IMPRCM Documentation

Release 0.4+52 (6c8bb53)

Vivacity Rail Consulting for RSSB: PJ, GM

Jun 18, 2019

Contents:

Welcome to the Rail Cross Industry Remote Condition Monitoring (Cross-industry RCM) resource: a guide to using the toolkit developed in RSSB's research project *T1010*. Here you will find guidance, reference material, tips, sample template documents and best-practice notes gained from applying this toolkit to an actual Cross-industry RCM scheme (*The IMPRCM Project* (page 6)).

Part I

Site Contents

CHAPTER 1

Introduction to Cross-industry RCM

This page is an introduction to Cross-Industry Remote Condition Monitoring (*XIRCM*) for general readers and people new to the field. It covers:

- *What is Cross-Industry RCM?* (page 3) - general definition of Cross-industry RCM and why it has particular problems which have limited its take-up.
- *Some Examples of Current Cross-industry RCM Schemes* (page 4) - notes on some examples of Cross-industry RCM.
- *Industry Initiatives* (page 5) - where Cross-industry RCM fits in to work being done by industry to meet future challenges.
- *The T1010 Project* (page 5) - the guidance and standards toolkit developed by the Cross-Industry RCM Strategy Group (*XIRCMSG*) to tackle impediments to Cross-industry RCM take-up.
- *Key Principles of the T1010 Advice* (page 6) - an introduction to some of the important concepts used in the toolkit to support a standard approach to Cross-industry RCM schemes.
- *Further Reading* (page 12) - suggestions on where to go for more information.

1.1 What is Cross-Industry RCM?

1.1.1 RCM - Definitions

In Remote Condition Monitoring (RCM), the health of an asset is observed by sensors which transmit their readings back to an office-based *IT* system. By processing the readings, the system can warn of actual or incipient failure of the asset and trigger maintenance / repair actions.

The potential advantages are well-recognised:

- continuous monitoring of the asset, rather than periodic inspection, potentially leading to earlier diagnosis of degradation or faults
- reduction of need to expose staff to dangerous on-track conditions
- a store of information which can be processed to improve understanding of asset behaviour.

Railway assets fall into two categories - fixed **infrastructure** and moving **trains**. Figure 1 below shows four typical scenarios for rail industry RCM - where the asset being monitored is an infrastructure or train one; and whether the monitoring is done from infrastructure-mounted or train-borne sensors.

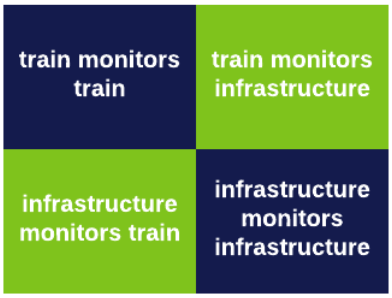


Figure 1: Rail RCM Quadrants

By "Cross-industry RCM" we mainly mean the two green quadrants: the scenarios which straddle the rail/wheel interface - train-borne equipment monitoring infrastructure, or track-side equipment monitoring trains.

These warrant particular attention because they have turned out to be much harder to achieve than the blue quadrants. The large potential benefits they offer the industry have yet to be realised. The blue scenarios, by contrast, are now commonplace: much critical infrastructure is fitted with remote sensors and all new trains monitor their own condition.

1.1.2 Some Examples of Current Cross-industry RCM Schemes

- is an example of infrastructure monitoring trains. A trackside array of microphones records the sound profile of vehicle bearings as the train passes it. Software analyses the acoustic fingerprint and can detect a variety of bearing wear scenarios, allowing them to be corrected before expensive and dangerous bearing failure. When coupled with Automatic Vehicle Identification (*AVI*) and a suitable vehicle database, the software can detect gradual degradation and generate maintenance work orders at appropriate times.
- (*UGMS*) is an example of trains monitoring infrastructure. Equipment mounted on rail vehicle bogies contains accelerometers and other sensors to measure track geometry. Readings are transmitted back to the ground with associated location information (*GPS* or other location sources available to the train). Ground-based software detects significant anomalies and generates alerts; or can consolidate multiple readings over the same piece of track to improve resolution and accuracy and assess the change in track condition over time.

1.1.3 Problems with Cross-industry RCM

Research work such as that done by the *XIRCMSG* (see *The Cross-Industry RCM Strategy Group* (page 63)) has exposed the types of problem that impede the take-up of Cross-industry RCM, even where a clear business requirement exists. These fall into these main categories:

- **business case**: it is hard to make a business case where one set of industry parties picks up the costs of a scheme but its benefits accrue to other parties
- **legal and commercial**: problems clarifying who owns the data, where do *IP* rights lie in the processes applied to it, where does liability lie for deficiencies in the data, what principles should apply to the use and further exploitation of the data
- **organisational**: who is in charge of the scheme and who needs to actually carry out tasks associated with it, including those associated with providing, maintaining and disposing of equipment; gathering, storing and transferring the data; and applying the data to improve business decisions.
- **technical**: how best to make data easy to use and easy to combine with data from other industry sources; how to support novel uses of the data; how best to ensure consistency of processing and interpretation across data from different sites, different suppliers and gathered at different times.

1.2 Industry Initiatives

Cross-industry RCM's problems reflect the need for more system-wide thinking that industry analysts see as necessary to drive the safety, capacity, availability and value-for-money improvements the railway needs to deliver.

The McNulty Rail Value for Money Study 2011 (page 62), commissioned to identify why the UK's railway was more expensive than comparators on the European mainland and slow to innovate, identified the lack of best practice in asset management as an industry weakness that needed to be addressed. To drive a 30% improvement by 2018/19, a *"whole-system, whole-life"* approach was recommended. The study also recommended a focus on better cross-industry information systems.

Looking more long-term, *The Rail Technical Strategy 2012* (page 62) (*RTS*) has a strategic 30-year goal to meet increasing demands on the railway by more radical change in philosophy. Several of its workstreams include the expansion of Cross-industry RCM use to help deliver them: infrastructure and rolling-stock maintenance moving to a more condition-based approach; IT changes to support integration, sharing and novel exploitation of RCM data.

Already in place at the time the *RTS* was written, and key to its delivery, is the *The Cross-Industry RCM Strategy Group* (page 63), a group tasked with leading the industry's efforts to facilitate Cross-industry RCM. The group has investigated in depth the impediments to Cross-industry RCM; its response was to commission *The T1010 Project* (page 5) to provide a toolkit setting out a structure, standards and guidance to help project practitioners assess the business case for their Cross-industry RCM scheme, make the necessary commercial arrangements and set up the technical aspects in a standards-compliant way. This website is supplementary guidance in the use of that toolkit.

Following the T1010 project, *XIRCMSG* sponsored *The IMPRCM Project* (page 6) to try out aspects of the T1010 toolkit in a real Cross-industry RCM project and deliver:

- additional guidance in the use of the toolkit, including best practice and experience gained (this website)
- a prototype data broker exhibiting the technical approach and standards proposed, to act also as a demonstrator of the capability to industry.

To implement the objectives of the *RTS*, the industry set up the *Capability Delivery Plan* (page 63) (*CDP*). This has twelve key capabilities, two of which overlap strongly with Cross-industry RCM: **"More Value from Data"** and **"Minimal Disruption to Train Services"**. *The IMPRCM Project* (page 6) will be liaising closely with the teams responsible for delivering these capabilities. This has now been rolled into Network Rail's research programme for *CP6*.

1.3 The T1010 Project

RSSB's research project *T1010* comprised four work packages, each delivering a report giving guidance and suggested standards for one of the aspects of Cross-industry RCM previously identified as being problematic. Exploratory work investigated existing Cross-industry RCM projects, existing RCM projects within industry sectors, approaches used in other industries and in the rail industry elsewhere and emerging information technologies; then four reports were produced:

- T1010-01: data architecture. A set of standards, guidelines and a suggested IT approach for interchanging RCM data between different parties, respecting their IP but maximising the openness and therefore future exploitability of the data.
- T1010-02: commercial / legal architecture. A process map and list of commercial principles, supplemented by a template contract and schedules, intended to assist parties in striking agreements for the provision, processing and use of Cross-industry RCM data.
- T1010-03: guidance and standards. An overview of the process needed to start, progress and complete a Cross-industry RCM scheme and of the existing standards and constraints that it would need to comply with.
- T1010-04: business case toolkit. A spreadsheet-based model and guidance to help quantify the overall business case and the party-specific business case for data-interchanging parties. This would indicate whether the scheme was at all viable and also provide a basis for commercial negotiations to address the issue that costs would likely be borne by one party but benefits accrue to another.

Although the main focus of the Toolkit is on the green quadrants in Figure 1, projects in the other two blue quadrants can also suffer similar challenges and thus potentially benefit from applying its guidance.

This website represents additional guidance at an overview level, above the detail present in the T1010 documents.

T1010 documents themselves can be accessed via the Reference area here: *T1010 Documentation* (page 61).

1.4 The IMPRCM Project

A recommendation of the T1010 project was that the approach should be tried out and the data principles prototyped in an actual small-scale Cross-industry RCM project to see if they were fit for purpose.

Stakeholders *TfL* and Network Rail Western Route identified a possible target project: the exchange of condition data about Network Rail track gathered by *UGMS* equipment mounted on Crossrail Class 345 trains running on the Western route, from Paddington initially to Heathrow Airport Junction but later to Reading.

A preliminary scoping exercise was carried out by *RSSB* to see how much of the T1010 toolkit would be useful for this project. Nearly all the commercial principles could be examined and the template contract tried; a large proportion of the technical approach and standards could be trialled; and the process guidance could be adapted to fit a project of this size. The business case toolkit could also be tried out, though the results would not impact the commercial arrangements because of the prototype nature of the work.

The resulting project is IMPRCM. The scope of the project includes:

- assisting the parties (chiefly NR Western Route and TfL) to scope the project, agree Heads of Terms and come to a commercial agreement
- developing a set of documentation to act as guidance for users of the T1010 toolkit (this website)
- building a prototype data broker, based on T1010 architectural principles, to mediate the exchange of data between *TfL* and *NR* Western Route
- carrying out a gap analysis between the current T1010 toolkit and the approach best fitting the pilot project, to identify scope for updating the toolkit and guidance
- demonstrating the benefits of the approach used to an industry body.

1.5 Key Principles of the T1010 Advice

The T1010 toolkit was put together to address reasons why Cross-industry RCM schemes have not taken off as well as desired. The reasons have been outlined above in *Problems with Cross-industry RCM* (page 4): they relate to the difficulties associated with industry fragmentation and siloisation.

Underpinning the guidance and standards defined in the 4 workpackages of *T1010* are a number of principles which all try to address this fragmentation. They can be divided into Commercial and Technical principles.

Additionally, some Philosophical Principles underpin the Cross-industry RCM goals.

1.5.1 Philosophical Principles

- data is presumed to be owned by the party sponsoring and paying for its collection
- liability should not attach to the data supplier for anything done with the data
- data supply should be accompanied by best estimates of its accuracy and precision, but data processors / consumers must deal with any deficiencies
- licensing restrictions should not be onerous and should not restrict novel use
- as far as possible, data should be freely accessible to any downstream users
- IP in data processing techniques can be safeguarded

1.5.2 Commercial Principles

The Commercial / Legal thread of T1010, T1010-02, set out a list of 13 commercial principles which the agreements between stakeholders should respect. The premise is that by adopting these principles, parties will save themselves effort, safeguard their interests and support future extension and enhancement of their schemes and other industry-wide ones.

The commercial principles are all referenced in *Commercial Guidance* (page 33). The gist of them is clear identification of:

- the parties involved and their roles and responsibilities
- the requirements for the data: content, quality and delivery arrangements
- the stages and gateways in the project lifecycle
- the ownership and license arrangements for data at all stages of its lifecycle
- who pays what and for what: equipment, data, storage and processing, maintenance, disposal.

1.5.3 Technical Principles

The Technical thread of T1010, T1010-01, sets out the principles to be followed in the data interchange. The goal here is to standardise as far as possible, to minimise the effort needed to develop a new data source, and to support the philosophical and commercial principles.

The key technical principles at work are:

- open and standard human-readable data formats. Data must be presented to consumers in plain text formats - *CSV*, *JSON* or *XML* - or a standard *MIME type* for audio or video media or similar.
- data must be easy to use - by end users using standard desktop software (e.g. MS-Excel or MS-Access); by analysts using standard analytic software (e.g. *MATLAB*, *Stata*, *R* or *Python/Pandas*); or by downstream computer systems using a *REST*-based *API*
- standard encodings for common data items: *ISO* standards for dates, times, *GPS* locations; railway standards for rail track locations, asset identifiers and operational data
- use of existing standards for rail data representation and transfer - including *RailML* and *RailTopoModel*.
- clear separation of RCM data (which is open) and RCM processing (which may be proprietary) according to the six-layer *ISO 13374* model (see *The ISO 13374 model* (page 7) for an introduction).
- simple use of industry reference data and translation / mapping services available by standard lookup mechanisms
- *metadata* available for all data in a standard format, providing information about data ownership, licensing, accuracy and precision so that downstream users know clearly what they can do with the data.
- support for a *data broker*-based industry architecture to mediate data interchanges. This will make it easier for data providers, data processors and data users to interact in novel ways, support the adoption of standards and reduce marginal costs.

1.5.4 The ISO 13374 model

1.5.4.1 Principles

ISO 13374 sets out to define a framework for the management of RCM in terms of data processing steps and data flows between them. The steps cover the whole spectrum of involvement of RCM data, right from its initial capture ("Data Acquistion"), through the generation of alerts and alarms about asset failure ("State Detection" and "Health Assessment") up to the most forward-looking use in remedial / preventative planning ("Advisory Generation").

Each type of processing that RCM data undertakes can be categorised as belonging to one of the steps. The steps can be:

- carried out by different parties or use different IT systems

- protected by *IP* safeguards.

The steps should have:

- openly available definitions of what they do and how they do it
- clearly-stated limitations on the dependability of the data they produce
- service level specifications on availability, performance, latency

The data flows between the processing steps should:

- use open and standard data formats
- contain data appropriate to the processing step

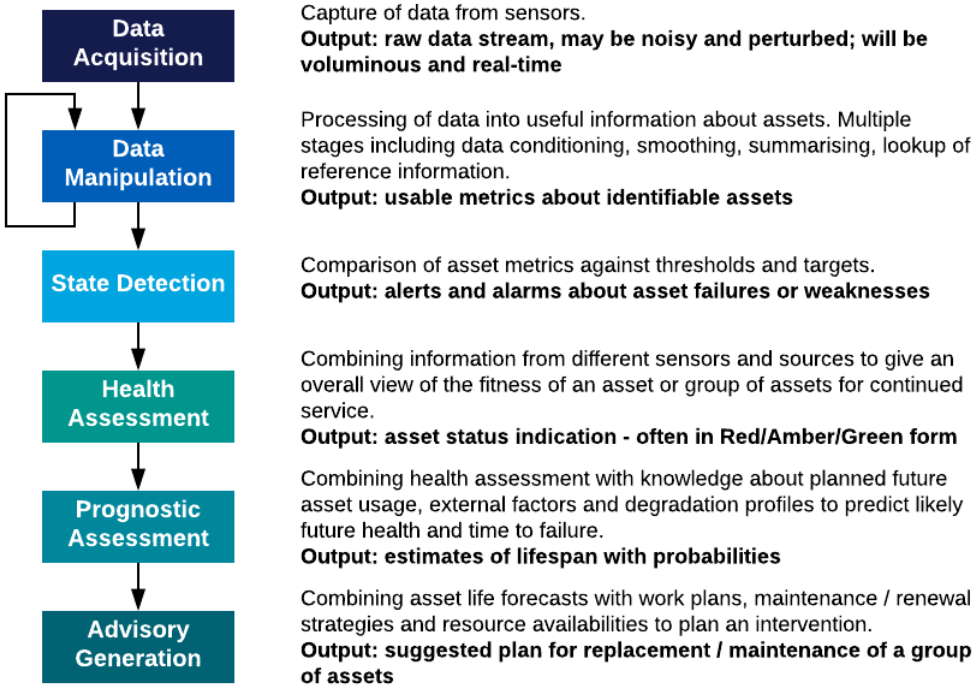


Figure 2: The ISO-13374 Processing Model

The processing steps are shown in Figure 2 and described below. They represent a sequential hierarchy in which each step builds upon its predecessors. At each step through the hierarchy:

- the volume of RCM data reduces, generally by orders of magnitude
- the amount of contextual information about the assets and their role in the railway increases.

Typical current *RCM* systems work internally at the first three levels - *Data Acquisition (DA)* (page 9), *Data Manipulation (DM)* (page 9), *State Detection (SD)* (page 10) - and generate results at a basic *Health Assessment (HA)* (page 10) level. They do not generally give access to the more detailed data of the first levels, citing concerns about data volume and bandwidth, *IP* and the difficulty of interpreting the data. They generally struggle to reach the higher levels - *Prognostic Assessment (PA)* (page 10) and *Advisory Generation (AG)* (page 11) - because of the difficulty in integrating data and business processes across different systems and different rail business domains.

The *XIRCMSG*'s standardisation efforts aim to extend this typical structure in two ways using the *ISO 13374* framework:

- to make data at the lower levels more readily available in standard forms so that novel uses can be made of it. This involves providing a technical framework and a contractual setup to address confidentiality and *IP* concerns.

- to support efforts to reach the higher levels by lowering barriers to integration of data from other rail domains.

1.5.4.2 Data Acquisition (DA)

This is the first stage, where physical characteristics of the asset are converted to electronic signals by sensors. Many characteristics could be captured. Some examples:

- **acceleration**, captured by accelerometers. Used to detect misalignments or anomalies in, for example, track, *OLE*, or vehicle wheels
- **noise**, captured by microphone. Used to detect wear in wheel bearings, for example
- **appearance**, captured by video camera. Used to detect missing components, uneven wear etc
- **temperature**, captured by thermometer or infra-red detector. Used to detect overheated wheel bearings etc

At this stage, the data readings are raw and subject to all kinds of noise and sensor misalignment / miscalibration and require processing to be of value in RCM. Also, they have little context: an accelerometer reading will have no knowledge of which piece of track or OLE it relates to; the timestamp associated with the data item may be inaccurate.

The data stream could be voluminous and occur continuously in real-time, particularly where the sensors are on vehicles and are monitoring the infrastructure. An important consideration is how much of this raw data stream is stored, and whereabouts is processing done to reduce its volume and extract meaningful information that can be transmitted from train to ground with the bandwidth available. There is value in preserving this raw data stream if possible, as it makes it possible to use it for new types of investigation in future.

1.5.4.3 Data Manipulation (DM)

In this stage, the raw data stream is processed to make it useful for asset management purposes. This is generally a multi-step process, the steps being different for the different types of asset data. Some typical steps are:

- **data cleansing**: removal or correction of anomalies caused by broken or malfunctioning sensors or data transmission glitches
- **calibration**: scaling the data values so that they match physical reality independent of the response of individual sensors
- **time correction**: adjusting for the differences between time recorded by the sensors and the actual time. (Sensors' clocks can drift and can suffer from other synchronisation issues including failure to account for Summer Time adjustments)
- **referencing**: attaching a track location (for infrastructure) or vehicle identity (for vehicles) to the data. This is an important topic, discussed in *Supporting Processes* (page 11) below. Other contextualisation, such as providing information about the asset being monitored
- **aggregation**: summarising the data stream, such as by calculating totals, time-based averages, standard deviations or other statistical measures
- **trending**: looking for change over time
- **feature extraction**: picking the elements of the data that suggest the presence of anomalies.

