National Rail Safety Data Scheme

AUSTRALASIAN RAILWAY ASSOCIATION

onrsr Office of the National Rail Safety Regulator

National Rail Safety Data Scheme

Version 2 – January 2022

Table of Contents

1 Purpose and overview

The rail safety data scheme set out in this document establishes a national approach for the capture and management of - and access to - rail safety data. The data covered by this scheme is principally concerned with rail safety occurrences and associated network and railway operations characteristics that create a dataset that can be used to inform and monitor rail safety performance across Australia.

1.1 Objectives

Established under the *National Rail Safety Data Strategy 2018-2022*, a joint initiative between the rail industry (represented by the Australasian Railway Association, ARA) and rail safety regulator (the Office of the National Rail Safety Regulator, ONRSR), this scheme establishes the principles, methods and requirements for the creation, management and use of a National Rail Safety Dataset that is collected from rail transport operators and accessible by a range of rail safety stakeholders.

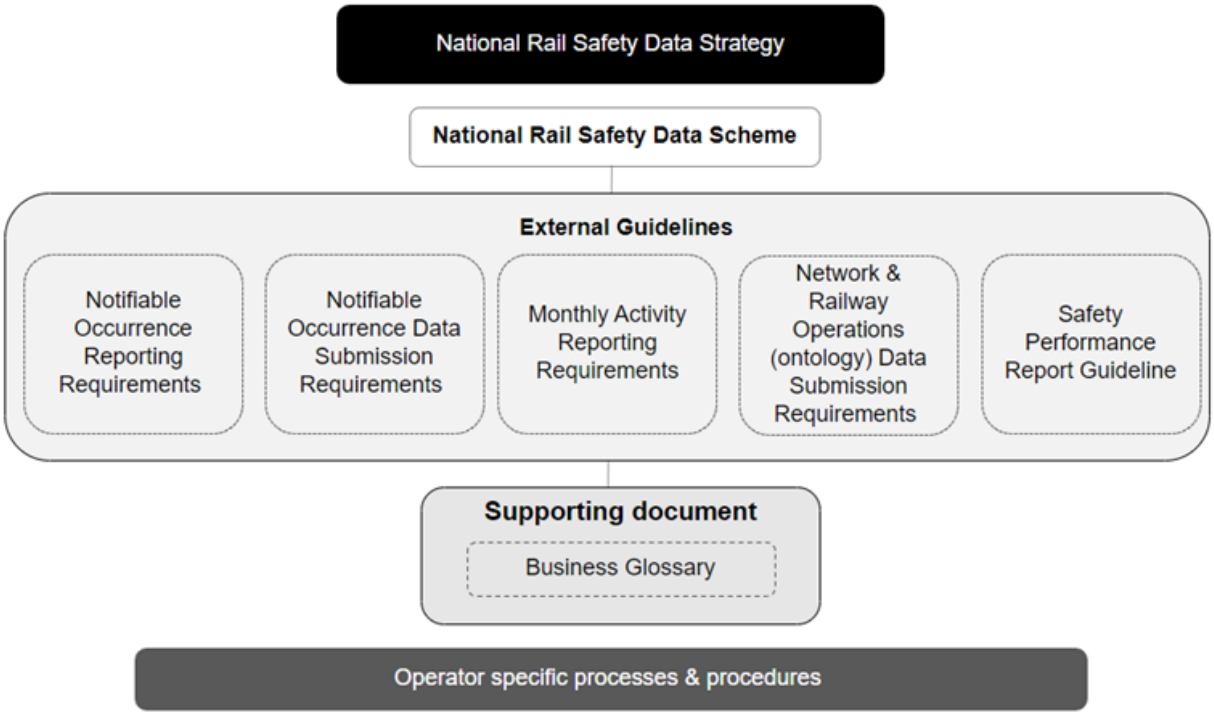
This scheme will:

> provide timely, accurate and relevant information about rail safety performance;

> balance the need for data with the appropriate burden of collection; and

> establish a National Rail Safety Dataset that can be accessed by rail transport operators and stakeholders.

1.2 Key documents



1.3 Data use

All stakeholders directly or indirectly involved in the management, monitoring and improvement of rail safety recognise the importance of consistent, quality and accurate data when making decisions relating to rail safety. The data collected and made available under this scheme is to provide a means of:

> encouraging and supporting good decision-making about rail safety; and

> guiding actions to improve rail safety by enabling monitoring and understanding of rail safety performance.

The data:

> is generated by accredited rail transport operators meeting the legislated reporting requirements for the reporting of notifiable occurrences, and other information, to ONRSR and the Australian Transport Safety Bureau (ATSB).

