

As a result these feed directly into use cases, so that every RCM activity must be linked directly, or indirectly via another RCM activity, to one or more use cases.

They cover requirements that the architecture must support:

- Availability of data for a set of current RCM systems
- Availability of supporting information to assist in the RCM process
  - vehicle, asset or location can be identified
  - internal and external information sources can be used
  - operational facts can be proved
  - diagnosis and allocation of cause can be carried out
- Extensibility so that data for other RCM systems can be available in the same way
- The 6 processing stages of RCM as defined in ISO13374 (see Section 4.1.1.1)
  - raw data can be acquired
  - manipulated data can be acquired
  - state detection can be carried out

- health assessments can be carried out
- prognostic assessments can be carried out
- advisory generation can be carried out
- As well as associated functional and non-functional requirements about how the data is available and how it can be used
  - general data quality requirements
  - data is available in real time

The findings in this group are also labelled according to which stage of the six level processing model they are involved in. See Section 6.1.1.1 for further detail on this processing model.

RCM activities are found in the Enterprise Architect package 'Consolidated Requirements' with the stereotype 'RCM Activity' assigned.

9.4 Recommendations

These findings cover a wide spectrum of sources, from recommendations made by the authors of reports to those made during the consultation period. They can be categorized as:

- Those concerned with implementation
  - testing and acceptance procedures
- Those concerned with the architecture itself
  - existing systems with which the architecture should be integrated – ORBIS, LINX, DRACAS
  - alignment with other sets of specifications such as MIMOSA and railML
  - recommended platform – service oriented architecture in line with the InteGRail platform,
  - data sharing level recommendations
  - inclusion and reuse of certain elements such as ontologies, railway functional architecture views, and definitions of security and integrity levels from other areas,
  - extensibility and modularity
  - use of universal unique identifiers
- Those concerned with standards
  - ISO standards relating to asset management data, condition monitoring and lifecycle data management (see Section 4.1)

    - European rail standards such as technical specifications for interoperability TAP and TAF (see Section 2.5.1.1)
    - railway industry standards such as the rolling stock data standard RIS
- Those concerned with the management of a project of this type
    - formal processes for selecting tools and languages used in architecture
    - checklists for requirements for this sort of project
    - methodologies for requirements gathering and consultations
    - methodologies for creation of schemas

Recommendations are found in the Enterprise Architect package 'Consolidated Requirements' with the stereotype 'Recommendation' assigned.

9.5 Dependencies

This includes any activities or technologies which the realisation of the full benefit of an integrated RCM data architecture is or may be dependent on. This is an important set of findings to document as their role as a dependency for RCM may support their own business cases. Where specific technologies are suggested in the literature, they are translated into dependencies which state the required outcome, not enabler (the 'what' not the 'how'), For example, rather than stating that a particular AVI technology is required we can state that a way of automatically identifying vehicles is required.

The dependencies fall into 3 main groups: improved analysis, improved information and improved identification.

In Enterprise Architect, the package 'Raw Dependencies' is used for the findings relating to dependencies as they come directly from the research. The 'Dependencies' package is used for the consolidated, de-duplicated and more generic, shorter set of dependencies.

9.5.1 Improved analysis

T844, the report mapping the current RCM activities, identified that the reduction in delays expected from wider use of RCM (generally 30 %, with 40 % for points and 75 % for signalling power supplies) would only be 10-20 % if not accompanied by more sophisticated analysis through the development of improved algorithms. These should be consistent visible and accessible, in other words not 'black box processes'.

The creation of suitable algorithms and the analysis they support is therefore central to the higher levels of RCM processing identified in ISO 13374. It is these more complex computations that will allow predictive activities to take place and thus realise the full potential of RCM.

Increased analytical capabilities should be recorded as an important dependency for the full realisation of the benefits of an RCM data architecture. [146]

9.5.2 Improved information

Of equal importance in the ability to carry out predictive activities is the libraries of information upon which algorithms will be developed, for example, asset condition, asset configuration and deterioration profiles. These libraries should conform to the agreed standards implemented by RCM and other relevant data initiatives, providing consistency of definitions, formats and interfaces. The requirements for analysis and information to support RCM activity are in fact more than dependencies. In fact they, along with the architecture, are key components of a predictive and advisory RCM system. Any data or interface requirements between the architecture and any existing or future data repositories will be defined in the architecture documentation.