Each of these steps requires the use of an algorithm or referencing process. Different parties may use different algorithms and different reference data sources which may not be current: this leads to lack of consistency in the data generated. There is thus scope for standardisation here, which the industry can support by:

- making standard services available to all RCM data users, e.g. for conversion from *GPS* co-ordinates to track co-ordinates (*ELR* / Track ID / yards or chains); or for vehicle identification.
- allowing data users to choose the algorithm they use from those provided by 3rd-party providers, by supporting a data broker which makes these readily available.
- standardising the presentation of metadata describing the algorithms and their characteristics and limitations.

In most current RCM systems, the processing and data associated with this stage are internal and not accessible to analysts or data owners; and the algorithms used are often perceived to be proprietary. However, one of the main goals of the Cross-industry RCMSG effort is to reassure data processors that they can make data at this level available while protecting the *IP* associated with the algorithms and processes.

The output of this stage is conditioned data that can be used to make decisions about the status of assets.

1.5.4.4 State Detection (SD)

In this stage, the data readings from the DM step are compared with boundary values or thresholds to determine whether they represent normal or abnormal state. The comparisons may result in

- indicators of state (OK / near threshold / over threshold)
- indicators of level of exceedence of threshold
- indicators of rate of change
- comparators against expected levels given by e.g. "bathtub" curves.

The thresholds or boundaries will be set by human intervention and may be refined over time as the knowledge of asset behaviour improves. A key issue here is ensuring that the volume of triggers generated is appropriate: enough to enable asset performance to be improved; not so many that they can't be handled by operators.

1.5.4.5 Health Assessment (HA)

The job of this stage is to assemble State and other data from the various sensors associated with an asset to come up with an overall view of its condition. This may involve logic to resolve differences of opinion from different sensors on the same asset or to embody professional knowledge on its fitness for service although sensors indicate cause for concern.

In the Business Case tool described in T1010-04, this level is divided into two - **HA-1** and **HA-2**[1].

- **HA-1** is a basic level of analysis in which different sensors on the same asset are combined - for example *HABD* and acoustic readings of the same bearing.
- **HA-2** is a more advanced level, in which sensors on different assets in a group or assembly are combined to give a view of the health of the whole assembly.

The output of this process is alerts and alarms about the capability or fitness for service of the asset or assembly. These may include indications of the likely fault condition or failure mode, with probabilities and explanation of the findings.

Most existing *RCM* systems generate output for their users at the HA (typically HA-1) level: they report on the health of assets and generate alerts or alarms for action by asset users / maintainers.

This is an area where asset owners may see sensitivity, not wishing to divulge to 3rd parties or their competitors the ways in which their products are prone to fail. It is also an area where industry-wide efforts to maximise the benefits of Cross-industry RCM will require standardisation of failure modes and fault-type coding.

1.5.4.6 Prognostic Assessment (PA)

In this stage, forecasts of the asset's future performance are generated from the asset's current state (from the HA stage), knowledge about the degradation rates, and estimates of future degrading factors such as usage or weather. The assessment may include:

- likely asset health at future dates
- likely time of failure
- likely failure mode(s)

and could be accompanied by probability distributions, confidence limits, explanations of the logic used.

[1] See definition in

This is the level at which condition-based maintenance of assets becomes a realistic possibility and so should be seen as the medium- to long-term goal of Cross-industry RCM efforts. It requires integration of data from different railway domains and external data sources in order to become effective:

- future asset usage for railway assets is driven by the timetable and resource plans (from the operational domain)
- the industry cost of asset failure or degradation depends on the expected impact on passengers and freight customers (from the commercial domain)
- degradation also depends on external influences such as weather (temperature / sunshine / rainfall / wind / storm surge) and tide.

It also requires a good understanding of the expected deterioration rates and characteristics of assets by their type, age, usage and maintenance history. This type of understanding is patchy and incomplete.

Both these requirements need industry support to make the data flows readily accessible and to support the accumulation of knowledge of asset behaviour.

1.5.4.7 Advisory Generation (AG)

This is the highest level of the stack, in which data about the future health of assets from the PA level is merged with contextual data which could include:

- means of maintenance and repair
- costs and impacts of the possible failure modes and different interventions
- availability of money and physical resources

to come up with an optimal remedial plan. It requires the highest level of data integration and the most radical re-work of business processes in order to be effective.

The output of the process is a set of recommended actions or alternative plans, with justifications. These may be interfaced automatically to maintenance planning or workbank systems.

1.5.4.8 Context

Figure 3 shows the processing blocks of the six-stage hierarchy in context, with data flows between them and other supporting sub-systems.

The subsystems are:

- **Configuration and Reference**. Each of the processing blocks needs supporting information about its own situation, the assets it is associated with, its observed own state of health, its calibration, its physical connections with others. These are held in a configuration block
- **History**. Past data about asset condition needs to be stored to make it available to calculate statistical summaries such as moving averages, to support queries of historic data and to detect rates of change. The historic data may be used by other processes in the hierarchy or by external systems or operators.
- **Visualisation**. The processing blocks need to be observed by humans for two reasons - firstly to ensure that they are working correctly (for technicians); and secondly to observe the results that they produce (for asset managers). Any real RCM system will have a user interface that offers this type of facility.

Each of the processing blocks can interchange data with these subsystems. It can also read data from its predecessors in the hierarchy.

1.5.5 Supporting Processes

For any Cross-Industry RCM initiative, there is a key problem which needs to be addressed: the identification of the assets being monitored. For the blue quadrants in *Rail RCM Quadrants* (page 4) - train monitors train and infrastructure monitors infrastructure - this is not difficult. However, for the green quadrants, it is a crucial problem that needs an enabling technology to make it possible.

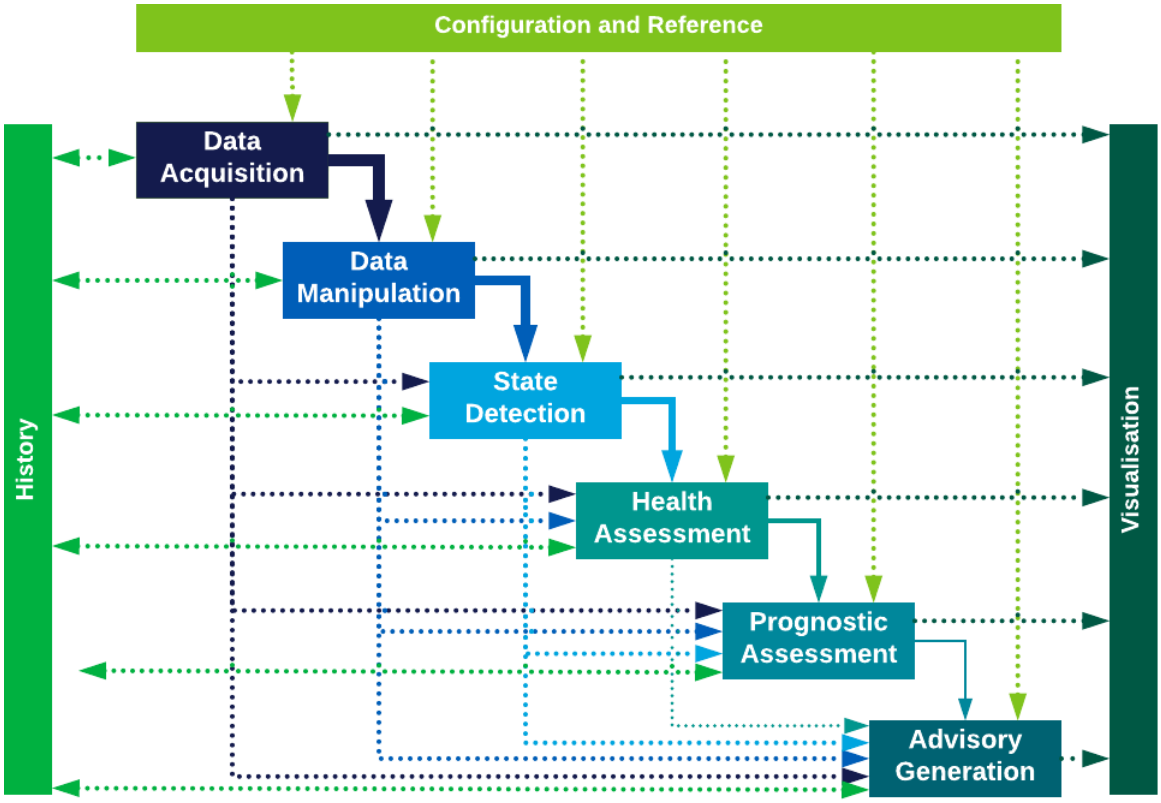


Figure 3: ISO 13374 - Links and Context

- For **train monitors infrastructure**, the issue is one of frames of reference and location precision: which piece of track or OLE does the measurement relate to? Train-borne sensors are not aware of the location of data readings in terms that Track or *OLE* engineers require, so a conversion or mapping is required. Typically the sensor data is tagged with a *GPS* location (which may not be very precise: typical precision is quoted as +/- 15m with 95% probability)' and may also have readings from the train's odometer, and possibly other supporting information about which fixed points such as *ETCS* balises or *AWS* magnets have most recently been passed. The engineers require location in the form of *ELR*, track ID and miles, chains or yards. There is a definite need here for a location-translation service to convert from one to the other which can be used by any RCM project.
- For **infrastructure monitors train**, the issue is one of vehicle identification. The RCM system needs to know which vehicle a reading relates to; and it must be possible to link that to a multiple unit or rake, to a train formation, to an operational train. The identification can be direct, say by using *RFID* tags attached to vehicles, supported by a database which maps tags to vehicles (*AVI*); or indirect, based on the synthesis of information from train describers and rolling stock dispatching systems. Currently, RCM systems of this type either have no such facility, or they have a project-specific one which can only identify vehicles belonging to a single client. There is a need for a standardised *AVI* service which can be shared by any RCM project.

1.6 Further Reading

See *Reference* (page 61) for more detail on the industry context, the T1010 documents and relevant standards.

CHAPTER 2

Process Map

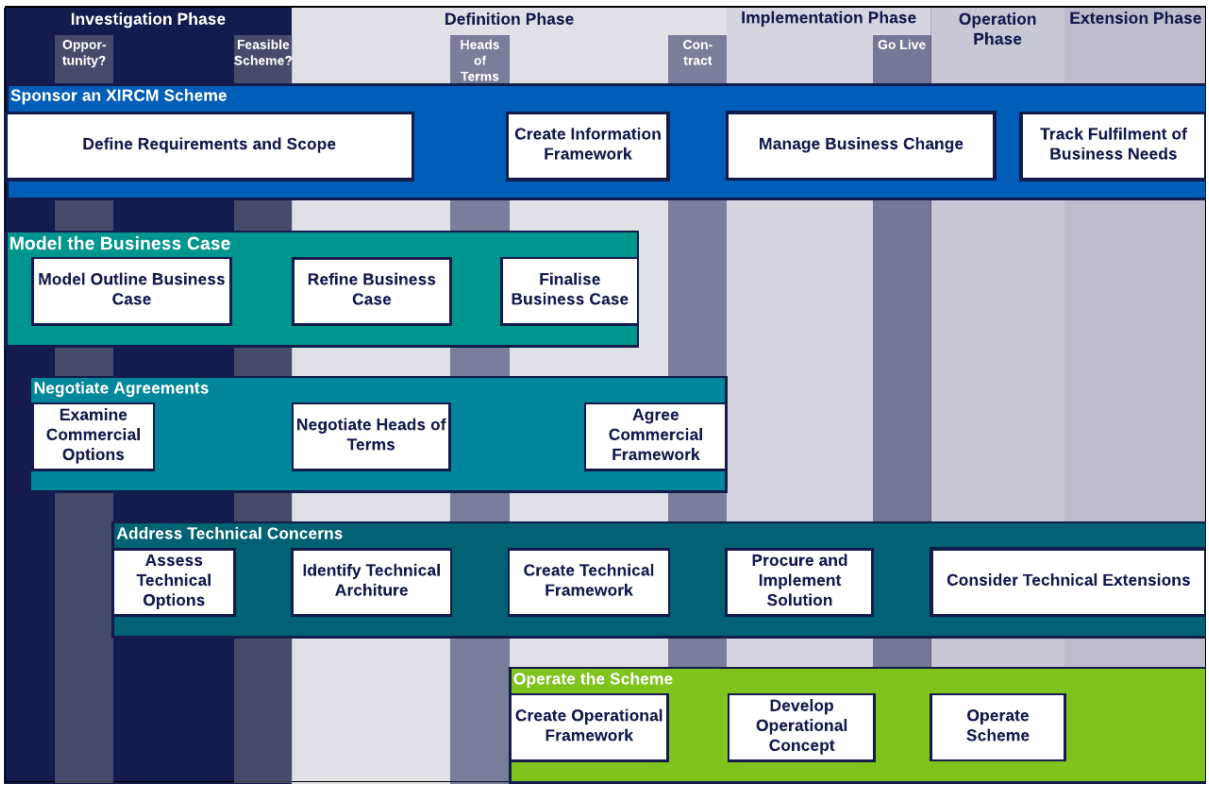


Figure 1: Cross-industry RCM Project Process Map - Overview

This diagram shows the overall lifecycle of a Cross-industry RCM scheme, in terms of phases, gateways, work-streams and the main tasks to be carried out.

It expands the overview map shown in getting-started and summarises the business process maps described in T1010-02 Appendix C2[1] and T1010-03 Best Practice Guide[2].

[1] Source: T1010-02 Appendix C2 - Process Map
[2] Soucre: T1010-03 Section 2 Project Lifecycle

2.1 Phases and Gateways

a Cross-industry RCM project can be considered as having a sequence of phases. At the end of each phase, and sometimes part-way through it, are gateways at which decisions are made whether and how to proceed with the project.

The phases of the project lifecycle and the associated gateways are described in the following sections.

2.1.1 Investigation Phase

This initial phase evaluates whether a Cross-industry RCM scheme is necessary and whether there is any chance of it being commercially and technically feasible. Outline work is done on the business need, the expected scope of the project, a business case and possible stakeholders and commercial arrangements.

2.1.1.1 Gateway "Opportunity?"

As early as possible during the Investigation Phase, a decision needs to be taken based on the scoping and business case work done so far as to whether a case has emerged for using Cross-industry RCM, or could the business need be met another way? If the answer is no, the scheme is abandoned.

If there is a Cross-industry RCM opportunity, work continues to identify technical options for sourcing, delivering, storing and using the Cross-industry RCM data. These options are costed in outline and the results used to inform the developing business case. Stakeholders and involved parties are identified and the basic elements of commercial / legal relationships with them included in the commercial options.

2.1.1.2 Gateway "Feasible Scheme?"

Once a potentially-feasible scheme can be envisaged, the next point of decision is whether to proceed to define the scheme and move towards a formal agreement to implement it. This is the "Feasible Scheme" gateway.

If there is no technically / financially feasible scheme, the process is halted. If there is, work now proceeds to the Definition Phase.

2.1.2 Definition Phase

In this phase, the proposed Cross-industry RCM scheme is worked up to a full set of commercial agreements with all the necessary schedules specifying the commercial, legal and technical elements of the scheme.

This involves moving into detailed documentation of requirements, identification of data needs and the technical resources needed to support them, identification of all the concerned stakeholders and contractual capture of their interests, and the compilation of documents which record these:

- *Information Framework* (page 45): the statement of information requirements, in terms of content and other relevant characteristics such as timeliness, completeness, accuracy, availability etc.
- *Technical Framework* (page 54): the statement of technical requirements required to provide the information defined in the Information Framework.
- The *Heads of Terms* (page 32): a statement of commercial intent
- The *Full Contract and Schedules* (page 32): a formal agreement between each of the parties.

Guidance on preparing these documents is available: - *Commercial Guidance* (page 33) gives general guidance on the technical documentation.

2.1.2.1 Gateway "Heads of Terms"

This gateway records the agreement of the *Heads of Terms* (page 32) between the parties and thus the willingness to procesd with the project at least as far as a commercial contract. If agreement cannot be reached at this stage, the alternatives are: - conclude the project - agree to work under a looser arrangement such as a non-disclosure agreement.

2.1.2.2 Gateway "Contract"

This gateway records marks the signing of formal agreements between the parties to the scheme, upon which the scheme will be implemented. The scheme now proceeds into the Implementation Phase.

2.1.3 Implementation Phase

In this phase, the parties work together to bring the scheme into live operation. This involves procurement, development, testing and deployment of the data sharing solution, the change management necessary for it to be adopted as part of business processes, and planning for live operation.

2.1.3.1 Gateway "Go Live"

This gateway is a formal review of the readiness of the solution to pass into live operation. The solution must have passed acceptance tests and be deployable; all stakeholders must be ready for the process change that the solution will enable or require; service level agreements and operational guidelines must be in place.

2.1.4 Operation Phase

This is the phase in which the scheme is operated. Tasks involve monitoring its technical performance, checking that the scheme is fulfilling the business goals it was set up for, and identifying future enhancements or modifications to the scheme, for example to meet new business needs or capitalise on new data sources.

2.1.5 Extension Phase

enhancing the scheme to move to more fundamental integration with business processes or novel uses of the information captured.

2.2 Workstreams

Workstreams represent sequences of activities carried out by stakeholders with a particular interest in aspects of delivering the project. The workstreams are associated with particular scheme roles: - Sponsor: business representative whose work is intended to benefit

2.2.1 Sponsor Workstream

Ensuring that the scheme addresses genuine business needs. Ensuring the scheme and its timetable and risks are managed in a proper business change framework. Identifying, engaging and managing parties who need to interact with the scheme or who are otherwise involved.

See *Sponsor a Cross-industry RCM scheme* (page 16) for full details of the role of the scheme sponsor.

2.2.2 Commercial / Legal Workstream

Evaluating the costs and benefits of the scheme. Tackling the commercial and legal arrangements between parties to the scheme. Drawing up the contractual documents: Heads of Terms and Contract.

See *Negotiate agreements with stakeholders* (page 30) for full details of the role of the scheme commercial/legal teams.

2.2.3 Technical Workstream

Managing the interchange of data, ensuring relevant standards and guidelines are respected.

See *Address the technical concerns of a Cross-industry RCM scheme* (page 56) for more on the technical workstream.

2.2.4 Operational Workstream

Running the scheme in live operation.

See *Operate the Cross-industry RCM scheme* (page 59) for more on the operational workstream.

CHAPTER 3

Sponsor a Cross-industry RCM scheme

A Cross-industry RCM project needs a business sponsor to take responsibility for

- defining the goals of the scheme
- identifying the stakeholders and funders
- overseeing the management of its definition, implementation and delivery
- managing the business change needed to make it effective
- ensuring it delivers the promised business benefit
- seeking opportunities to enhance the benefit in future work on the project and others that could learn from it.

The *T1010* toolkit offers some advice and guidance to project sponsors which is summarised in this page. Links are provided to the source T1010 documents as appropriate.

The sponsor will need to work within their organisation's project control process (such as Network Rail's *GRIP* process). This page is not a guide to these processes but will indicate where Cross-industry RCM considerations might feed into them.

3.1 Context

See *Introduction to Cross-industry RCM* (page 3) for a lay-person's guide to Cross-industry RCM.