> is used by ONRSR to:

  - be informed of notifiable occurrences in a timely manner to enable an appropriate regulatory response.
  - monitor the safety performance of regulated rail transport operators in order to strategically determine the regulatory approach and key safety issues to be addressed under ONRSR's risk-based approach to regulation; and
  - produce publicly available performance monitoring reports (e.g. ONRSR's Annual Report and Rail Safety Report*).*

> allows rail transport operators to benchmark performance with similar railway operations and to inform risk management decision making, which may include the population of risk models.

> assists rail industry stakeholders, particularly those that can influence or are involved in safety improvement investment decision making, to inform themselves of rail safety performance.

1.4 Legislation

The *Rail Safety National Law* RSNL) and *Rail Safety National Law National Regulations 2012* (the Regulations) set out the reporting requirements for data under this scheme and, where permitted by the law, ONRSR sets the manner and form in which data needs to be submitted. The Regulator's manner and form is established through the provision and maintenance of a data capture system, the ONRSR Portal.

The RSNL and the Regulations gives effect to this scheme through the following provisions[1]:

Notifiable occurrence data	> s121 of the RSNL, which requires the reporting of notifiable occurrences in a manner and time prescribed by Regulations. > Regulation 57, which sets out the definition of Category A, B and C occurrences, the required reporting periods and provides that the Regulator can determine the manner and form of reporting. > Schedule 1A of the Regulations*,* which sets out the incidents and events that are Category A, B or C notifiable occurrences.

[1] This document anticipates the introduction of Regulation changes proposed in the *Rail Safety National Law Nation Regulations (Reporting Requirements) Variation Regulation 2021*.

Network and railway operations (ontology) data	> s120 (3) of the RSNL, which requires the reporting of prescribed information for a period and in a time, manner and form approved by the Regulator. > Regulation 56A, which sets out the nature of information to be provided and reviewed annually, or as otherwise required by the Regulator.
Rail operations activity (monthly) data	> s120 (3) of the RSNL, which requires the reporting of prescribed information for a period and in a time, manner and form approved by the Regulator. > Regulation 56, which sets out the details of the information to be provided for a monthly period.

1.5 Data submission system – ONRSR Portal

The ONRSR Portal is the primary and preferred means for operators to report occurrences and submit other data and information required under this scheme.

The ONRSR Portal provides an intuitive interface that enables the entry of information by an operator to be:

> guided through a smart-form type interface, presenting the user with selection choices driven by the information already entered; and

> tailored to the operator - providing selection options based on the operator's specific scope and nature of operations.

Further information on the ONRSR Portal is available on the ONRSR website: www.onrsr.com.au.

Paper form submission of data is discouraged and will only be acceptable in limited circumstances after discussion with ONRSR.

1.5.1 API data exchange

The ability for operators to submit data via an application programming interface (API) is being considered by ONRSR but this option will not be available until sometime after the commencement of this scheme.

Utilising an API will enable an operator to establish the functionality required to report notifiable occurrences and submit relevant data within their own technology system. While ONRSR will provide documentation of the API, establishing this functionality is at the operator's own expense.

1.6 Commencement

The collection of data under this scheme commences in July 2022, with the data made available to rail transport operators and other stakeholders in the first quarter of 2023.

As the scheme is given effect through legislation, and in acknowledgement of the investment required by all parties involved in generating and using the data, the scheme has been designed to be durable into the future. It establishes key principles for rail safety data management (Section 3) that are intended to remain core principles for the ongoing maintenance, review and amendment of the scheme.

A formal review of the scheme will occur in July 2023.

2 National Rail Safety Dataset

The National Rail Safety Dataset created under this scheme comprises three distinct data subsets:

Notifiable occurrence data	**Network and railway operations (ontology) data**	**Rail operations activity (monthly) data**
Description of rail safety notifiable occurrences, including likely causes and consequences (section 4.1).	Geographical network and network and train operational characteristics data providing an understanding of the nature and scope of an operator's railway operations (section 5).	Operational activity data providing a view of the extent of an operator's activity in key operational and management areas (section 6).
This data is used for:	This data is used for:	This data is used for:
> reporting of notifiable occurrences.	> improving the accuracy and efficiency of reporting of notifiable occurrences by providing input data to data field selections, e.g. location information.	
	> normalising notifiable occurrence data for comparative analysis.	> normalising notifiable occurrence data for comparative analysis.
> regulatory decision making for regulatory response to the occurrence.		
> ATSB decision making for investigative response to the occurrence.		
> regulatory decision making for prioritising regulatory effort and activities.	> regulatory decision making for prioritising regulatory effort and activities.	> regulatory decision making for prioritising regulatory effort and activities.
> stakeholder use and analysis to support safety related decision making.	> stakeholder use and analysis to support safety related decision making.	> stakeholder use and analysis to support safety related decision making.
> determination of fees under the *Rail Safety National Law.*	> determination of fees under the *Rail Safety National Law.*	> determination of fees under the *Rail Safety National Law.*

3 Key principles

The concepts and principles presented in this section set the foundation for the national rail safety data scheme, building on the strengths and addressing the limitations of the historical approach to the collection and use of rail safety data.