Improved asset information base should be recorded as an important dependency for the full realisation of the benefits of RCM data architecture. [147]

9.5.3 Improved identification

In addition to the information requirements above, there are two important information gaps at the lower level of the processing model, the automated matching of data to the asset from to which it relates. This can be done in two ways – identifying the location of a mobile asset when monitoring a fixed asset, and identifying a particular mobile asset as it passes a fixed monitor.

9.5.3.1 Location

The literature reviewed reveals a key dependency on improved location identification to provide the precise location of a train when it detects, for example, a fault with track. T853, Mapping the RCM Architecture, notes the requirement for an ‘unambiguous definition of the infrastructure assets a train is interacting with’[148] and the current absence of this function ‘is

146 Dependency D02

147 Dependency D01

preventing systems where trains are monitoring infrastructure from being exploited fully. [149] It states delivery of 'full implementation of AVI / GPS / Train Describers to allow this. In addition, the railways systems workshop in December 2012 found that the provision of GPS data was of most interest in the field of RCM to provide 'fine-grained positioning of vehicles'. [150]

Accurate location information is an important part of the vision, and enabler of, the Rail Technical Strategy 2012, in fact, this is required for a number of industry initiatives and is being addressed by the Train Location Service project. This aims to improve on the accuracy of a known train location on board and at a base station, by combining data from GPS, odometers, signalling systems and trackside locators and repeatedly iterating it.

To conclude, it is apparent that accurate location information is a key dependency of an effective RCM data architecture. The degree of precision and accuracy are to be documented as part of the data requirements. The dependency does not need to specify what the source of this data ought to be, however, given the moves in the industry towards increasing the use of GPS, and the potential data streams involved in the train location service, it will be within our remit to advise on the suitability of this for the purposes required and to specify the architecture accordingly.

9.5.3.2 Vehicle Identification

This is the corresponding dependency for the monitoring of mobile assets and concerns the identification of a train or vehicle as it passes over a fixed monitor. If a fault is detected, this information is of increasing value the more accurately the location of the fault can be assigned. We can state that accurate identification of assets is a key dependency of an effective RCM data architecture. T853 and T844 both find it to be an 'essential enabler' for train monitoring[151].

More specifically, Automatic Vehicle Identification (AVI) is described by T844 as an 'essential enabling facility[152]' in enabling remote monitoring of mobile assets beyond the 'alarm' stage, given that a consistent and quick vehicle identification system is a current gap.[153] Assets can be currently identified by

148 T853 Mapping the Remote Condition Architecture p33

149 T853 Mapping the Remote Condition Architecture p31

150 T853 Mapping the Remote Condition Architecture p11

151 T853 Mapping the Remote Condition Architecture p26

152 T853 Mapping the Remote Condition Architecture p43

combining a number of other data sources however this can is neither quick nor automatic. 'Beyond the alarm' stage means recording normal performance and long term trends needed to build up a repository of asset condition data. Additionally, AVI was selected by the groups working in the information systems workshop in December 2012 as a key actor or enabler in many systems and a current gap.

9.5.4 Constraints and context

Context includes things that provide a useful background to the task and as such may well contribute to the direction of the solution. Typically these may be about the motivations behind the initiative, problems to be solved, expected benefits and details about current activity and thinking in the field.

Constraints include some of the 'problems to be solved,' other things that would constrain the solution, such as storage and transmission capabilities, or implementing in the most cost effective method.

Constraints and context are found in the Enterprise Architect package 'Constraint & Context'

9.6 Use cases

Further to the analysis of the research findings, the sets of business goals, RCM activities and recommendations were used to formulate a list of use cases and associated actors. The aim of this exercise is to consider the business processes the RCM data architecture would need to support to help identify the areas where the RCM data architecture would add most value.

Actors represent the roles of people or systems who carry out actions in pursuit of a business goal. By analysing the roles rather than actual IT systems, actual job descriptions or industry positions, we can identify the core of the business process they participate in which would survive any reorganisation of the rail industry or part of it.

Use cases represent the tasks and operations associated with RCM data. They are similarly defined in general terms, irrespective of the types of data or the types of asset being monitored.

153 T853 Mapping the Remote Condition Architecture p31, T844 Mapping Current RCM Activities to Systems Reliability Framework p38

The aim of this approach is to help identify the roles the RCM data architecture will need to fulfil in order to assist current RCM activity and enable extension to new activities that support the same business processes.