The figure Figure 1 shows the Sponsor workstream in the context of the project phases and gateways and the other workstreams. More detail on these aspects of a Cross-industry RCM project can be found in *Process Map* (page 13) and the pages devoted to the other workstreams.

3.2 The Sponsor's Tasks

3.2.1 Define Requirements and Scope

The requirements for the scheme need to be identified early on in the *Investigation Phase* (page 14), refined during the *Definition Phase* (page 14) and validated as part of the testing in the *Implementation Phase* (page 15).

This is essentially a business process re-engineering task in which a beneficial business change is identified and the changes to processes, data flows, stakeholder interactions and job-specifications are worked out.

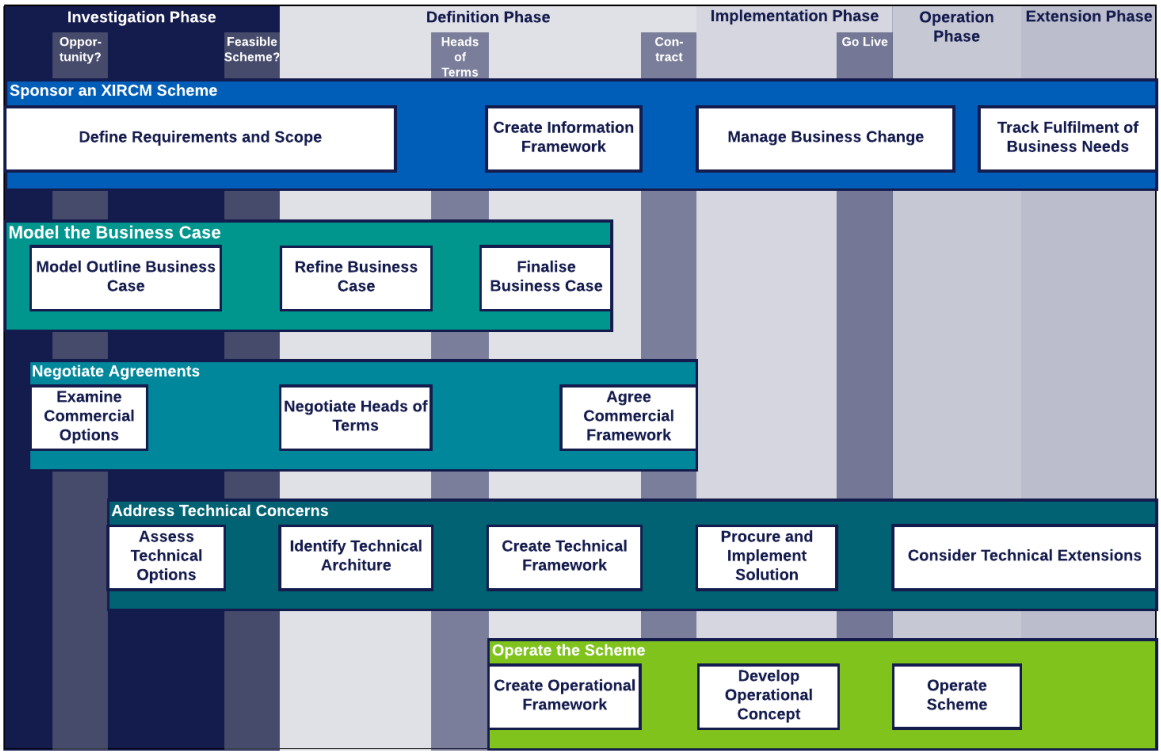


Figure 1: Cross-industry RCM Project Process Map - Overview

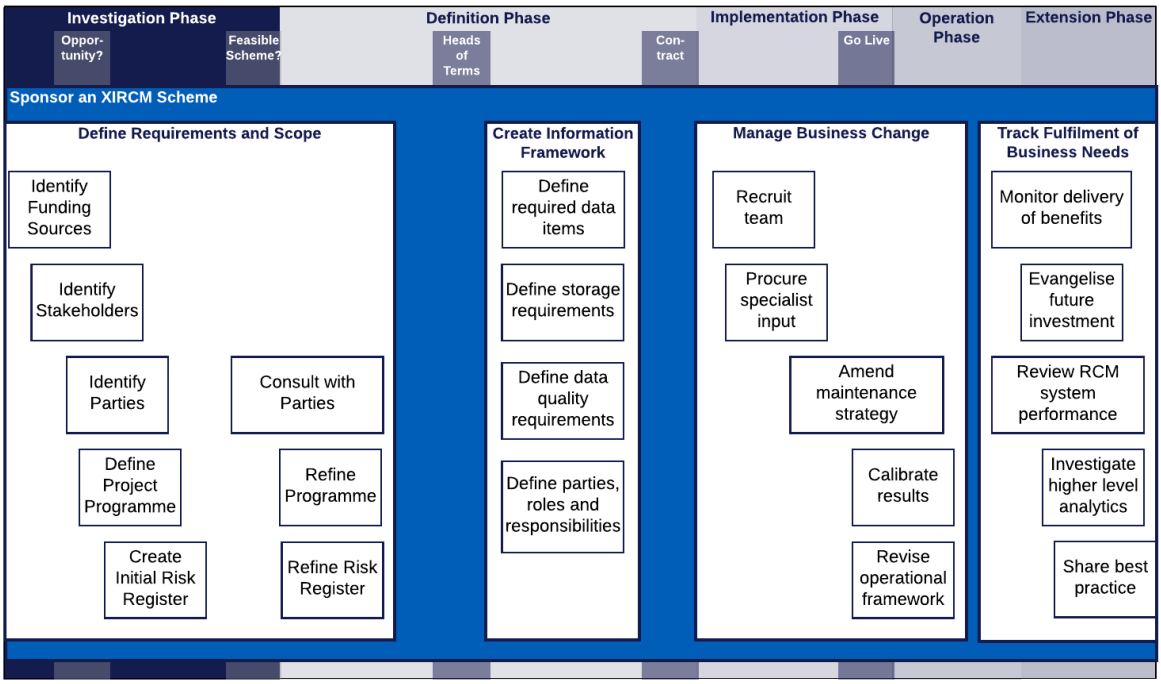


Figure 2: Sponsor Workstream

The task will need to encompass the detailed sub-tasks shown in Figure 2 since these generate important scheme artefacts.

The artefacts that this task creates are:

- A set of **User Requirements** and the outline of an **Acceptance Test Strategy**.
- A **Funding Statement** indicating, in outline at least, where the funding for the project will come from at each of its stages.
- A **Stakeholder Matrix** indicating the roles, responsibilities and interests of the stakeholders, plus an engagement plan for how to interact with them through the project.
- A **Project Programme** specifying the timings of project activities and the interactions and dependencies between them.
- A **Risk Register** identifying the risks to the project's outcome, budget and timetable and external risks, plus mitigation measures.

Tip: In the IMPRCM pilot, these analysis techniques proved useful in defining the scheme requirements and scope:

- **User Story analysis**. We created a set of candidate user stories based on discussions with project stakeholders, particularly those associated with the potential uses for the data gathered: track maintenance engineers and asset management specialists. The user stories relate directly to desired business benefits and therefore act as a good starting point for more detailed analysis and the creation of user acceptance tests. From the set, we focussed on one that could generate benefit in the short run without fundamental business process change. See further guidance and example here: *User stories* (page 20).
- **Stakeholder Interaction Matrix**. This is a matrix showing the sources and targets of the information flows that would be needed to satisfy the selected user story. It was a helpful aid to discussion with project stakeholders to help define the data flow and the relationships implied. This enabled all the stakeholders to be identified, plus the essential characteristics of the interfaces between them. Each of these interfaces could potentially require a contractual relationship to be established. See further guidance and an example matrix here: *Stakeholder Interaction Matrix* (page 20).

Tip: **Scheme requirements** were collected from several sources: railway standards, the *T1010* data architecture, other data sharing initiatives including *In2Rail*, and specific demands of the selected user story. See further guidance here: *Requirements and Acceptance Tests* (page 23). The Sparx Enterprise Architect tool (Enterprise Architect) was used to capture and organise the requirements.

3.2.2 Create Information Framework

The Information Framework is a document which defines the content and other characteristics of the data flow needed to satisfy the scheme requirements. It covers:

- the **data content** - the data items required. These will include the condition measures, but will also contain vital context such as details of the assets being monitored, the time and date of data collection, calibration and status information for the capturing and processing equipment, and useful background and context such as the identities of trains or locations.
- **data quality requirements** - timeliness, accuracy, precision, consistency, completeness, readability, standards compliance, metadata required. The quality requirements should be set at levels that satisfy the business need, but do not "gold-plate" or add complexity or cost unnecessarily (see *Gold Plating*).
- **data storage requirements** - how long should data be stored for, at what level of completeness or detail, how the data should be accessed, how quickly it should be accessible, what are the consequences of data loss, backup / restore arrangements, redundancy and disaster recovery.
- **data pipeline** - what is the data pipeline from source through to consumption at the final target? What storage, transmission, manipulation and processing is done en route and by whom?.

For more guidance and best practice on preparing the Information Framework document, see *Information Framework* (page 45).

3.2.3 Manage Business Change

The success of the scheme will depend on how effectively the data flow it offers is embedded into the working practices of the data users. The business change necessary can have several important aspects that the Sponsor needs to manage:

- alignment with, or changes to, existing Railway Group Standards or Company Standards, to allow the use of the new data flow and analysis from it in safety-related areas such as asset maintenance.
- changes to other linked or associated IT systems
- changes to documented processes
- changes to staff job descriptions
- staff re-skilling and acquisition
- special skills required to assist in transition
- parallel operation, fallback, backup processes
- fine-tuning and calibration of the scheme outputs
- review of the system operational framework after bedding-in and some live running
- phase-out of replaced data streams and systems
- management of reputation and publicity.

Note: These aspects have not been covered to any depth in the IMPRCM work.

3.2.4 Track Fulfilment of Business Needs

The success of the scheme depends on how well it delivers the benefits forecast for it during the investigation and definition phases, and to what extent it is seen as a strong platform for future enhancement. The Sponsor needs to:

- check the extent to which the scheme is delivering the benefits
- monitor and review the performance of the RCM scheme against its SLA and expectations
- look for opportunities to get more benefit by moving up the *ISO 13374* maturity levels: bringing in new contextual data sources and doing more sophisticated analysis of the current data sources.
- work to share good practice and lessons learned, and to absorb them from other projects to improve the effectiveness of Cross-industry RCM both for the current scheme and more broadly.

3.3 Interactions with other workstreams

As the prime mover of the Cross-industry RCM scheme, the Sponsor's work drives that of the other scheme workstreams. In particular, the documents created by this workstream are inputs to the more detailed workstreams looking at the business case, the commercial and legal agreements and the technical aspects of the scheme.

3.3.1 Business Case workstream

The business case workstream is described in *Model the Business Case* (page 26).

The Sponsor workstream establishes the business goals of the project which will be expressed as benefits that get realised by the reduction in costs of management, maintenance and repair of assets on the other side of the wheel/rail interface. The business case effort seeks to quantify these benefits.

This workstream also aims to identify how the scheme changes the *ISO 13374* maturity level and therefore what level of investment in technnology and business process change will be necessary. This has an impact on the scheme costs and timing.

3.3.2 Negotiate Agreements workstream

The commercial negotiation workstream is described in *Negotiate agreements with stakeholders* (page 30).

The Sponsor workstream identifies several important aspects of the scheme that affect the commercial workstream:

- The **stakeholders** and the interfaces between them, which may require commercial agreements
- The **requirements** which define the purpose of the scheme and the regulatory context
- The **Information Framework** which feeds into the Data and Service Level Agreement schedules
- The **project programme** which feeds into the contract text and Programme schedule
- The **risk register** which feeds into the Risks schedule and contract conditions.

3.3.3 Address Technical Concerns workstream

The technical concerns workstream is described in *Address the technical concerns of a Cross-industry RCM scheme* (page 56).

The Sponsor workstream directly drives the work of the technical workstream:

- Any **interface between stakeholders** requires a technical component to deliver it
- The **data items** in the Information Framework form the data content of interfaces
- The **data storage requirements** drive the hardware requirement
- The **data quality** considerations influence the architecture and scale of technical approaches
- The implied *ISO 13374* maturity levels drive the standards and data interchange mechanisms to be considers.

3.3.4 Operate the Scheme workstream

The operation workstream is described in *Operate the Cross-industry RCM scheme* (page 59).

The Operational Framework describes how the scheme will run day-to-day. This is driven by the data quality requirements defined in the Information Framework. Also important are the various aspects of storage accessibility, data safety and access management which flow from the other aspects of the Information Framework.

3.4 Guidance - IMPRCM experience

3.4.1 User stories

The IMPRCM team used Github issues as a simple way to create, share, link and track user stories. Figure 3 shows a sample user story in github.

Background on the use of User Stories as a requirements-gathering technique can be found in various places. This is a good starting point: User Stories.

3.4.2 Stakeholder Interaction Matrix

The IMPRCM team built a Stakeholder Interaction Matrix using Excel. Figure 4 shows a screenshot.

Tip: Use the cells of the matrix to highlight key aspects of each interface for further consideration - these may be sensitivities of stakeholders to liability, use of their data or similar; or particular constraints on the interface or the relationship that will need to be reflected in any agreement.

Story

As a Track maintenance supervisor

I want to be assured that repair work of previously-detected Switch/Crossing alignment defects, recorded in the *Ellipse* system, has been done correctly and that the repairs are resilient

So that I can return the route to full capability

Testing / Acceptance Criteria

precursors / dependencies

- data from different runs by different units over the same track sections can be calibrated
- data from different runs can be spatially aligned

direct requirements
I need to be able to verify that:

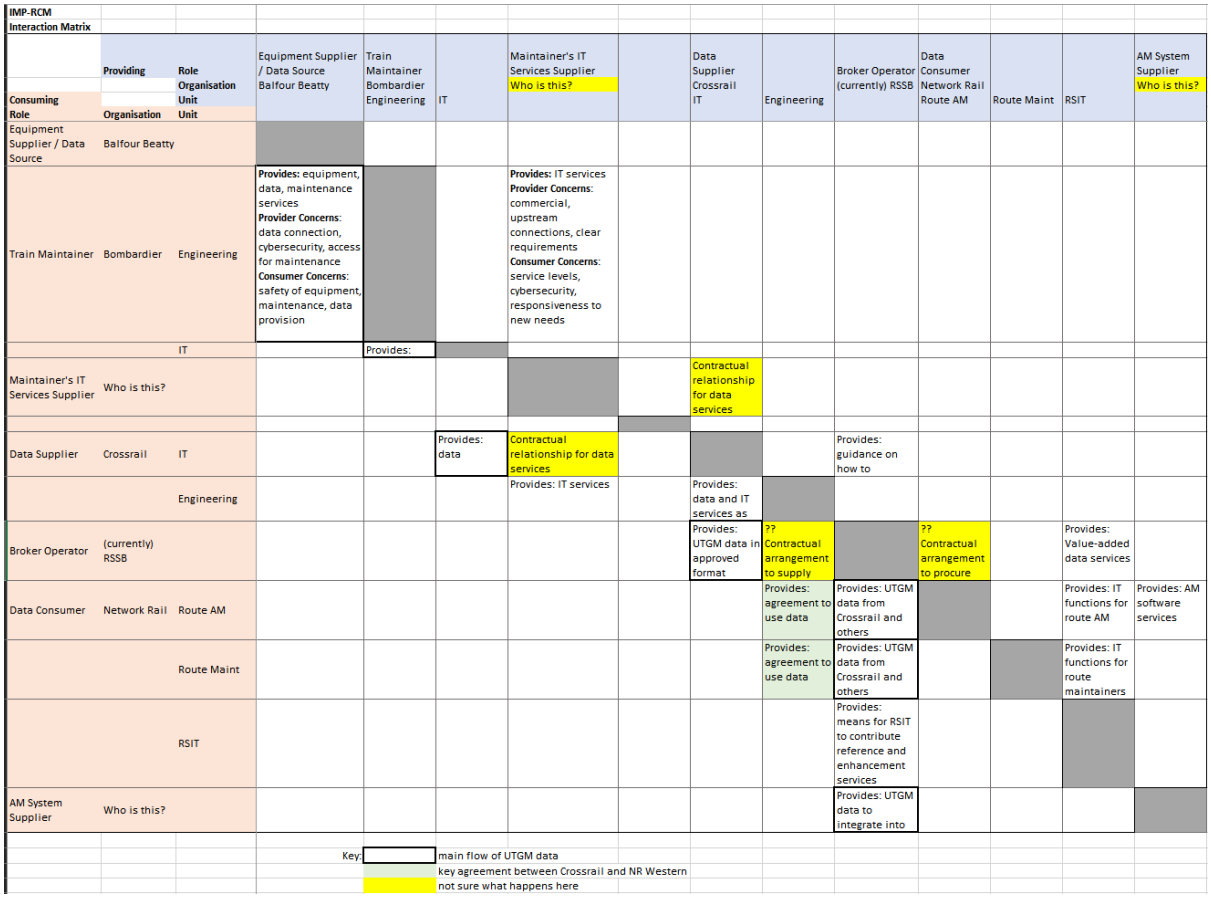
- a previous defect was detectable in the UTGM data
- a record of a repair intervention has been made in Ellipse and made available to the Data Broker
- a sufficient number of subsequent UTGM runs over the same previous defect show that it has been removed and the track is now demonstrably of acceptable quality
- track geometry does not rapidly degrade

Notes

This is a story which will require the development of several key capabilities before it can be satisfied:

- the listed precursors
- a data input from Ellipse showing work done
- rapid degradation can be identified (see #8)
- presentation of before / after / improvement data to engineers
- verification of data quality and the reliability of the conclusion that the track is now fixed.

Figure 3: Sample User Story



IMP-RCM Interaction Matrix														
	Providing	Role	Equipment Supplier / Data Source Balfour Beatty	Train Maintainer Bombardier Engineering	IT	Maintainer's IT Services Supplier Who is this?		Data Supplier Crossrail IT	Engineering	Broker Operator (currently) RSSB	Data Consumer Network Rail Route AM	Route Maint	RSIT	AM System Supplier Who is this?
Consuming		Organisation												
Role	Organisation	Unit												
Equipment Supplier / Data Source	Balfour Beatty													
Train Maintainer	Bombardier	Engineering	**Provides:** equipment, data, maintenance services **Provider Concerns**: data connection, cybersecurity, access for maintenance **Consumer Concerns**: safety of equipment, maintenance, data provision			**Provides:** IT services **Provider Concerns**: commercial, upstream connections, clear requirements **Consumer Concerns**: service levels, cybersecurity, responsiveness to new needs								
		IT		Provides:										
Maintainer's IT Services Supplier	Who is this?							Contractual relationship for data services						
Data Supplier	Crossrail	IT			Provides: data	Contractual relationship for data services				Provides: guidance on how to				
		Engineering				Provides: IT services		Provides: data and IT services as						
Broker Operator	(currently) RSSB							Provides: UTGM data in approved format	?? Contractual arrangement to supply		?? Contractual arrangement to procure		Provides: Value-added data services	
Data Consumer	Network Rail	Route AM							Provides: agreement to use data	Provides: UTGM data from Crossrail and others			Provides: IT functions for route AM	Provides: AM software services
		Route Maint							Provides: agreement to use data	Provides: UTGM data from Crossrail and others			Provides: IT functions for route maintainers	
		RSIT								Provides: means for RSIT to contribute reference and enhancement services				
AM System Supplier	Who is this?									Provides: UTGM data to integrate into				
			Key:		main flow of UTGM data									
					key agreement between Crossrail and NR Western									
					not sure what happens here									

Figure 4: Sample Stakeholder Interaction Matrix

The actual matrix used by the IMPRCM project is available for download here: `Stakeholder Interaction Matrix IMPRCM Example`.