Appendix A discusses the key changes to the historical approach.

3.1 Data analysis and performance monitoring responsibility

This scheme is aimed at the creation of and access to a National Rail Safety Dataset to provide information to stakeholders to guide and improve risk and safety management decision making. In doing so, it does not detract from the primary safety responsibilities established by the RSNL. The RSNL places the responsibility to manage risks to safety, so far as is reasonably practicable, on rail transport operators (principally) and other duty holders.

This means that alongside this scheme the Regulator has an expectation that operators are monitoring their own operational performance through their collection and analysis of data and information relevant to the scope and nature of their railway operations. This expectation is reinforced by the RSNL requirements for operators to establish key performance indicators for measuring safety performance and determining the effectiveness of the safety management system, along with having processes for the dissemination of safety information.

Operator's should make use of the dataset in meeting risk and safety management duties and responsibilities through reviewing and explaining data analysis and outcomes in the annual Safety Performance Report.

The national dataset is established to assist a range of stakeholders with risk and safety management decision making, providing the opportunity to access a dataset to undertake performance benchmarking or to inform risk assessments (particularly quantitative risk assessments) or the setting of risk tolerability limits. However, the dataset is not the only source available to rail transport operators for this purpose.

The national dataset is not considered sufficient for the effective monitoring and management of rail safety risks at an operational level, nor is the capture of the national data intended to enable operators to meet other regulatory (e.g. work, health and safety law) or organisational reporting requirements.

3.1.1 Safety Performance Reports

The RSNL (s103) establishes a requirement for accredited rail transport operators to submit a Safety Performance Report to the Regulator for the purpose of informing the Regulator of the operator's activities associated with safety management system effectiveness and addressing safety performance matters. While not a component of the national rail safety data scheme, the Safety Performance Report provides the formal mechanism for presenting relevant data analysis and resulting safety management decision outcomes to the Regulator.

Essentially, the report provides the opportunity for the operator to describe to the Regulator the important safety related anomalies, adverse or improving trends that appear in the operator's data, the safety issues arising and corresponding improvement actions that are being considered or have been taken. Operators should be mindful that the Regulator will have access to the data within the National Rail Safety Dataset, particularly in relation to notifiable occurrences, and the opportunity should be taken to demonstrate to the Regulator that any trends or spikes in this data are understood and are being addressed.

Guidance on the expected content of a Safety Performance Report, including an example, is set out in the ***Safety Performance Reporting*** guideline document.

3.1.2 ONRSR data analysis and reporting

The Regulator has functions (in relation to this scheme) under the RSNL to work with rail transport operators and others involved in railway operations to improve safety nationally, to collect and publish information relating to rail safety and promote and coordinate the sharing of information to achieve the safety objectives of the RSNL. These functions are, in part, discharged through the publication of data forming the National Rail Safety Dataset.

ONRSR does and will continue to provide information directly to operators and publish information relevant to the industry for the purpose of communicating safety issue concerns and in the promotion of continuous improvement of rail safety.

However, ONRSR does not undertake analysis of data for the purpose of providing a service to inform operators of emerging issues or trends in relation to their railway. This task is expected to be undertaken by the operator.

3.2 Data capture

Principles driving the capture of data under this scheme are:

Capture data once and use many times	Aimed at minimising the need to capture the same information multiple times, the scheme is structured to capture data once and then call on this data when it needs to be used to support other data capture, analysis or reporting needs. This means that data will be shared across operators and stakeholder groups within the confidentiality constraints of s244 of the RSNL. The primary example of this principle is network and railway operations (ontology) data, which is about capturing, once, core data that does not change significantly over time and then using this data to streamline occurrence reporting and providing a common location reference (by making rail infrastructure manager network location data available to relevant rolling stock operators). The capture of data is managed by ONRSR as the single collector.
Capture only what is used	The burden of capturing data that is under-used or not used at all is recognised, with the scheme focused on the national collection of data that has an identified operational or strategic need by the Regulator, rail transport operators and other industry stakeholders. Data is not collected for speculative reasons.
Data is suitable for multiple analysis needs	To enable multiple analysis applications of the National Rail Safety Dataset, the dataset comprises fields of codified data comprising factual (measurable or observable) information. This means the dataset: > minimises the use of free text, removing the need for stakeholders to invest in free-text mining analytics solutions; > is set up to enable users to use the data as building blocks to generate datasets that are relevant for their intended use, effectively to establish their own classification schemes; > contains an Amalgamated Record of the occurrence (section 4.2.3) providing a single record for an occurrence.