9.6.1 Actors

By considering the requirements expressed as business goals, the types of user known to be associated with current RCM activities, and probable future data and system roles, we arrive at a candidate list of actors who will interact with RCM data and so take part in the use cases the architecture it will support.

9.6.1.1 Human actors

The human actors identified are listed in Appendix E. The mapping between these and the business goals identified in the requirements is defined in the Enterprise Architect repository and shown in Appendix E Section 11.

The human actors fall into the categories of:

- infrastructure manager
- IT
- rolling stock maintenance
- rolling stock owner
- railway undertakings

Within these categories there are a variety of different actors who have a range of responsibilities from managing data, asset management, maintenance planning, safety, and the provision of services and equipment. A table showing the full list of actors and their responsibilities is found in Appendix E.

9.6.1.2 System actors

Categories or roles of IT system involved with RCM data that supports business goals were identified. The system actors were grouped into categories including:

- railway system
- fixed RCM system
- on board RCM system

Within these groups there were a large number of different system actors, representing a wide range of functions in all aspects of railway activities, such as storing and processing data, identifying assets, enabling communication of data, and managing systems such as timetabling, stock allocation or work planning. In addition the RCM systems have a set of responsibilities specifically for monitoring certain aspects of asset, and processing and transmitting the resulting data.

The full list of system actor categories, actors and their functions are listed in Appendix E Section 2. A representative list of actual UK rail IT systems and their categorisation is described in Appendix F.

9.6.2 Use cases

By considering the goals of the actors and their roles in using RCM data and systems to improve the operation of the railway, the following use cases were derived which the RCM architecture will need to support.

The mapping between actors and use cases i.e. which actors are involved in which use cases is shown in the use case diagrams in Appendix E Section 12.

The use cases will be used during the architecture requirements definition phase of the project to verify that the requirements users needs.

We have categorised the use cases to reflect the type of user goal they support. The categories used are shown in Appendix E.

Table 9 - Use case categories

Use case category	Description
Content	Use cases relating to the different types of data content that actors might require or supply. The data content types are defined in terms of the ISO 13374 stack described in Section 4.1.1.1
Delivery	Use cases relating to the different ways actors may provide or receive RCM data
LifeCycle	Use cases relating to the life cycle of individual assets: addition, movement, disposal
Process	Use cases relating to the gathering and processing of RCM data. We have used the layers of the ISO 13374 stack described in section 4.1.1.1 to distinguish the different types of data processing and usage that take place.

Table 9 - Use case categories

Use case category	Description
Reference	Use cases relating to the lookup of valid descriptive data about the assets and the railway
Responsibility	Use cases relating to ownership, responsibility and commitments to supply and use the data as agreed
Security	Use cases relating to the need to maintain security of railway IT systems and the integrity of the RCM data handled by the architecture
System	Use cases relating to the management of the RCM data architecture and the connection of external systems to it

9.6.2.1 Content use cases

Content use cases cover the definition of the data content of interchanges that the architecture needs to be mediated.

The use cases are defined in terms of the data content implied by the layers of the stack defined in ISO13374, described in Section 4.1.1.1, rather than by their specific RCM function. This helps ensure that the use cases will be valid for any type of RCM data irrespective of the type of system actually used to gather or process it. In these use cases, the actual method of providing or getting access to the data is not considered – the Delivery use cases described below cover that aspect.

The full list of content use cases is available in Appendix E Section 3 and includes cases such as getting sensor data, getting data from other systems, and setting configuration for data sensors and parameters for their data manipulation functions.

9.6.2.2 Delivery use cases

These use cases cover different ways in which RCM data can be accessed or transferred. They reflect the different ways in which railway systems could make data available and the different ways in which users may wish to query the data. In these use cases, the actual content of the RCM data is not considered – the content use cases cover that aspect.

The full list of delivery use cases is available in Appendix E Section 4 and includes cases such as getting data by format, by parameter such as location, time or event, from fixed or moving sensors, and the delivery mechanism – is it scheduled, on demand, event-driven.

9.6.2.3 Asset lifecycle use cases

These use cases cover the changes that occur to assets on the railway. They are expressed in general terms so that they can apply equally well to infrastructure assets such as track, switches or overhead line equipment; or to moving assets such as rail vehicles, units and trains.