3.4.3 Requirements and Acceptance Tests

For the IMPRCM prototype data broker, the requirements for the data interchange and the prototype broker were assembled into a spreadsheet, and subsequently into a Sparx Enterprise Architect repository[1], from the following sources:

- the requirements of the T1010-01 data architecture deemed to be relevant to the IMPRCM broker, given that it was a proof of concept and not a fully-specified production-quality solution.[2]
- requirements of the *In2Rail* programme data and application architecture.[3]
- requirements driven by the chosen User Story: verification that S&C repairs have been correctly carried out.

The requirements were analysed and a set of "features" created: deliverable elements of the prototype solution. For each of these, a number of acceptance tests was devised. These were specified using the *Gherkin* language.

[1] A downloadable version of the requirements repository in Enterprise Architect format is here: `Requirements Repository`.

[2] T1010-01 requirements are in the "Architecture Principles and Requirements" document linked to here: *The Technical Architecture* (page 62). Of these, roughly 70% were seen to be relevant to the protoype broker.

[3] *In2Rail* requirements were supplied by Network Rail's representative on the programme. General details on the programme are here: In2Rail. The requirements for the Application Framework and the Canonical Data Model are listed here: In2Rail Application Architecture requirements.

CHAPTER 4

Investigate a Cross-industry RCM opportunity

This page contains all the guidance, or links to guidance, that may be needed by someone investigating a potential Cross Industry Remote Condition Monitoring scheme.

4.1 Planning a potential XI RCM scheme

This area explains the lifecycle and tasks involved in getting an XI RCM scheme under way.

The investigation phase evaluates whether a potential Cross-Industry RCM scheme is viable, via the gateways of Opportunity and Feasibility. Outline work should be carried out on business requirements, expected scope, a business case and possible stakeholders and commercial arrangements.

4.1.1 Investigation Phases

4.1.1.1 Gateway "Opportunity?"

In the first instance, the sponsor must be able to demonstrate that the scheme is necessary. This can be based on initial business case and scoping work, and must demonstrate that there is a case for using Cross-Industry RCM that could not be met by existing capabilities or a simpler solution than cross industry monitoring. A decision then needs to be taken on whether the scheme investigation should proceed, or be abandoned.

4.1.1.2 Gateway "Feasible Scheme?"

Secondly, it must be demonstrated that the proposed scheme is likely to be commercially or technically feasible. Work continues to identify technical options for sourcing, delivering, storing and using the Cross-industry RCM data. These options are costed in outline and the results used to inform the developing business case. Stakeholders and involved parties are identified and the basic elements of commercial / legal relationships with them included in the commercial options.

Once a potentially-feasible scheme can be envisaged, a decision is made as to whether to proceed to define the scheme and move towards a formal agreement to implement it. If there is no technically / financially feasible scheme, the process is halted. If there is, work now proceeds to the Definition Phase.

4.1.2 Guidance

The T1010 03 Good Practice Guide provides a brief explanation of the expected lifecycle of an RCM project. This is a useful introduction to what tasks are involved in an RCM scheme, and what sort of things will need to be considered. The guide is structured around the project stages of planning, implementation and operation.

More detailed, structured and role-based guidance for planning an XI RCM scheme is available in the *Process page and diagram* (page 13), around which this suite of guidance is structured.

CHAPTER 5

Model the Business Case

This page contains all the guidance, or links to guidance, that may be needed by someone modelling a potential Cross-industry RCM business case.

5.1 Context

See *Introduction to Cross-industry RCM* (page 3) for a lay-person's guide to Cross-industry RCM.

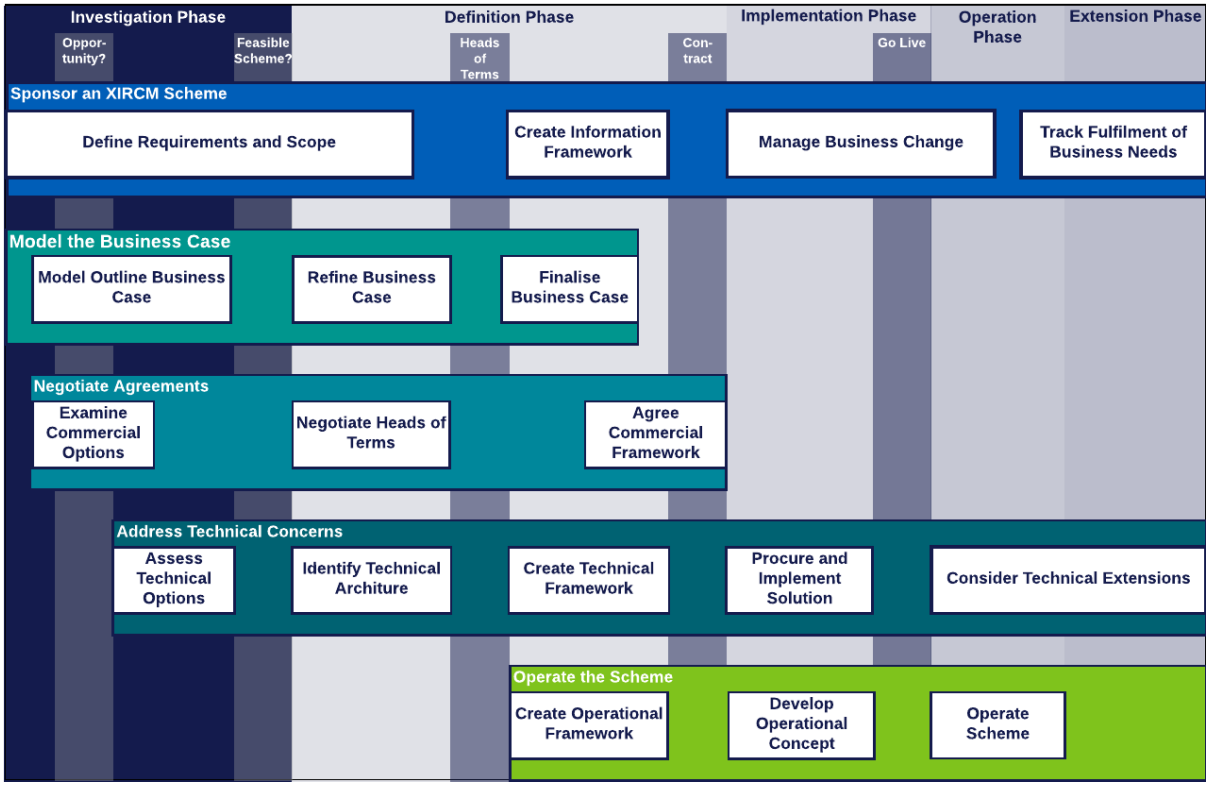


Figure 1: Cross-industry RCM Project Process Map - Overview

The figure *Cross-industry RCM Project Process Map - Overview* (page 26) shows the Model the Business Case workstream in the context of the project phases and gateways and the other workstreams. More detail on these

aspects of a Cross-industry RCM project can be found in *Process Map* (page 13) and the pages devoted to the other workstreams.

5.2 Tasks

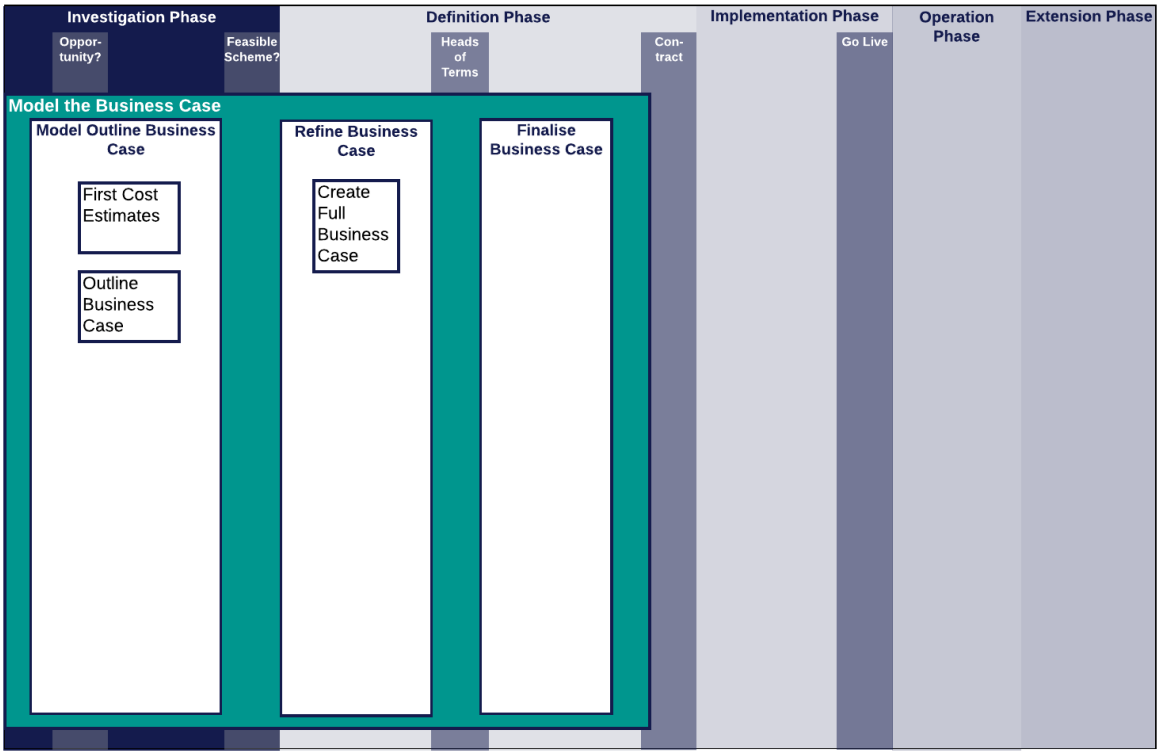


Figure 2: Model the Business Case Workstream

5.2.1 Model the Outline Business Case

This task is carried out near the start of the project during the Investigation phase, to see if there is the possibility of a business case for the project. It may involve the use of *The Business Case Tool* (page 27).

5.2.2 Refine Business Case

This task is carried out in the early part of the Definition phase. *The Business Case Tool* (page 27) is intended for use at the level of detail required at this time. It enables the overall case for the project to be examined, optioneering to be done and the outline of the relative costs and benefits to the involved parties to be identified.

5.2.3 Finalise Business Case

This task involves compiling a Business Case which fulfils all the requirements of the Sponsor's organisations project control process. It may require the use of other planning tools, but the data compiled for input to the Business Case Tool and its interim results will be useful inputs.

See also:

5.2.3.1 The Business Case Tool

5.2.3.1.1 Overview of the Tool

The Business Case Tool is an Excel spreadsheet model to help calculate an outline business case for a Cross-industry RCM scheme. It is part of the *T1010* toolkit, created as part of the T1010-04 workstream.

It takes inputs as follows:

- the assets being monitored in the scheme - as a count or number of miles of, e.g., track
- the stakeholders involved in the scheme
- current risks and costs associated with the assets being monitored as part of the scheme, categorised by safety (lives / property at risk), service (costs of delays and cancellations) and other (any type of additional risk or cost that might be affected by the scheme), for each stakeholder
- asset management costs: current rates of failure, time / cost per fault; maintenance costs (time-based and/or usage-based)
- current and future RCM costs, estimated at "maturity levels" corresponding to *ISO 13374*
- the impact of the Cross-industry RCM scheme, in terms of the move from lower to higher ISO 13374 maturity levels, with timings of their implementation
- the benefits of the scheme, in terms of percentage changes to the costs previously entered by category and stakeholder.

From these, the model calculates a set of outputs which give an estimate of the overall potential value of the scheme and the distribution of costs / benefits between the various stakeholders:

- an overall summary by maturity level
- a year-by-year breakdown
- a cash flow forecast.

5.2.3.1.2 Tool and Documentation

You can download the template spreadsheet from here: `Business Case Tool Model`.

The user guide for the tool is available here: Business Case Tool User Guide. This gives step-by-step guidance in using the tool.

The theory behind Cross-industry RCM business cases and the working of the model is explained in the T1010-04 report here: Business Case Guidance.

5.2.3.1.3 Worked Example

As part of the preliminary work for *The IMPRCM Project* (page 6), the model was used to create a worked example based on the data interchange envisaged for that pilot project: track geometry data captured by Crossrail trains about Network Rail Western Route infrastructure between Paddington and Reading.

The spreadsheet file for the worked example is available for download: `Business Case Tool Worked Example`

The worked example used the assumptions documented in Table 1 below. The table is based on the pages of the model's "Wizard", which takes the user step by step through making inputs to the model. The Wizard includes guidance and tips, but refer to ref:*Tools and documentation<tools-and-documentation>* for a fuller set of instructions.

Table 1: Worked Example

Wizard Stage	Purpose	Detail
Start the Wizard	The Wizard will step through the inputs into the model with helpful comments	By opening the model, and clicking the "Menu" button from any page, then "Wizard" button.
Summary Description	Enter a short description of the purpose of the model	The summary for our example was "Use of track condition information gathered by Crossrail trains fitted with measuring equipment will improve proactive maintenance of NR track"
Sensitivity Config	This will set the "Upper" and "Lower" bounds of entries in the rest of the wizard automatically, though they can always be overwritten.	
Asset General Definition	Set the name of the asset, parties names and general financial basis of the model.	The four parties in our scheme were Crossrail, NR Western, Balfour Beatty and RSSB.
Implementation Timescales	This step contains several of the RCM processing levels from *ISO 13374* (page 7). Enter the numbers of years it will take each stage to reach maturity.	In our worked example we assumed two years to realise the benefits of *State Detection* (page 10) a further two each for the *Health Assessment* (page 10) and *Prognostic Assessment* (page 10) stages, and a final three years to achieve *Advisory Generation* (page 11). So achieving the end result of an advisory generation scheme based on this RCM initiative would take 11 years.
Asset Information	There is a page to enter Asset Information for each of the parties. Upper and Lower bounds for each entry are calcualted automatically, and a "Expected" figure can also be used.	We entered data for NR Western only, with Infrastructure Assets Count of "60" indicating 60 miles of track. We also entered the average utilsation of each mile of track per year (1,000,000)
Asset Risk	Again, there is a page for entering the risks associated with the asset monitoring, for each party. Fill in at least one of the three categories of Safety, Performance and Other Risk.	We showed Performance Risk for Crossrail as follows: 20 incidents a year, each causing 100 minutes of delay at a cost of £25 per minute, and with a cancellation rate of 1%, with each cancellation costing £10,000. We also entered a rate of 0.1 major incidents per year, at a cost of £50,000 under "Other Impacts".
		For NR Western we set the Performance Risk at 50 incidents a year, each causing 100 minutes of delay at £50 per minute. We also entered a rate of 0.3 major incidents per year, at a cost of £100,000 per incident.
Asset Whole Life Cost (WLC)	These pages define the current costs of reactive and planned maintenance for the asset being monitored, for each party	Here we defined a one-year maintenance schedule and costs for inspection, tamping and relaying for NR
RCM Current & Future Costs	In these pages you can specify up to four systems to compare; one per page of the wizard. Enter the installation and maintenance costs and responsibility % for the proposed RCM scheme/s. "Future" refers to future site mix rather than to a post RCM implementation costs scenario.	On the first page we specified the on-train costs of a UTGM scheme. We envisaged one fault per site each year, at a cost pf £2,000 to Crossrail, and centralised operating costs of £15,000 per year to NR.
		We used the next page to specify the costs additional costs relating to a data

CHAPTER 6

Negotiate agreements with stakeholders

This page contains an overview of the process, documents and guidance that may be needed by someone working towards commercial agreements required for a Cross-industry RCM scheme.

6.1 Context

See *Introduction to Cross-industry RCM* (page 3) for a lay-person's guide to Cross-industry RCM.

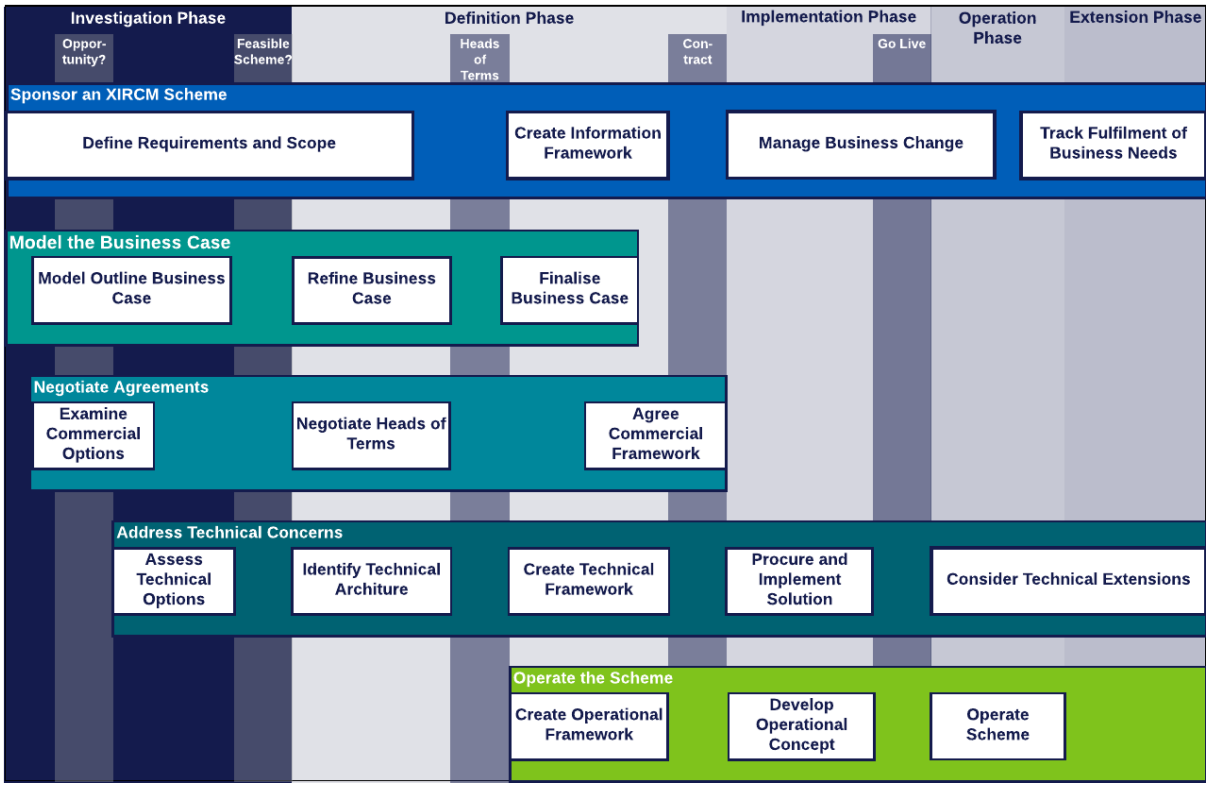


Figure 1: Cross-industry RCM Project Process Map - Overview

Figure 1 shows the Negotiate Agreements workstream in the context of the project phases and gateways and the

other workstreams. More detail on these aspects of a Cross-industry RCM project can be found in *Process Map* (page 13) and the pages devoted to the other workstreams.

6.2 Tasks

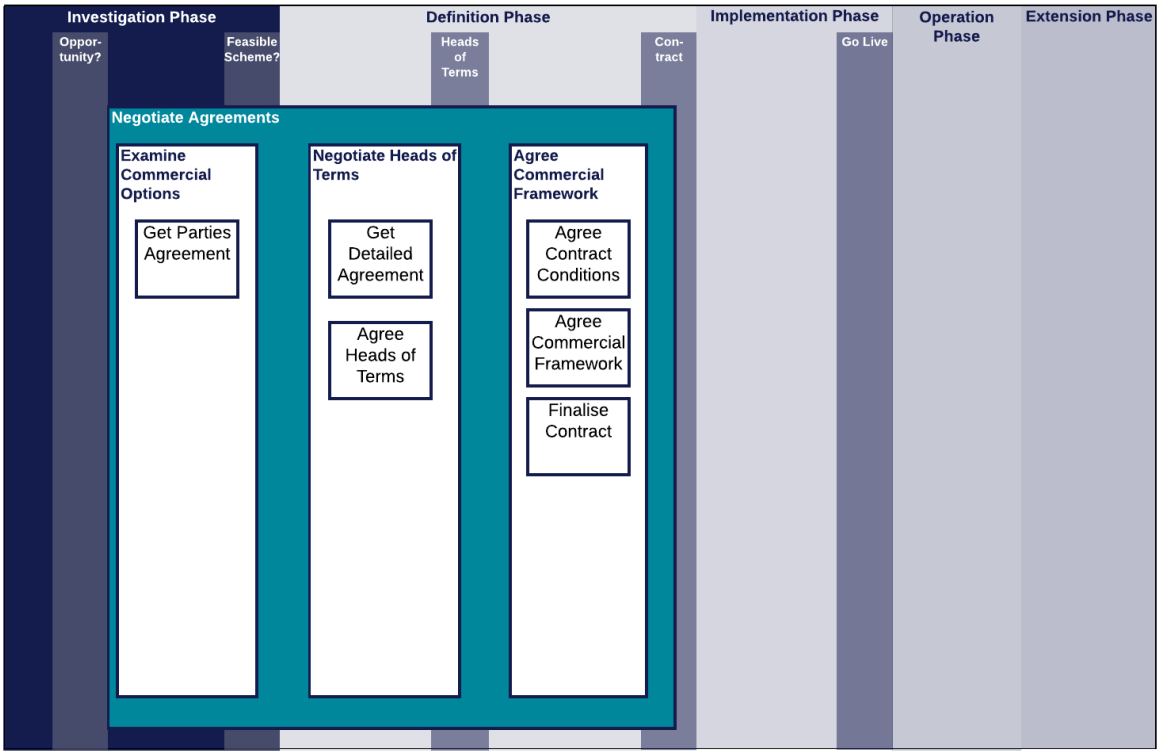


Figure 2: Negotiate Agreements Workstream

6.2.1 Examine Commercial Options

This initial task is a brainstorming / scoping effort carried out to see if there is a viable commercial framework available. It is dealt with in the *Investigate a Cross-industry RCM opportunity* (page 24) page.

6.2.2 Negotiate Heads of Terms

This task is carried out early in the Definition phase. The Heads of Terms is a non-binding understanding about the commercial principles that will be embodied in an agreement.

6.2.3 Agree Commercial Framework

This task represents all the work that needs to be done to reach a full contractual agreement between parties.

6.3 Guidance

T1010 provides guidance and templates to support the tasks above. The documents list the decisions that need to be taken at the Heads of Terms and Contract stages and give some help on to how to complete them.

6.3.1 Commercial Decisions Matrix

This is a *matrix of all commercial considerations* (page 40) to determine in outline the decisions required for an XI RCM scheme as it progresses towards contractual agreement. For each scheme, items on the matrix may or

may not be required for the heads of terms, and for the full contract. Items can be marked as required or not, responsibility for each item can be assigned, and decisions can be recorded in the matrix itself. The Commercial Guidance page provides guidance for all of these items.

Download the template from here: `Commercial Decisions Matrix Template`

6.3.2 Commercial Guidance

The *commercial guidance page* (page 33) contains guidance for all of the decisions listed in the Commercial Decisions matrix at the outset of work towards commercial agreement.

6.3.3 Heads of Terms

The heads of terms to the scheme will be formed by the answers in the *Commercial Considerations Matrix* (page 40). A word template for a Heads of Terms document was developed for use with the IMPRCM case study, and can be downloaded here.

Download the template from here: `Heads of Terms Template`

6.3.4 Full Contract and Schedules

A template contract and schedules for a Cross-industry RCM scheme is available to use. This is the output of previous research project T1010 02 - Appendix E to the report. Note the principles linked to below; the template is not intended to be a fully developed agreement but includes modules covering a range of alternative approaches that can be chosen by the parties to suit different types of RCM activity and different cross-industry relationships.

Download the template from here: `Contract and Schedules Template`

6.3.5 Principles of use for T1010 02 template contract

These commercial principles, taken from the report from the T1010 02 project, explain how the template should be used: `T1010 02 Report`

- The templates were not intended to be fully developed agreements but would include modules covering a range of alternative approaches that can be chosen by the parties to suit different types of RCM activity and different cross-industry relationships
- The templates were planned to target agreements between parties interested in common endeavour (and thus appropriate for immediate application to current initiatives) and not the 'cloud of data' concept where other parties can 'reach in' (a future ambition)
- An overall scheme may involve multiple parties with a single agreement or, more likely, multiple agreements some of which will be bilateral supply or service agreements. The templates are intended to have multiple uses
- Where purely commercial supply or service agreements are based on existing company terms of supply then certain of the template modules may still be used in schedules to ensure obligations still flow back-to-back between the parties
- At an early stage preliminary Heads of Terms might be agreed pending full development of the concept and the business case (which could themselves require considerable inputs)

6.3.6 Supporting Documents for the template contract

Other outputs from T1010 02 are available for reference.

- `T1010 02 Report` - including Commercial Principles, Key Features & Recommendations.
- `T1010 Appendix F` - the "How To" for the contract document itself.

CHAPTER 7

Commercial Guidance

This page contains guidance for working towards *Heads of Terms* (page 32) and *Full Contract and Schedules* (page 32).

It comprises:

- a description of each decision required in the *Commercial Considerations* (page 40) template which will form the basis of an agreement
- a link to the source material from the T1010 project
- any relevant additional reading

7.1 How to use the Matrix of Commercial Considerations template

The *Commercial Considerations* (page 40) template and this page of guidance is intended to be used at the outset of a scheme to provide the reader with an understanding of the scale and scope of required work.

The exact requirements for a contract, and to some extent for a heads of terms, will vary from scheme to scheme, and this matrix allows the user to select the relevant items, assign a responsibility and fill in any decisions or comments. It should be therefore be noted here that the template contract and schedules, created in the T1010 research project, are intended to be used as a **starting point** for the final contract, and there is no requirement for the final agreement to correspond exactly to the template.

The remainder of this page comprises guidance for the decisions required in the matrix.

7.2 Objectives

> Precursor to Heads of Terms. There is no requirement for the objectives to be entered into the contract however a full understanding of the aims and objectives of all parties should be ensured at the outset of discussions.

A preamble is needed setting out the overall aims of the scheme. There may be multiple aims for different parts of the industry which need to be referred to, whether achieving maintenance efficiency gains, better performance or other improvements. The intention to achieve these aims by common endeavour of the parties to the agreement needs to be expressed. If there is to be an 'alliance' relationship with a risk/reward sharing mechanism this should also be mentioned.[1]

[1] Source: T1010-02 Commercial Principle 2 (Objectives)

7.3 Parties and Governance

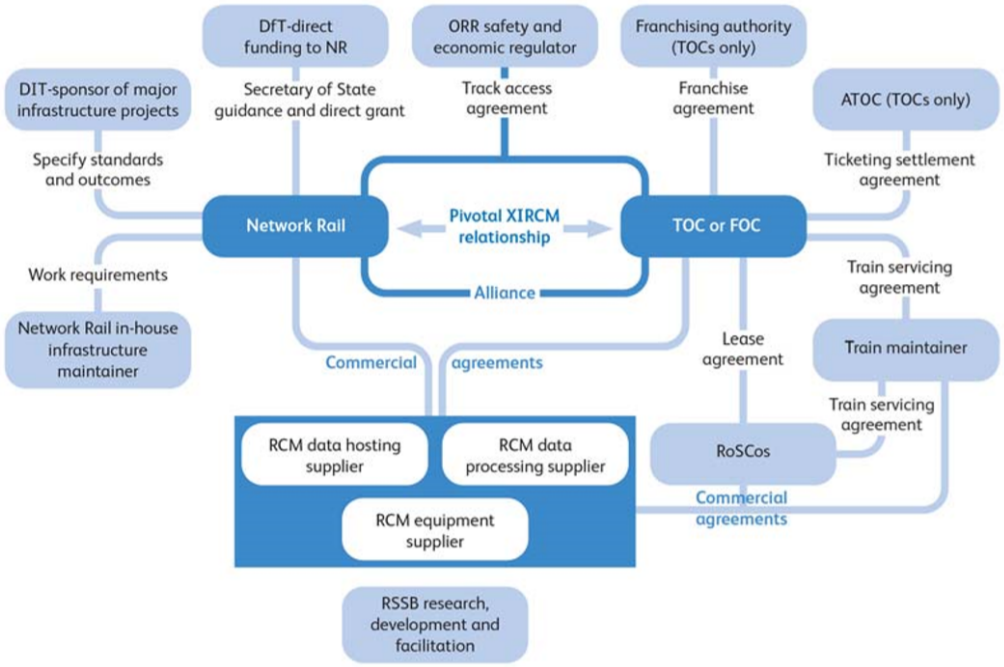
> There is no requirement for the information on parties to be entered into the contract however a full understanding of the responsibilities and requirements of those parties involved should be ensured at the outset of discussions.

For a data collection and sharing scheme to be successful there needs to be clarity on who is the promoter, who is supporting that and the roles and responsibilities of the different parties. Roles identified include:

- Scheme Lead or Joint Scheme Leads (scheme promoter(s), who can be expected also to have other roles listed below)
- Suppliers: Equipment and software suppliers/installers/maintainers
- Data collectors/ hosters/ processors/ analysts/ distributors
- Data receivers/ users/ beneficiaries
- Scheme facilitators

The organisations fulfilling these roles can include Network Rail, other infrastructure owners such as HS1, TOCs, FOCs, ROSCOs, train manufacturers, train maintainers, monitoring equipment manufacturers, software houses, data hosting firms and technical consultants. Scheme facilitators can also include governmental, regulatory and industry bodies with contractual of regulatory relationships with any of the above.

Figure 1, showing interested parties within the industry and their relationships, illustrates this item.



Where there are Joint Scheme Leads, provision may be needed for establishing a governance forum (management group) for scheme oversight including details of how decisions are to be taken. If an alliance is involved this forum would have oversight of the operation of the risk/reward mechanism.[2]

7.4 Business Case

> This information would not be required in the final contract however could be stated in the heads of terms, if necessary.

[2] Source: T1010-02 Commercial Principle 1 (Organisation and Parties)

Where a scheme business case has not been completed prior to an agreement (or more likely a Heads of Terms agreement) being reached between the parties, the agreement may need to include reference to responsibilities of the parties for its completion. This may be at industry level and parallel business cases at the level of the individual parties involved (allocating contributions and benefits to them) will be an enabler for agreement to be reached.[3]

See also:

Guidance and templates (page 26) for the creation of an outline business case.

7.5 Programme

Programme can be recorded in the heads of terms and in schedule 3 of the full contract.

This should include programme for the supply and installation of equipment, and may include timescales for the supply or transmission of any data required. It may also include provisions relating to the schedule for monitoring condition of assets, for example regarding train-borne equipment, where particular units will be located, which areas of track this will monitor and over what timescale.[4]

7.6 Equipment

Equipment specifications can be recorded in the heads of terms and in schedule 2 of the full contract.

Equipment should be specified in the following ways.

7.6.1 Ownership and Use

- Installation- what is to be installed, where and by whom.
- Who owns the installed equipment? Schedule 2 broken down to be more specific
- Who has permission to use the equipment? Schedule 2 broken down to be more specific
- Are there are any conditions relating to the removal or decommissioning of the equipment? Schedule 2 broken down to be more specific[5]

7.6.2 Operation, maintenance and safety

- Specify the tasks required to accept the system, undertake system safety assessments, put the system into operation and keep the system operational (day to day system management and maintenance, including managing equipment servicing and warranties)
- Specify which parties are responsible for managing and undertaking these tasks
- Specify which parties are responsible for paying for these tasks, how much and when, including allocation of operating cost and other O&M risks to the parties best placed to mitigate and manage those risks.
- Specify any liabilities if any Equipment fails and has knock-on repercussions[6]

7.6.3 Other

- Are any approvals required? Who will be responsible for gaining these, and by when? (Eg from bodies such ROSCOS, DfT, ORR)
- Should any industry standards apply to the equipment installation or operation?
- Is any asset protection agreement or similar arrangement required?[7]

[3] Source: T1010-02 Commercial Principle 4 (Business Case)
[4] Source: T1010-02 Contract Template Schedule 3 (Programme)
[5] Source: T1010-02 Contract Template Schedule 2 (Specification)
[6] Source: T1010-02 Commercial Principle 6 (System Operation, Maintenance)
[7] Source: T1010-02 Contract Template Schedule 2 (Specification) and T1010-02 Commercial Principle 5 (Initial Investment, Renewals)

7.7 Data

Data specifications can be recorded in the heads of terms and in schedule 2 of the full contract.

Details are needed of:

- the data required, including any metadata.
- data collection - by whom, with what frequency
- The steps required in the information lifecycle (including initial processing, cleansing, normalising and transmission) and who is responsible for each task
- Data provision - by whom, to whom, in what format, with what frequency and in what time
- Data storage - by whom, for how long, under what terms ? .. todo:: what does records of the collected data mean? See schedule 2
- Levels of precision and accuracy
- Which parties are responsible for undertaking each of the data handling tasks
- Which parties pay for each of the data handling tasks, how much and when, including allocation of data handling cost/other data handling risks[8]

7.8 Data Uses, Sharing, IP

Use of data and IPR can be defined in the Heads of Terms and in schedule 4 of the contract. Data sharing can be defined in the Heads of Terms and in schedule 2 of the contract.

7.8.1 Data Use and Sharing

Details are required of:

- the intended ownership of data and of value added information at each stage made available post-processing.
- the expected use of the data at each stage by each party
- what data is to be shared with which parties and where ownership is transferred
- how data is to be shared and any protocols relating to data sharing[9]

See also:

Also see the *Principles* (page 37) section below.

7.8.2 IP (Intellectual Property) Rights

Details are required of title to data and licensing of IPR agreements.

Title to data - three options are described in schedule 4 of the contract.

OPTION 1 – to be used where a particular Party retains all IPR to the Data as it is recorded – perhaps with licence to another Party for various purposes (see below)]

OPTION 2 – title to the original Data remains vested in a particular Party, but another Party has title to any modifications or improvements or processed output made in respect of the Data]

OPTION 3 – title to IPR developed or created in the course of the Agreement, to be owned by the party that developed or created the IPR

Licensing of IPR - two options are described in schedule 4 of the contract, for the granting of a licence of IPR for a permitted purpose.

[8] Source: T1010-02 Contract Template Schedule 2 (Specification) and T1010-02 Commercial Principle 7 (Data Collection, Processing, Hosting, Transmission)

[9] Source: T1010-02 Commercial Principle 8 (Data Ownership (IPR), Sharing, Usage and Confidentiality (NDAs))

OPTION 1 - Permitted Purpose means for any and all purposes, including to improve and/or modify such Intellectual Property Rights.

OPTION 2 - Limited license to use the Data for a particular purpose only[10]

7.8.3 Principles

The principles suggested to be put in place for this are:

- Raw data:
  - should remain the intellectual property of the party for whom it is collected (that is normally the organisation responsible for the assets being monitored, whether it is a TOC, FOC, manufacturer, maintainer or infrastructure owner), even though it is considered that raw data should generally be shared freely within the industry given appropriate licensing conditions (see below)
  - may not be ascribed value other than the allocated cost of collecting it
- Processed information:
  - should become the intellectual property of the party for whom the data is processed (in whom IPR should rest)
  - may be ascribed value in addition to the allocated cost of processing where it enables demonstrable net savings or value added, so leading to net industry costs being reduced
- Data sharing:
  - the principle proposed is that data ownership is not transferred but that the sharing is by means of a licence to receive and use the data for specified purposes (to be described) and subject to non-disclosure agreement (NDA) terms to be set by the owner.
  - any further processing by a party with whom the data has been shared, and any subsequent sharing with further parties, should then be subject to the same conditions and limitations.[11]

7.9 Payments

> The principles should be recorded in the Heads of Terms. The amounts, payment arrangements and other terms should be defined in schedule 5 of the contract.

7.9.1 Principles of Charges

Agree the principles of charging for the Heads of Terms:

- who is going to pay whom for what
- on what basis are the costs for raw / processed / augmented data to be decided?
- what payments are required for equipment supply and installation
- what payments are required for equipment maintenance and operational - either time based or on another basis
- is payment required for software development

7.9.2 Details of Charges

For the full contract agree the amounts and timeframes for charging for:

- Data
- equipment supply and Installation
- equipment maintenance and operation

[10] Source: T1010-02 Contract Template Schedule 4 (Data Ownership and IP)
[11] Source: T1010-02 Commercial Principle 8 (Data Ownership (IPR), Sharing, Usage and Confidentiality (NDAs))

- software

Agree allocation of risk

Agree terms on general items - refunds, invoicing, terms[12]

7.10 Service Level Agreements

> Service level agreements should be considered in the heads of terms and fully defined in Schedule 6 of the contract.

Details are required of the levels of service agreed between the service providers and the users or beneficiaries. These are likely to relate principally to data or processing. The extent of detail required is most likely to be proportional to the size, scale and importance of the project and the level of investment/payment involved. If the size of the agreement is small or the Parties are not paying each other for data, there may not be an SLA.

Recommended items for consideration within a service level agreement:

- Availability
- Timeliness (or frequency)
- Quality (integrity, precision, accuracy)
- Transfer dependability
- Security
- Fault tolerance
- Response times

In addition, it may be advisable to define what should happen if anything goes wrong:

- Steps to be taken in response to any service delivery issues
- Escalation procedures
- Compensation for downtime
- Disaster recovery

Parties should also consider whether a performance regime may be appropriate, and what remedies there may need to be for any failure to meet the Specification – for example, liquidated damages, increased monitoring, remedial plans, liability caps for breach etc.[13]

7.11 Insurance and Warranty

> Any insurance or warranty arrangements should be defined in Schedule 7 of the contract.

As a principle it is not generally accepted that data that is shared by an owner with another party should carry any level of warranty. This is particularly the case with raw data. It is believed that the only duty on the owner is to be open with the receiving party about the source of the data and methods used to collect it. The view is taken that parties with responsibilities in the rail industry are sufficiently knowledgeable to assess the data and the risks around it and to take responsibility for any use to which they put it.