There are 2 key concepts which underpin these use cases:

- Assets and segments. This terminology is borrowed from MIMOSA (see Section 6.1). A segment refers to a designed position on the track or on a rail vehicle; an asset refers to the specific item of equipment or plant that occupies that position at a given time. For example, the front wheelset on the front bogie of a given rail vehicle would be considered a segment; the actual numbered wheelset installed there would be considered an asset.
- Unique identifiers. Segments and assets can be referred to by different types of identifier by different railway systems; and some systems do not make a distinction between the segment and the asset occupying it. In order to maintain continuity and handle all the cases that can occur of reconfiguration of the network and the introduction, moving and removal of pieces of equipment, it is important to have a single unique identifier for each segment and each asset. The actual way in which the unique identifiers are assigned and managed is to be determined, but it may be necessary to set up a specific register of unique IDs to satisfy the needs of the RCM data architecture.

The full list of asset lifecycle use cases is available in Appendix E Section 5 and includes cases as described above.

9.6.2.4 Process use cases

ISO 13374 describes RCM data processing activities in a stack, moving from the lowest level of the collection of raw sensor data, through conditioning and smoothing of that data, to the comparison of the data against expected norms and alert and alarm thresholds. Higher levels of the stack consider the assessment of asset health, the forecasting of future performance and health, and the generation of maintenance responses to the detected expected asset deterioration.

This way of thinking about RCM data processing is independent of the exact type of RCM data being considered, so is appropriate for the consideration of the data processing steps a general-purpose architecture needs to include.

Important use cases in this set are those that refer to calculating the impact on a group of assets of the state of one of its constituent assets. In a tightly-connected area such as the rail network, it is important to be able to take a clear view of the impact of, for example, a fault in a point motor on the capacity and capability of a whole section of the network.

The full list of process use cases is available in Appendix E Section 6 and includes cases such as generating health assessments and advisories and generating deterioration profiles and prognostic parameters.

9.6.2.5 Reference use cases

The use cases in this group cover the important issue of cross-referencing and linking. They include the key cross-industry issues of converting between different ways of referring to the same location; tying together the physical rolling stock with the operational train; and identifying the components such as wheelsets comprising a composite asset such as a rail vehicle.

These use cases imply lookup activities taking place against standard reference sources such as a track model or an asset configuration management system.

The full list of reference use cases is available in Appendix E Section 7 and includes such cases as identifying assets or vehicles passing sensors, getting location of moving sensor data and accessing standard reference data.

9.6.2.6 Responsibility use cases

Providers of RCM data will need to be able to know what is happening to it; consumers of data need to know who owns it and what restrictions apply to its use.

The data quality characteristics of the data are important: users may need to request data of a known good quality; and will need to know the characteristics of the data they are supplied with.

The full list of responsibility use cases is available in Appendix E Section 8 and includes such cases as identifying the owner or responsibility holder for an item of RCM data or a process, and identifying the data quality requirements that these actors have.

9.6.2.7 Security use cases

These use cases cover the security of the rail systems connected to the architecture and the architecture itself. They also cover the integrity of data being transferred using the architecture.

The full list of security use cases is available in Appendix E Section 9 and includes use cases applying to the security of the architecture, the data (which may be privately or publicly viewable) and the connected railway systems.

9.6.2.8 System use cases

These use cases consider the management of the architecture itself and the support of projects wishing to connect to it to transfer RCM data.

The full list of system use cases is available in Appendix E Section 10 and includes use cases to the integration of various systems with the architecture and with the facilitation of access to historical data.

10 Appendices

Appendix A Train Location Strategy Data

In separate document “Appendix A Train location strategy”

Appendix B RCM activities identified in T844

In separate document “Appendix B RCM Activities Identified in T844”

Appendix C Project Workshop Findings

In separate document “Appendix C Workshop Findings”

Appendix D Enterprise Architect

In separate document “Appendix D Enterprise Architect”

Appendix E Actors and Use Cases

In separate document “Appendix E Use Cases”

Appendix F Systems and Data

In separate document “Appendix F Systems and Data”

Appendix G Suggested Applications for a Railway Data Framework

In separate document “Appendix G Suggested Applications for a Railway Data Framework”

Appendix H Example Vocabulary Lists from Specification for a System Wide Data Framework

In separate document “Appendix H Example Vocabulary Lists”

Appendix J Record of Consultations

In separate document “Appendix J Record of Consultation”

RSSB R&D Programme
Block 2 Angel Square
1 Torrens Street
London
EC1V 1NY

enquirydesk@rssb.co.uk

http://www.rssb.co.uk/research-development-innovation/research-and-development