When the Commercial Principles were put in place for the creation of the template contract, there was] no information on insurances being employed specific to Cross-industry RCM schemes but provision could be made for clauses adaptable to any perceived needs. [Similarly there was] no information on limitations on liability being applied in Cross-industry RCM schemes but these are expected to be in place on many commercial contracts and provision can be made for suitable clauses to be made available.

[12] Source: T1010-02 Contract Template Schedule 5 (Payments)

[13] Source: T1010-02 Commercial Principle 9 (Service Level Agreements) and T1010-02 Contract Template Schedule 6 (Service Level Agreement)

The contract schedule 7 allows for the following liabilities to be entered:[14]

- Product
- Third party
- Employers

7.12 Alliance Risk / Reward

> Any targets or risk allocation relating to a risk / reward arrangement can be entered in Schedule 1 of the contract.

While in usual rail industry business, commercial agreements will set out the services being supplied and costs and risk allocations, there are good reasons for more collaborative commercial arrangements to be entered into in certain circumstances. Examples include the Wessex alliance, bringing SWT and NR Wessex Route together in a shared risk and reward model intended to drive up such things as efficiencies and performance.

In such arrangements, the targets are set out, as are how to share the risks and rewards of achieving, exceeding or underperforming those targets. Cross-industry RCM schemes, with clear aims for outcome improvement could also be incentivised in this way. Provision can be made in the template agreements for clauses that would enable alliance arrangements (unspecified) to be covered.[15]

7.13 Term & Franchise End

> Any terms should be defined in Schedule 1 of the contract.

Depending on the nature of the scheme and its intended longevity, it may be desirable to identify with the franchise authority arrangements to allow carryover of a scheme into the next franchise. It has been suggested that this might be overcome by having the equipment provision added to the requirements of the track access agreement or train service agreement - this making it feature in the data room when a franchise competition occurs.[16]

7.14 Disputes and Termination

> Any terms should be defined in Schedule 1 of the contract.

Suitable arrangements for resolving disputes and for termination, consistent with industry norms, will be needed along with specific provisions for decommissioning equipment.[17]

[14] Source: T1010-02 Commercial Principle 10 (Warranties, Insurances, Liability Caps) and T1010-02 Contract Template Schedule 7 (Insurance)
[15] Source: T1010-02 Commercial Principle 11 (Alliance Risk and Reward Arrangements)
[16] Source: T1010-02 Commercial Principle 12 (Term of Agreement, Franchise-end Provisions)
[17] Source: T1010-02 Commercial Principle 13 (Disputes, Termination, Decommissioning)

CHAPTER 8

Commercial Considerations

This table shows items of agreement that are likely to be required in a contract between two parties in a Cross-industry RCM scheme. Of course it may be that some items are not required in certain schemes and additional items may be required.

The Heads of Terms column indicates those items which are likely to be required for a Heads of Terms agreement.

A full description of how to approach these considerations can be found under *Commercial Guidance* (page 33).

Table 1: Commercial Considerations

Area	Item to agree	Heads Of Terms	Template Contract ref
Objectives (page 33)	Overall aims. Aims may be different according to party.	yes	Preamble
	Common endeavour or risk / reward mechanism involved?	yes	Preamble
Parties and Governance (page 34)	Identify roles including scheme lead / promoter	yes	n/a
	Hardware / software suppliers, installers, maintainers	yes	n/a
	Data collectors, processors and users	yes	n/a
	Scheme facilitators	yes	n/a
	Agree governance group meetings		Schedule 8
	Agree programme management - monitoring, risks, issues, decisions		Schedule 8
Business Case (page 34)	Who is responsible for completion	yes	n/a
Programme (page 35)	Timeframe for supply and installation of equipment	yes	Schedule 3
	Timeframe for supply of any data required	yes	Schedule 3
	Timeframes for monitoring of assets	yes	Schedule 3
Equipment (page 35)	Specify installation - what, where and by whom	yes	Schedule 2
	Who owns equipment	yes	Schedule 2
	Who may use equipment	yes	Schedule 2

Continued on next page

Table 1 – continued from previous page

Area	Item to agree	Heads Of Terms	Template Contract ref
	Are any approvals required? Eg from ROSCOS, DfT, ORR	yes	Schedule 2
	Should any standards apply?	yes	Schedule 2
	Operation, maintenance & safety - what tasks are required?	yes	Schedule 2
	Operation, maintenance & safety - who is responsible?	yes	Schedule 2
	Operation, maintenance & safety responsibilities -what are payment arrangements?	yes	Schedule 2
	Any specifications on equipment removal?	yes	Schedule 2
	Asset Protection agreement - define responsibilities	yes	Schedule 2
Data (page 36)	Data required - what (including metadata?), in what format, any other requirements such as open source	yes	Schedule 2
	Data acquisition - by whom, with what frequency	yes	Schedule 2
	Data lifecycle - processing steps, cleansing, transmission, archiving	yes	Schedule 2
	Data storage - by whom, for how long, under what terms?	yes	Schedule 2
	Levels of precision / accuracy required (to be formalised in SLA)		Schedule 2
	Tasks required for data lifecycle and who is responsible	yes	Schedule 2
	Costs associated with data lifecycle tasks and who is responsible.		Schedule 2
Data Uses, Sharing, IP (page 36)	Expected use by party	yes	Schedule 4
	Data sharing to be specified	yes	Schedule 2
	IPR - title to data option 1, 2 or 3, or title to licence option 1 or 2	yes	Schedule 4
Payments (page 37)	Agree charging principles - which parties are paying whom for what, in relation to:	yes	
	Equipment supply and installation	yes	Schedule 5
	Equipment maintenance and operational (time based or other)	yes	Schedule 5
	Data	yes	Schedule 5
	Software development	yes	Schedule 5
	Define amounts and timeframes for:	yes	
	Equipment supply and installation	yes	Schedule 5
	Equipment maintenance and operational (time based or other)	yes	Schedule 5
	data	yes	Schedule 5
	software development		Schedule 5
	General - refunds, invoicing, terms		Schedule 5
Service Level Agreements (page 38)	Availability	yes	Schedule 6
	Timeliness	yes	Schedule 6
	Quality (integrity, precision, accuracy)	yes	Schedule 6
	Transfer dependability		Schedule 6
	Security		Schedule 6

Continued on next page

Table 1 – continued from previous page

Area	Item to agree	Heads Of Terms	Template Contract ref
	Fault tolerance		Schedule 6
	Response times		Schedule 6
	Steps taken re service delivery issues		Schedule 6
	Escalation procedure		Schedule 6
	Compensation for downtime		Schedule 6
	Disaster recovery		Schedule 6
	Performance regime?		Schedule 6
Insurance and Warranty (page 38)	Product liability - not less than and aggregate / annual maximum		Schedule 7
	3rd party liability - not less than and aggregate / annual maximum		Schedule 7
	Employers' liability - not less than and aggregate / annual maximum		Schedule 7
Alliance Risk / Reward (page 39)	Set out targets	yes	Schedule 1
	How to share risk		Schedule 1
	How to share reward of underperforming, achieiving or exceeding target		Schedule 1
Term & Franchise End (page 39)	Likely to extend past franchise end?		Schedule 1
Disputes and Termination (page 39)	Specify any alternative to clause 15	yes	Clause 15 / Schedule 1
	Who can voluntarily terminate		Schedule 1
	Notice period		Schedule 1
	Removal of equipment - who responsible and who pays		Schedule 1
	Any clauses survive termination?		Schedule 1
	Basis and payment of costs to be paid by defaulting party		Schedule 1

CHAPTER 9

T1010 02 Appendix E Contract Template and Schedules

This page represents the contract template itself.

Will be held as a PDF file. Whether or not this header page continues to exist tbc. May be of use to have this and other header pages retained to provide overview of doc?

CHAPTER 10

How to use the Contract

This page to hold links to and explanation of how to use the template contract and schedules.

It is a how to use the document and not a how to make decisions etc

RCM Contract Template and Schedules (T1010 02 Appendix E)

How to use the template contract (T1010 02 Appendix F)

CHAPTER 11

Information Framework

11.1 Introduction

The Information Framework defines the data requirements of the scheme in *business* terms - i.e. what is needed for the scheme to satisfy its business needs.

It is a key output of the *Investigation Phase* (page 14) and *Definition Phase* (page 14) of the scheme.

It defines the *"what"* element of the information flows. The *"how"* element is then defined by the *Technical Framework* (page 54) and *Operational Framework* (page 55) and the schedules in the contract dealing with scope and service level agreements.

The *T1010* documentation provides guidance on aspects of the Information Framework:

- it sets out some *recommended principles* (page 6) for cross-industry data interchange to simplify and clarify legal and technical issues and improve re-usability of the data.
- The T1010-02 business process maps specify the content of the document and refer to the *ISO 13374* data / processing model and maturity levels.[1]
- The T1010-01 data architecture has requirements and recommendations about relevant data standards, data item formats, transfer methods, identification and keying, metadata and governance.[2]

Relevant parts of that guidance are included in the text below.

Also included are tips which reflect lessons learned during the development of the prototype data interchange via broker in *The IMPRCM Project* (page 6).

11.2 Framework Content

11.2.1 Data Items

The Information Framework needs to define the data items to be included in the data flow. This should be done in the form of a *schema* listing the data items and defining for each one:

- **name** - the name of the data item. This should conform to a naming convention which makes the name usable in as many contexts as possible. Guidance: *Naming conventions* (page 49).
- **title** - a descriptive title of the data item. This is a short phrase defining what the data item is.

[1] Source: T1010-02 Process Map - C
[2] Source: T1010-01 Data Architecture Requirements

- **description** - a lengthier description of the data item containing contextual information such as source and meaning.
- **data type** - the type of the data item: numeric, string, date/time etc. Should be defined as specifically as possible. Guidance: *Data item types* (page 49).
- **format** - the format of the data item. Particularly important for items such as dates / times / geographical positions / track positions where standardisation is useful. Guidance: *Data item formats* (page 49).
- **size** - the size of the data item in characters or bytes (if relevant). This is important for correct handling of textual ("string") data and for the calculation of data volumes (see *Data Volumes* (page 47)). Guidance: *Data item formats* (page 49).
- **unit of measure** - the engineering units in which the data item is measured. Should be *SI* units where available, otherwise recognised rail industry standard units. Guidance: *Engineering units* (page 50).
- **uniqueness** - is the value of the data item unique - is it an identifier of, e.g. an asset?. Guidance: *Uniqueness and mandatory/optional* (page 51).
- **mandatory/optional** - must the data item always be present. Guidance: *Uniqueness and mandatory/optional* (page 51).
- **grain** - the level of detail of the data item. This typically relates to how frequently it is captured, or whether it appears at the level of "header" or "detail". This impacts the data volumes and storage - see *Data Volumes* (page 47); guidance: *Grain* (page 51).
- **reference** - source or reference for the data item - such as any parent or shared schema that it belongs to.
- **notes** - any further notes or guidance to help providers or users of the data item.

The data items will fall into groups:

- the *RCM* **metrics** being gathered or results of data processing
- identifiers of the **measuring equipment or processing agent** responsible for the data
- identifiers of the **asset being monitored**
- **time / date stamps** for the capture and processing of the data
- **contextual** data, for example about the measuring equipment, the surroundings, the train being operated etc
- **metadata** about the data and processing: file names, calibration data, ownership / licensing information, data quality indicators.
- **non-standard** data that may be added by equipment vendors or software suppliers.

It is important to identify which of the groups of data are required to meet the business goal. Most data flows will contain all the data groups.

It is likely that some of the groups will be common across different data flows and there may have been a schema already defined for them and shared. In this case the shared schema should be used as a starting point.

The T1010-01 data architecture mandates some aspects of the data content: these are discussed in *Data items* (page 49) below.

11.2.2 Data Flow Characteristics

To fulfil the goals of the scheme, the data flow will need to behave appropriately. The Information Framework should define what characteristics the flow should have. These include:

- **accuracy** - how close the readings or positions are to actual truth, within the stated precision
- **precision** - how precisely readings or positions are measured - e.g. 1 part per 1000, or to the nearest millisecond, or to the nearest 10cm
- **completeness** - how complete the data set needs to be; how tolerant the business process is of missing data; whether a complete record is needed or only the most recent reading(s)

- **consistency** - how well different aspects of the data stream agree with each other; whether the business process will tolerate repeated data or conflicting readings from different sensors
- **timeliness** - how up-to-date the data needs to be; how well the business process will tolerate data that arrives late; what sort of propagation time is acceptable from measurement of an event to the data about it being available
- **availability** - whether the data flow needs to occur 24 x 7 all days; how tolerant the business process is of an absence of data for a given number of minutes or hours per day or month
- **integrity** - how important it is to the business process that the data is not corrupted or observed en route
- **security** - how important it is that only authorised parties see the data or gain access to the equipment and processing steps involved in the data flow
- **openness** - how important it is that the data stream uses open standards and commonly-used formats
- **conditions of use** - what restrictions on the use of the data can be tolerated, directly or indirectly.

The T1010 commercial principles and data architecture mandate some of these characteristics: see *Data Characteristics* (page 52) below.

The more thought is put into defining these characteristics in the Information Framework, the more likely the technical solution addressed in the *Technical Framework* (page 54) will be correct.

11.2.3 Data Volumes

Data need to be stored after collection, during processing and for analysis purposes in most schemes. The type and size of data storage can have a large impact on the complexity and cost of the solution.

Two aspects of data volume need to be considered:

- **data in motion**: what is the rate of data generation / transmission? This governs bandwidth and processing capabilities of the solution.
- **data at rest**: how long does the data need to be stored for, at what level of detail? This governs data storage volume.

Detailed calculation of the sizing is a matter for the :ref:'technical-framework', but the inputs to that calculation are business-driven. These include:

- the raw rate of data generation by the scheme's sensors. This may be measured in data points per second, per event or per metre of distance covered
- the number of data items in each data record
- the steps in the processing pipeline (see *Data Processing Pipeline* (page 47)). Each step may need input, internal and output data stores; and may result in significant changes to data rates - whether reducing them by e.g. summarising or feature detection; or increasing them by adding contextual or calculated information.
- the length of time data needs to be stored for. This may be for the immediate needs of the scheme; it may also be for later analysis of historic data or to make the data available for other purposes.
- how many copies of the data need to be kept - say for backup or disaster recovery purposes.

11.2.4 Data Processing Pipeline

11.2.4.1 Overview

The Information Framework needs to define any processing steps needed to transform the data from its initial state to the form required by the scheme. Each such step needs the following to be considered:

- the *RCM* data coming in to the step: items, formats, volume, timing. This consistutes a data interface: see *Interfaces* (page 48).
- any contextual data required by the step to help it do its work: reference data, other data streams
- the processing being done by the step (see *Processing steps* (page 48))

- the performance requirements of the step: how quickly does it need to run; how often; how much data can be batched up.
- the data coming out of the step. Again, an interface: *Interfaces* (page 48).

T1010 recommends that pipelines are built in accordance with the principles set out in *ISO 13374*. This defines levels of processing (also referred to as "maturity levels" and a clear distinction between these and the data flows between them). See the guidance in *Processing Pipelines* (page 52).

11.2.4.2 Processing steps

Any pipeline will have some processing steps which are standard and may be shared by other similar pipelines; and some which are specific to their own application.

Steps which may be done in a standard way include:

- outlier detection and removal
- calibration (scaling of sensor or location data to correct for incorrect or drifting readings)
- quality identification (detection of invalid data due to e.g. sensor error or missing values)
- location referencing (establishing a track position from a geographical one)
- asset identification (including providing context data such as route section, train id etc)
- feature extraction (identifying events of interest from an othewise unremarkable data stream)
- aggregation (calculation of summaries such as averages, totals, standard deviations, exceedence counts)

Where the steps are standard, a module or algorithm may already be available and should be used in preference to creating a new one.

In a general case, different parties may be involved in carrying out different processing steps. This implies a possible need for a commercial / legal arrangement and a service level agreement with these parties: these should be included in the stakeholder matrix as described in *Define Requirements and Scope* (page 16).

The provider of a processing stem may retain *IP* rights in the algorithms or processes used.

11.2.4.3 Interfaces

T1010 requires that the interfaces into and out of processing steps are clearly documented and the data flows use standard data formats and interface methods. This is to ensure that the data travelling between steps is potentially accessible to other data users or for novel purposes; and to open the possibility of different algorithms, perhaps from different suppliers, to be used.

See the guidance in *Processing Pipelines* (page 52).

11.2.5 Ownership, IP, liability

The Information Framework needs to make clear who owns data and *IP* at each stage in the data lifecycle; and where responsibility lies for the provision of data and conformance to any expected quality characteristics or for decisions taken based on the data. This information is fed into the commercial terms for each agreement between parties.

T1010 has principles and guidance in this area, introduced in brief in *Key Principles of the T1010 Advice* (page 6) and detailed below in the guidance: *Ownership and IP* (page 53).

11.3 Guidance from T1010 / IMPRCM

These guidance notes highlight specific points taken from the *T1010* documentation relevant to the Information Framework. Where the IMPRCM project shed additional light, this is mentioned as a tip.

11.3.1 Data items

11.3.1.1 Naming conventions

Data item names should follow a consistent standard. They should be constructed in a way that makes them as easy as possible to re-use in different circumstances. This means:

- use a consistent case throughout - not mixed case. Lower case is generally more readable
- don't have spaces in data item names. Hyphens can cause trouble as well as they can be interpreted as minus signs in some contexts
- include the engineering units, where relevant, as part of the name so it is clear to users what they are

Tip: "snake_case" with word elements separated by underscores worked best in the IMPRCM prototype.

adapter column name	title	eng unit	data type	f
gps_longitude_deg	GPS longitued, decimal degress. +ve = East, -ve = West	deg	numeric	
speed_m_s	Speed, metres per second	ms^-1	numeric	
gps_ground_speed_m_s	GPS ground speed, metres per second	ms^-1	numeric	
eng_line_reference_rec	Engineer's Line Reference (ELR)		string	
track_id_geogis_rec	Track ID in GEOGIS format		string	
track_id_sect_appx_rec	Track ID in Sectional Appendix format		string	
distance_miles_decimal_rec	Elapsed distance in decimal miles	miles	numeric	
speed_as_recorded	Speed, as recorded		numeric	
speed_as_recorded_unit	Speed units, as recorded		string	
elapsed_distance_as_recorded_m	Elapsed distance as recorded, metres	m	numeric	
gps_acquisition_mode	GPS acquistion mode		string	
gps_detail_of_satellites	GPS satellite detail		string	
gps_dgps_station_id	GPS differential GPS station identifier		string	
gps_ground_speed_km_h	GPS ground speed, kilometres per hour	kmh^-1	numeric	
gps_height_of_geoid_m	GPS altitude	m	numeric	
gps_hdop_m	GPS horizontal dilution of precision	m	numeric	
gps_magnetic_track_made_good	GPS magnetic track made good	deg	numeric	
gps_number_of_satellites	GPS number of satellites in view		integer	
gps_pdop_m	GPS position dilution of precision	m	numeric	
gps_position_mode	GPS position mode		string	

Figure 1: Example from IMPRCM: column naming

Figure 1 shows an example from the master schema for the IMPRCM prototype broker.

11.3.1.2 Data item types

Where the type of data is known, it should be specified as precisely as possible. Some pointers:

- numeric data types should be defined as integer, float, decimal, hexadecimal or boolean as appropriate.
- where possible the size of the data item should be specified as well: integer or bigint; float or double-precision
- dates, times and timestamps should be identified as such, not described as character strings
- character strings should be given a maximum length

11.3.1.3 Data item formats

Standard representations exist for some data items which will greatly improve re-usability of the data if adopted. Data specifications should indicate these formats for clarity.

11.3.1.3.1 Dates and Times

Dates and times should be specified to be always formatted per *ISO 8601*. This means

- dates look like `20181226` or `2018-12-26`
- times look like `22:34`, `22:34:16` or `22:34.16.232451` (depending on precision)
- timezone must always be shown, either explictly with `+00:00` or `Z` for *GMT* / *UTC* or `+01:00` for *BST*.
- for preference, times should be shown in UTC using the `Z` specifier.
- timestamps look like `2018-12-26T22:34:16Z` or similar, with a `T` separating the date and time components.

11.3.1.3.2 Geographical locations

Latitudes and longitudes, such as GPS locations, should be specified to be formatted in a standard way. *ISO 6709* is mandated by T1010-01 as the standard representation for locations, but this is a permissive standard so the following restricted compliant approach is suggested:

- latitude should be shown in decimal degrees, like `51.4183`, with up to 7 decimal places of precision. (7 decimal places gives approximately 1cm resolution)
- positive latitudes are north; negative are south
- longitude should be shown in decimal degrees, like `-0.1365`, again with up to 7 decimal places of precision
- positive longitudes are east of Greenwich; negative are west.

11.3.1.3.3 Track Locations

Network Rail track locations are generally specified in Miles + Chains or Miles + Yards along an *ELR*, the miles being measured from mileposts. There is currently no Network Rail standard for representing these in data, so T1010-01 proposed one as follows:

- a distance in miles and chains should be shown as mmm:cc, miles not padded, chains padded to 2 places, e.g. `23:17` or `1:04`. The name of the data item should indicate that the distance is in miles and chains.
- a distance i miles and yards should be shown as mmm+yyyy with the yards padded to 4 places, e.g. `23+0374` or `1+0088`.
- track IDs should either be GEOGIS ones of the form `2100` or Sectional Appendix ones, of the form `DF`. The name of the data item should indicate which type of track id it is.

Warning: Work is under way to define a new method of defining track locations at Network Rail. The information in this section should be seen as provisional and likely to change. Track locations should be expressed in a Network Rail standard way, once defined.

11.3.1.4 Engineering units

Numerical items should be shown in *SI* units. Where there is also a standard but different way (e.g. speeds being shown in mph), a separate data item may be included to show that value.

Where a scaling factor is used, it should be clearly indicated in the data item name.

The data item name should indicate which unit is in use. So where a speed is recorded, it should always be shown in metres per second, but may also be shown in mph.

Tip: There is no single standard way for expressing the engineering units in textual form as part of the data item name. T1010-01 specifies the use of the *MIMOSA* list and, optionally, the *QUDT* ontology. Neither of these is particularly easy or clear to use. There is another attempted standard, *UCUM*, whose goal is to provide standards for naming units for interchange between computer systems, but doesn't give any clear way to represent, for

example, combination units like metres per second. For the IMPRCM data schemas, we used a simple standard like this:

Unit	Shown in name as
metres	_m
metres per second	_m_s
metres per second squared	_m_s2
miles / chains	_miles_chains
miles / yards	_miles_yards

11.3.1.5 Uniqueness and mandatory/optional

11.3.1.5.1 Unique identifiers

Some data items identify items such as the assets, railway locations, measuring equipment sets. Often, assets have many different real-world identifiers and there is scope for confusion. To simplify the handling of identifiers, T1010-01 follows the practice of *MIMOSA* in requiring every such thing to have a guaranteed unique identifier which is represented by a *UUID*. UUID-based identifiers should be included in the data specification.

11.3.1.5.2 Mandatory vs optional data items

T1010-01 generally takes the view that data items are optional - any data source or any data row from a source may leave it blank or show a null value. There are important exceptions, though, which should be indicated in data specifications:

- every data record must have a timestamp of the time of its creation or of any subsequent processing step
- the source equipment or software code of every data record must be identified by a *UUID*
- every railway asset must be identified by a *UUID*.

This is so that

- data from different sources can be merged together based on the identifiers of the things it describes
- a complete audit trail can be built up of the data's lifecycle.

T1010-01 also states that there should be a central service to issue UUIDs where they aren't available.

Tip: In the IMPRCM project, the data schema required UUIDs for the data files, the multiple unit on which the *UGMS* equipment was mounted, the timetabled train, the train run (train ID + date), data row and processing data adapter.

As part of the project, a UUID service was created, which could issue UUIDs and store the real-world identifier each one referred to. It then allowed the UUID to be looked up using the real-world identifier in future. This was a simple prototype of the service called "SDAIR" - Structured Digital Asset Interoperability Registry - proposed by *MIMOSA*.

It is quite easy to generate UUIDs - most computer languages have a library function to do it.

11.3.1.6 Grain

The grain of a data flow is determined originally by the data gathering equipment which will have a natural operation frequency or spatial resolution. Other grains may become important in the lifecycle:

- the data flow may be batched into a header / detail structure. The header has its own grain, which may be at the level of a file of data
- the data files may be batched up and transmitted once per train run or per day

- data processing algorithms may summarise the data into distance-grouped or time-grouped totals, averages or counts.

All the levels of grain at which the data are of interest should be noted in the information framework.

Tip: In the IMPRCM project, the raw data rows were at the grain of one row per 0.2m of track covered. Since little actual sample data was available, a test data generator was built which could simulate different ways of batching the data, all the way from emitting the rows in quasi-real time, through to batching them up into files representing the whole of a days's operation of a UGMS unit.

The project also built code modules to aggregate the data to the level of each *S&C* unit on the route section of interest, and to each train run over each such unit. This, of course, greatly reduces the volume of data being handled by the downstream processing.

11.3.2 Schemas

The data schema represents all the data items in the data flow and records all the characteristics noted in the sections above. T1010-01 goes into some detail about how schemas should be managed, most of it technical and thus covered in the *Technical Framework* (page 54).

An important aspect for the Information Framework is to note where schemas already exist for particular types of data. These should be used in preference to defining new schemas, as they help with standardisation and can reduce the cost of compliance. T1010-01 suggests a shared industry repository of these schemas and a variety of methods of making sure that they remain usable and can evolve.

The schema set up in the Information Framework has an important role in the *Technical Framework* (page 54), as it forms the core of the definition of the actual database structures and interface formats needed to store and move the data flow.

Tip: In the IMPRCM project, an Excel spreadsheet was set up to hold the master data schema with all the data items involved in the data interchange. Columns in the spreadsheet indicated where each data item was used in the data flow: this meant that the master schema could also be used to create sub-schemas for the interfaces between elements of the data processing pipeline.

This spreadsheet can be downloaded for inspection as an example: `IMPRCM Master Schema`.

11.3.3 Data Characteristics

T1010 does not mandate or recommend any particular set of data quality attributes or levels for accuracy, precision, completeness, consistency, timeliness or availability: these depend on the particular data flow and the business problem it addresses.

For **integrity**, T1010-01 specifies some standards that can be mandated to safeguard the data in transit. These include the use of checksums and encryption.

For **security**, T1010-01 specifies a number of methods of access control and restriction of access by user type. These should be defined in the Information Framework if have particular importance to the business process.

For **openness**, T1010-01 specifies that data must be openly available and readable in standard formats at each stage in its lifecycle.

For **conditions of use**, the T1010-02 commercial principles permit licensing conditions and restrictions (see *Data Uses, Sharing, IP* (page 36)). T1010-01 defines a way in which this information can be represented in a standard way and included or associated with the data stream (see *Metadata* (page 54)).

11.3.4 Processing Pipelines

Figure 2 shows a typical (and relatively simple) data processing pipeline conforming to the *ISO 13374* principles: a Data Acquisition step followed by a chain of Data Manipulation steps and a State Detection step. (This is the data pipeline used in the IMPRCM prototype data broker).

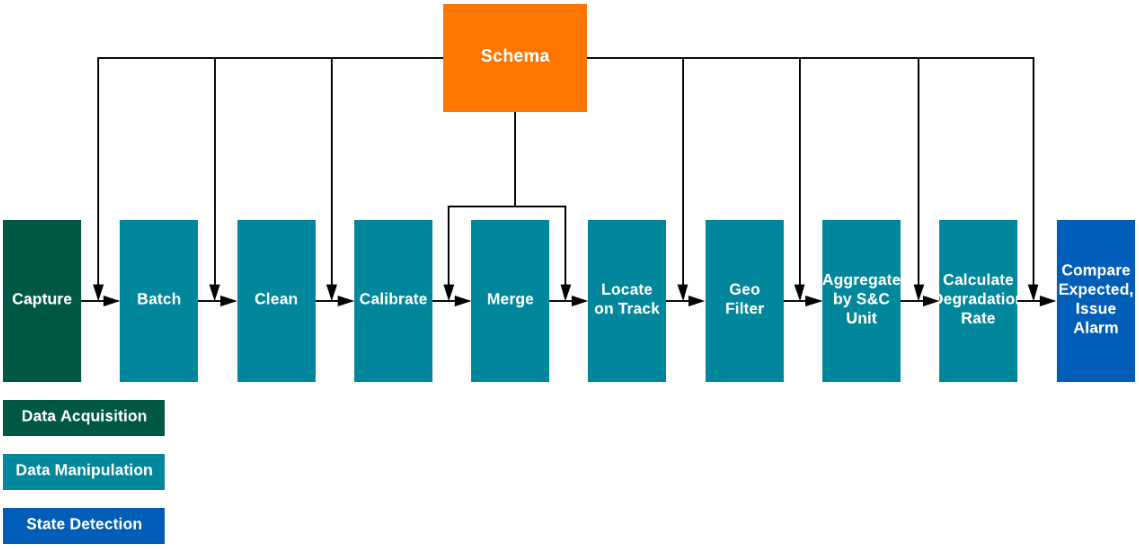


Figure 2: Sample Data Processing Pipeline

The data flows between each processing step are open and can be specified in the Information Framework by means of a *schema*. The processing steps need not be open - they may be procured from third parties - and so can safeguard any intellectual property associated with the processing.

It is likely that for some of the steps, standard algorithms or data services will be available: these should be used in preference to creating new bespoke ones for each flow. And conversely, any new processing step should be specified in the Information Framework to be designed to be re-used. This means the use of an *API* at the input and output sides of the processing step. T1010-01 gives some technical guidance on how to define these APIs (see *APIs* (page 54)).

The broker-based approach (described in *Technical Principles* (page 7) in the Introduction and tried out in *The IMPRCM Project* (page 6)) provides a good architecture for constructing these data pipelines in a generic and re-usable way.

Tip: The IMPRCM project built a prototype data broker in which some of the processing elements in the pipeline were services accessed using a *REST*-based *API*. The APIs were specified using *swagger*, a well-known standard method of defining APIs.

11.3.5 Ownership and IP

As mentioned above, the principles of *ISO 13374* require that data moving between processing steps be open and without any *IP*. IP can reside in the processing steps themselves.

CHAPTER 12

Technical Framework

12.1 Overview

The Technical Framework defines the various technical aspects of the RCM data interchange. It links closely with other frameworks:

- the *Information Framework* (page 45) defines what data need to be exchanged and the characteristics it needs to have to be useful
- the *Commercial Framework* (page 33) includes schedules defining the technical deliverables and service level agreements

These frameworks define the "what" of the RCM data interchange; the Technical Framework provides the "how".

Eacb data interchange project can define its own Technical Framework by direct agreement between the participants. However, it can be beneficial to adopt relevant standards and procedures

12.2 Components of the Technical Architecture

12.2.1 Data Item Definition

12.2.2 Metadata

Metadata.

12.2.3 APIs

APIs.

CHAPTER 13

Operational Framework

In future there may be a requirement for a framework for holding all of the documentation relating to operation agreements and activities.

CHAPTER 14

Address the technical concerns of a Cross-industry RCM scheme

This page contains all the guidance, or links to guidance, that may be needed by someone managing technical aspects of a Cross-industry RCM scheme.

14.1 Context

See *Introduction to Cross-industry RCM* (page 3) for a lay-person's guide to Cross-industry RCM.

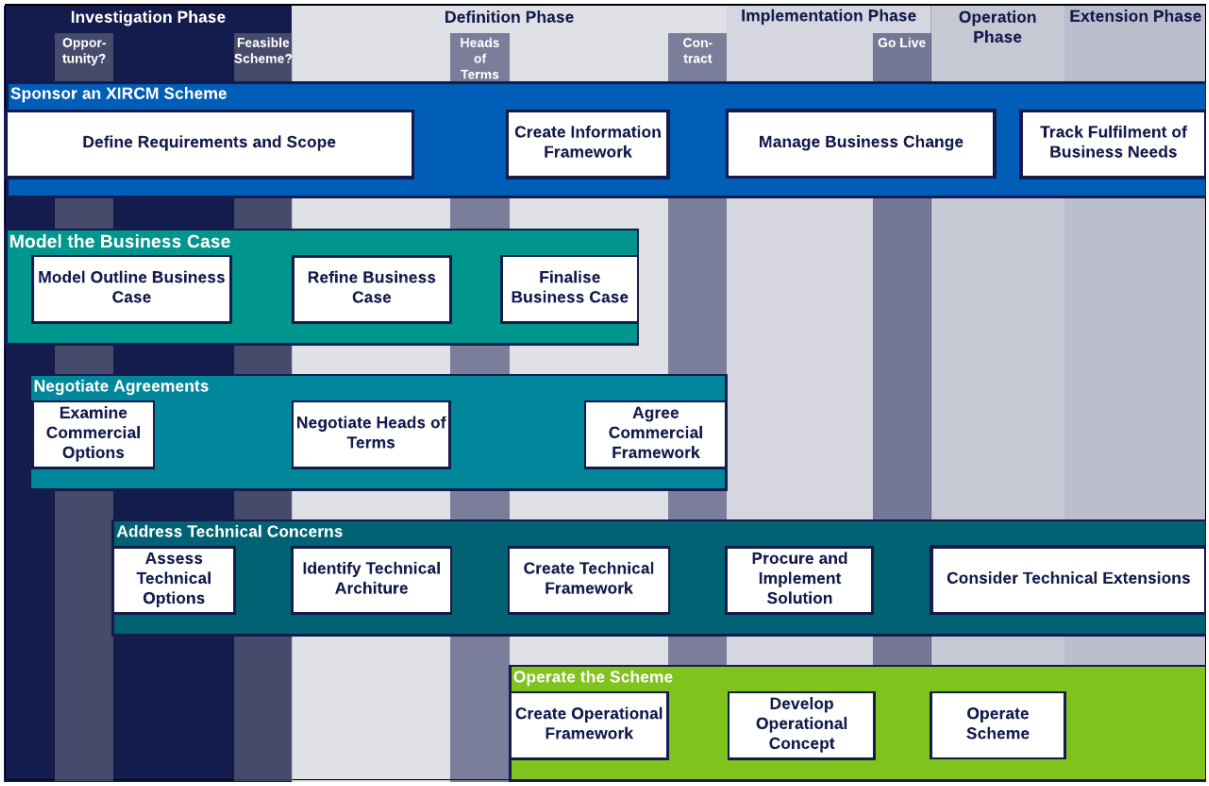


Figure 1: Cross-industry RCM Project Process Map - Overview

The figure *Cross-industry RCM Project Process Map - Overview* (page 56) shows the Address Technical Concerns workstream in the context of the project phases and gateways and the other workstreams. More detail on these

aspects of a Cross-industry RCM project can be found in *Process Map* (page 13) and the pages devoted to the other workstreams.

The technical concerns have two main focuses:

- the data to be interchanged in order to meet the business needs of the data consumers, and the constraints associated with the interchange: data quality, timeliness, reliability etc. These are expressed in the Information Framework (*Information Framework* (page 45)). The Information Framework feeds into the agreements between parties, specifically in Schedule 2 (Data) and Schedule 6 (Service Level Agreement)
- the means by which data are interchanged. This covers the data collection equipment, telecommunications and any hardware and software elements used to transfer, process, store and present the data. These are documented in the Technical Framework (*Technical Framework* (page 54)), which feeds into the agreements between parties in Schedule 2 (Equipment) and operational documentation.

The *T1010* data architecture documentation in T1010-01 is particularly relevant to the technical concerns. It sets out proposed standards for data formatting and content and approaches to interchange.

14.2 Tasks

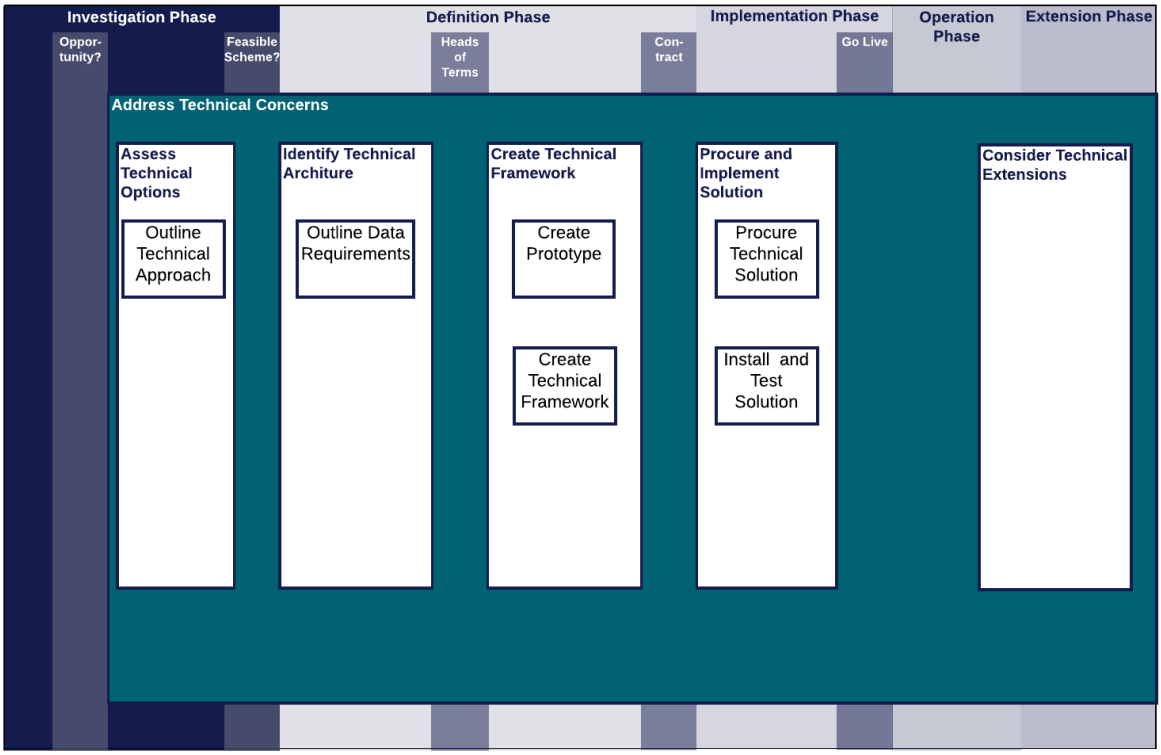


Figure 2: Technical Workstream

The figure *Technical Workstream* (page 57) shows the key tasks at each phase of a Cross-industry RCM project, as far as the Go Live gateway. These are described below.

14.2.1 Assess Technical Options

Following the definition of business requirements and scope in the *Sponsor a Cross-industry RCM scheme* (page 16) workstream, as expressed in *Information Framework* (page 45), this task considers the technical options for delivering the business requirements.

An early consideration should be whether there is actually need for a Cross-industry RCM scheme - whether existing data sources or sensors are already available.[1]

[1] Source: T1010-03 Section 3.1 Consider need for Cross-Industry RCM

14.2.2 Identify Technical Architecture

In this task, the data items required and a candidate technical approach to delivering them are identified and the costs, risks and timetable outlined to support the Sponsor, Business Case and Agreements workstreams. A key output of this task, produced along with the Sponsor workstream, is the *Information Framework* (page 45).

14.2.3 Create Technical Framework

In this task, the formal technical documentation is compiled to support the commercial agreement and the procurement of the required solution.

14.2.4 Procure and Implement Solution

In this task, the technical solution is procured, installed, tested and brought into use.

14.2.5 Consider Technical Extensions

Following implementation, this task involves looking for opportunities to extend the project and examining the technical implications. The extensions may be of scope - such as the enhancement to a higher level of the *ISO 13374* stack or the integration of other data sources - or of scale: deployment to a larger geographical area or a new type of asset.

14.3 Linked pages

CHAPTER 15

Operate the Cross-industry RCM scheme

This page contains all the guidance, or links to guidance, that may be needed by someone planning the live operation of an XI RCM scheme.

15.1 Context

See *Introduction to Cross-industry RCM* (page 3) for a lay-person's guide to Cross-industry RCM.

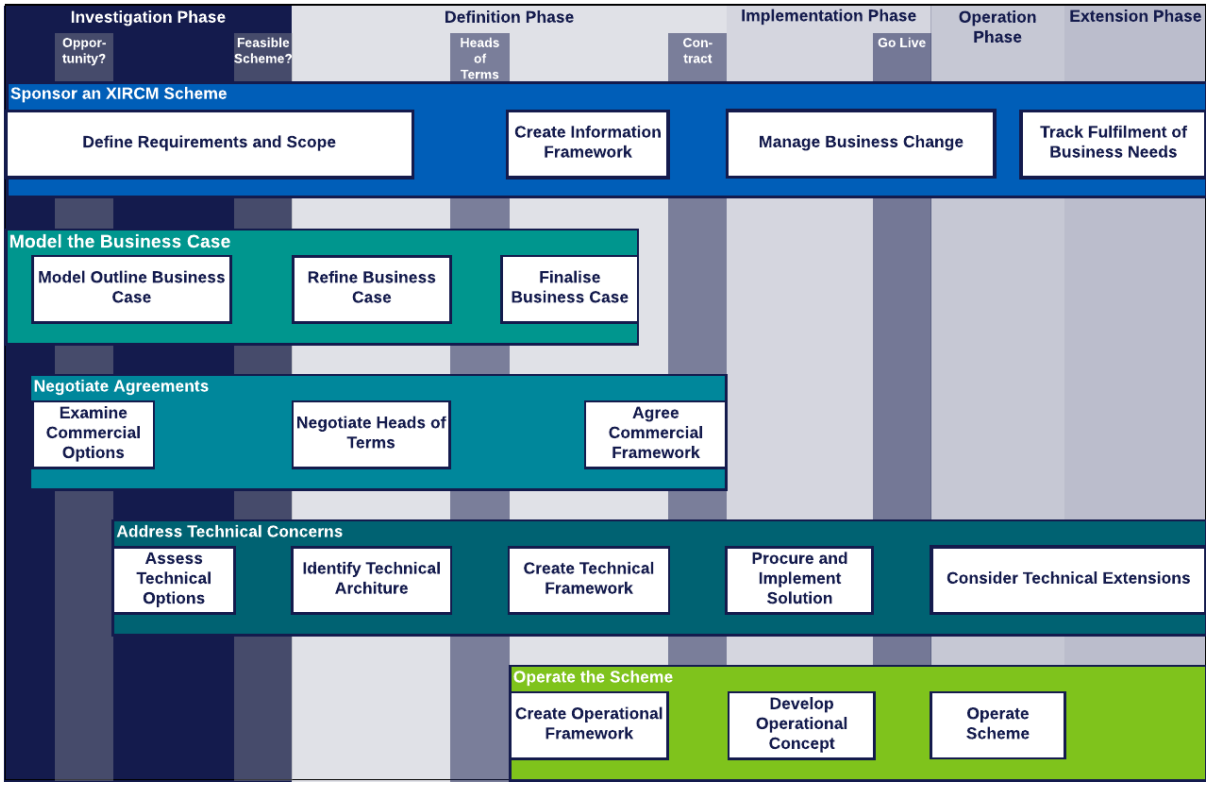


Figure 1: Cross-industry RCM Project Process Map - Overview

The figure *Cross-industry RCM Project Process Map - Overview* (page 59) shows the Operate the Scheme workstream in the context of the project phases and gateways and the other workstreams. More detail on these aspects

of a Cross-industry RCM project can be found in *Process Map* (page 13) and the pages devoted to the other workstreams.

15.2 Operate the Scheme: Tasks

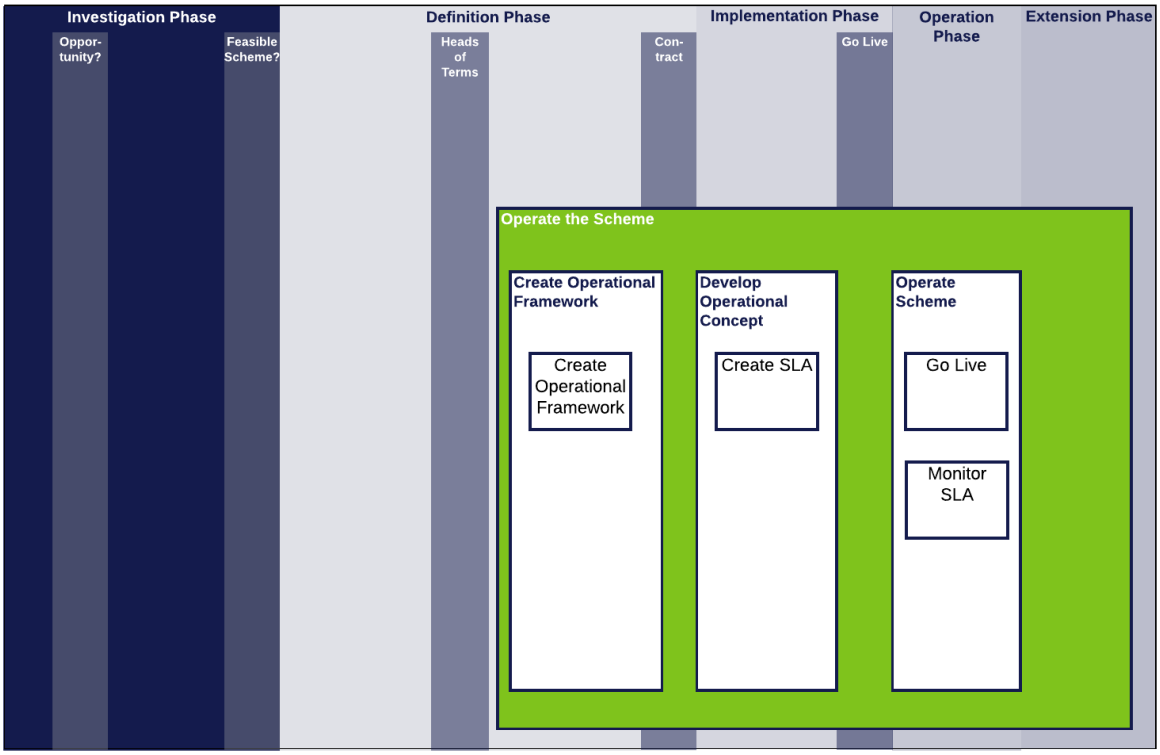


Figure 2: Operate the Scheme Workstream

15.2.1 Create Operational Framework

The Operational Framework is the part of the contractual agreement that deals with the live operation of the scheme: the mechanics of collecting, processing and storing the RCM data and making it available.

15.2.2 Develop Operational Concept

The Operational Concept is the set of processes, practices and documentation needed to operate the Scheme.

15.2.3 Operate Scheme

This is the tasks that would need to done to manage the scheme in live operation.

15.3 Guidance

15.3.1 Guidance for implementing an XI RCM scheme

At this stage just a placeholder for future work.

CHAPTER 16

Reference

This page contains links to the original source documents which inform the approach being taken in the toolkit.

16.1 T1010 Documentation

These are the documents produced by the RSSB research project *T1010* (Cross-industry RCM) which define the business case, commercial and technical toolkit for cross-industry RCM projects. They are the direct source for the work on this website.

16.1.1 Guidance and Orientation

- Good Practice Guide (T1010 03 GPG). An overview of the process of defining and running a cross-industry RCM project.
- Standards Guide (T1010 03 SG). This shows which existing standards apply to Cross-industry RCM work and suggests some future standards and industry-wide initiatives that might be helpful.

16.1.2 Commercial and Process-related

- Commercial Process Maps (T1010 02 Appendix C2).
- Commercial Report (T1010 02) Background and commentary on the commercial process and supporting documents.

16.1.3 The Business Case Tool

- Business Case Tool MS-Excel workbook (T1010 04, version 3.0)
- Business Case Tool User Guide (T857)
- Business Case Guidance - updated (T1010 04)

16.1.4 The Template Contract

- RCM Contract Template and Schedules (T1010 02 Appendix E). Text of the contract template and pro-forma schedules
- How to use the template contract (T1010 02 Appendix F). Guide to using the template contract.

16.1.5 The Technical Architecture

- Technical Architecture Overview (T1010 01 TAO). An overview of the guiding principles behind a standard approach to Cross-industry RCM data interchange.
- Architecture Principles and Requirements (T1010 01 APR). Requirements that data sharers and any central sharing infrastructure need to comply with to conform to T1010 standards.

16.2 Cross-industry RCMSG First-phase work

These documents are the results of the first phase of work sponsored by the *XIRCMSG*. They set the basic requirements and defined the industry constraints that the *T1010* work responded to.

- T844 Mapping current remote condition monitoring activities to the system reliability framework (T844). T844 investigated a large number of then-current RCM activities and worked out what impact improved cross-industry working would have on them. It made a number of recommendations about future work that should be done to support this goal.
- T857 06 Overview (T857).

16.3 Industry Context

These documents define the context within which the work of the *XIRCMSG* takes place.

16.3.1 The McNulty Rail Value for Money Study 2011

The McNulty study[1]

16.3.2 The Rail Technical Strategy 2012

The Rail Technical Strategy was produced by the Technical Strategy Leadership Group (TSLG) to set out a 30-year vision of a transformed railway that will continue to satisfy its customers. To do that while meeting goals of increased capacity, reduced cost and lower carbon consumption it will need to *"remove barriers and find more innovative and cost-effective ways to increase capacity and improve performance"*[2].

The *RTS* sets out a strategy to achieve this which has six workstreams. *XIRCM* is a key part of three of them:

- **Infrastructure**. The vision is for *"world-class asset management using train-borne inspection to provide accurate, timely information for condition-based intervention"*[3] using *"... remote condition monitoring ... across the railway to save significant costs and improve performance. Recommendations include ... integrate existing and isolated RCM models and practices ... use infrastructure fixed points to monitor rolling stock and use rolling stock to monitor fixed infrastructure"*[4].
- **Rolling Stock**. The vision is for rolling stock's whole-life costs to be reduced through reliability improvement and better interfaces with infrastructure. Among the objectives are *"a whole-system approach which ... improves condition monitoring..."*[5]. Enabling these objectives will be *"a cross-industry commercial and technical framework ... to allow asset condition data to be shared and exploited across all infrastructure and rolling stock asset types"*[6] and *"... Cross-interface RCM - trains monitor infrastructure and vice-versa"*[7].

1

2

3

4

5

6

7

- **Information**. The vision is for much more effective use of the industry's data, specifically: *"Agreed architectural standards . . . support connectivity, improve the capability to share the information resources and facilitate a whole-system approach"*[8]. Among the strategic goals are *"A suite of common and open standards"* and *"A common information framework"*[9]. Cross-industry RCM specifically needs these types of IT.

The *RTS* identifies three common foundations which link across the workstreams: a whole-system approach, innovation and people. Of these:

- the **whole system approach** directly impacts Cross-industry RCM. The report explicitly mentions the Cross-industry RCM Strategy Group (*XIRCMSG*) (see *The Cross-Industry RCM Strategy Group* (page 63) below) and its work; and in the vision for the future it includes *"Improved remote asset monitoring including RCM . . . increases the reliability of infrastructure and rolling stock"*[10]. Part of the strategy is the development of *"Aligned asset management plans"*[11] to bring coherence across all industry parties.
- **innovation** is necessary for Cross-industry RCM to flourish: it relies on novel equipment, processes and commercial structures. The strategy for innovation includes the key goal of *"commercial models to allow intellectual property (IP) to be accessed, shared and evolved for cross-industry benefits"*. These models should, among other things, *"support a quicker development cycle . . . [to] . . . allow products and services to enter the market more quickly and provide opportunities for wider exploitation throughout the rail industry and beyond"*[12].

16.3.3 The Cross-Industry RCM Strategy Group

The remit for the *XIRCMSG* is here:

> The remit of the *XIRCMSG*: XIRCMSG Remit

16.3.4 Capability Delivery Plan

This is the industry's plan for delivering the objectives of the *RTS*.

16.3.5 In2Rail

16.4 Standards

These are standards used in the definitions of the data architecture - BS ISO 13374 (*ISO 13374*) The ISO standard that defines the RCM processing model. See commentary in the *Introduction to Cross-industry RCM* (page 3).

16.5 Glossary of Terms

API Application Programming Interface - a way of standardising programs making requests for data or processing and the responses to those requests. The use of APIs aims to hide the complex details of the data source or processing from the user. Often provided using a *REST* approach.

AVI Automatic Vehicle Identification

AWS Advance Warning System - a method of warning train drivers of signals at caution and stopping the train if the driver does not acknowledge the warning.

BST British Summer Time

8

9

10

11

12

CDP Capability Delivery Plan - the industry's plan for delivering the objectives set out in the *RTS*. See *Capability Delivery Plan* (page 63) for details.

CP6 Network Rail's Control Period 6 - the monitoring period which runs from April 2019 to March 2024

CSV Comma-separated values - a simple textual file format for data, readable by standard desktop programs

Cucumber A tool for building and executing automated business-oriented tests written in the *Gherkin* language. See Cucumber test tool for details.

Data Broker A type of IT system architecture in which data providers and data consumers each connect once to a central hub, the broker, which mediates the data transfers between each provider and the consumers interested in its data.

ELR Engineers' Line Reference - a way of identifying a section of rail route between junctions or terminals.

ETCS European Train Control System - an in-cab signalling system based on European standards which will ultimately allow lineside signals to be replaced.

Gherkin A simple text-based language for defining acceptance tests in business-friendly language, that also lends itself to automatic testing using tools such as *Cucumber*. Tests have the structure **Given:** <some preconditions> **When** <a stakeholder interacts with the system under test in a particular way> **Then** <this will be the result>. See Gherkin language for details.

GMT Greenwich Mean Time

Gold Plating Over-specification or over-delivery of features of a technical solution beyond that necessary to fulfil its function. See gold plating for details.

GPS Global Positioning System

GRIP Governance for Railway Investment Projects. See for details.

HABD Hot Axle-box Detector

HTTP Hypertext Transmission Protocol - the protocol used to build and transfer web pages, characterised by the ability to link from one page to another.

In2Rail An EU initiative intended, among other things, to standardise various types of data interchange across European railways by the use of a set of data standards and a broker-based data sharing mechanism based on a central "Canonical Data Model". There is considerable overlap and good consistency between these requirements and the ones set out in *T1010*.

IP Intellectual Property

ISO International Standards Organisation

ISO 13374 ISO standard which defines the six-layer processing model and a data architecture for RCM data. See *The ISO 13374 model* (page 7) for a description.

ISO 6709 ISO standard that defines how to format geographical positions. See ISO 6709.

ISO 8601 ISO standard that defines how to format dates and times. See ISO 8601

IT Information Technology

JSON Javascript Object Notation - a textual format for encoding data, frequently used in web-based data interchanges and readable by most computer languages. Can be validated using *JSON Schema*.

JSON Schema A way of defining the structure and content of a *JSON* file so that it can be validated.

MATLAB A commercial data analysis / simulation program

metadata "data about data" - information about matters such as data format, units, provenance and quality

MIME type One of the standard file formats accepted for transmission over the web. In addition to the data formats described elsewhere here, it can also include media formats for sound and video.

MIMOSA The Machinery Information Management Open System Alliance, a standardisation body which defines an open framework for the interchange of measurement information between parties. The framework is an implementation of the standard approach defined in *ISO 13374*. See MIMOSA.

NR Network Rail

OLE Overhead Line Equipment

Python/Pandas An open-source language and toolkit often used for data analysis

QUDT A standardisation effort to provide clear specifications for Units of Measure, Quantity, Dimensions and Data Types. See QUDT

R An open-source statistical analysis language

RailML An *XML* data format designed for the interchange of rail rolling stock, infrastructure and operational information

RailTopoModel A standard way of describing the topology and layout of a railway network and the assets associated with it

RCM Remote Condition Monitoring

REST Representational State Transfer - a type of *API* which uses the *HTTP* standards of the web to manage interactions between software programs, including the interchange of data. Very commonly used and widely understood.

RFID Radio Frequency Identification - a way of identifying assets by tagging them with radio transponders which broadcast an identifier as they pass a reader.

RSSB The Rail Safety and Standards Board

RTS The Rail Technical Strategy

S&C Switches and Crossings

schema A shared definition of a set of data items - a key component of a Canonical Data Model.

SI System Internationale: the standard set of units of measure based around the metre, kilogram and second

Stata A commercial statistical analysis program

swagger A standard method of defining an *API*. See Swagger Spec.

T1010 RSSB's research project T1010, which developed the Cross Industry RCM toolkit. See an introductory description here: *The T1010 Project* (page 5); the project reports are here: *T1010 Documentation* (page 61).

TfL Transport for London

TSLG The Technical Strategy Leadership Group

UCUM A standard for naming units of measure to support communication between IT systems. See UCUM

UGMS Unattended Track Geometry Measurement System

UTC Universal Co-ordinated Time - equivalent to GMT for all practical purposes. See UTC.

UUID Universally unique identifier - a binary number and standard string representation, of a defined format, guaranteed within any reasonable set of circumstances to be unique. See RFC 4122 - UUIDs.

VRC Vivacity Rail Consulting Ltd. See .

XIRCM Cross-Industry Remote Condition Monitoring

XIRCMSG The Cross-Industry RCM Strategy Group

XML Extensible Markup Language - a textual format for encoding data, commonly used in industrial data interchange protocols. Supports validation using an *XML Schema* and programmatic translation into other formats.

XML Schema A way of validating an *XML* file to verify that it has the correct structure and only contains allowed and correctly-formatted data items

/crs

`GET /crs/{crs}`, **??**

/reason

`GET /reason/cancelled`, **??**
`GET /reason/cancelled/{code}`, **??**
`GET /reason/late`, **??**
`GET /reason/late/{code}`, **??**

/tiploc

`GET /tiploc/{tiploc}`, **??**

/toc

`GET /toc/{toc}`, **??**

/via

`GET /via/{at}/{dest}/{loc1}`, **??**
`GET /via/{at}/{dest}/{loc1}/{loc2}`, **??**