Data reflects what is known today	To maintain the dataset as an accurate record of what has occurred and what is in place, the data remains open for amendment at any time to address new information from investigations, errors or omissions. Supporting this approach, the dataset is managed to enable the ability to retrieve data as it was at a past point in time. Note: The National Rail Safety Dataset is presented, by default, as a current dataset, with what is valid now data.
Capture requirements and methods are stable	The scheme recognises the investment needed for all parties to establish systems and processes for the collection and use of the data generated. The scheme is intended to reset the national approach to the management of rail safety data on an ongoing basis. Certainty and stability of the scheme is established through regulation that brings with it a structured process for managing change. Some amendments to the scheme and data requirements will occur over time but these changes should remain aligned with these principles and concepts.

3.3 Data utilisation – National Rail Safety Dataset

The following principles drive the content and form of the National Rail Safety Dataset:

A single national dataset for all users	The data is captured with the aim of providing a published National Rail Safety Dataset that can be used to guide risk and safety decision making. The National Rail Safety Dataset will: > provide a single, national dataset that is available to all users; > provide operators and other stakeholders access to a national picture of rail safety performance, but not access to the performance of a specific operator (i.e. the dataset will be de-identified); and > be accessible as far as permitted within the confidentiality constraints of s244 of the RSNL. Note: Data will be retained as an historical record when an operator surrenders accreditation and no longer operates.
Dataset is for giving access to data, not reporting.	The National Rail Safety Dataset provides access to national rail safety data, which can be extracted by stakeholders for their own use and analysis. The product provided will allow benchmarking.
ONRSR and prescribed authorities (RSNL) will have full access.	To undertake their regulatory and investigative roles under the respective legislation, ONRSR and the ATSB will have access to - and use - the full dataset generated under this scheme. The Rail Industry Safety and Standards Board (RISSB) has access to the full dataset in order to populate its Australian Rail Risk Model. The ATSB and RISSB are prescribed authorities under the RSNL for the purpose of sharing information.

4 Notifiable occurrence data

Notifiable occurrence data comprises the capture of data that describes key characteristics of a specific set of incidents that must be reported to the Regulator.

A notifiable occurrence is defined in s4 of the RSNL as an accident or incident associated with railway operations that has, or could have, caused significant property damage, serious injury or death. Established through Regulations, the specific accident or incident types that must be reported are described in the ***Notifiable Occurrence Reporting Requirements*** document.

For the purpose of establishing the required time period for an operator to report an occurrence to the Regulator, notifiable occurrences are categorised as Category A, B or C.

Category A Notifiable Occurrence	**Category B Notifiable Occurrence**	**Category C Notifiable Occurrence**
Rail safety incidents and events that: > are likely to warrant immediate regulatory response, (including drug & alcohol testing); or > are likely to attract significant external stakeholder attention due to seriousness of the occurrences or its consequences; or > involve a fatality or serious injury related to train operations.	Rail safety incidents and events that: > may trigger a follow-up regulatory response; or > are indicative of a significant failure of a safety management system.	Rail safety incidents and events that: > unlikely to warrant regulatory follow-up for a single occurrence; or > when considered as multiple events, provide an indicator of the operator's safety performance and/or effectiveness of the safety management system, which may warrant regulatory follow-up due to adverse trending.

The classification of occurrences as Category A, B or C is set out in the ***Notifiable Occurrence Reporting Requirements*** document. The data that needs to be submitted for a notifiable occurrence is set out in the ***Notifiable Occurrence Data Submission Requirements*** document.

For the purpose of regulatory oversight of rail safety management and performance, accredited rail transport operators are:

> required to capture and submit occurrence data under this scheme; and

> expected to complement the data submission requirement by providing commentary of their own analysis, assessment and responses for rail safety data relating to their operations as part of their Safety Performance Report (section 3.1.1).

4.1 Notifiable occurrence reporting

Rail transport operators must report notifiable occurrences to the Regulator within the following time periods.

Category A Notifiable Occurrence	**Category B Notifiable Occurrence**	**Category C Notifiable Occurrence**
> Immediately report an occurrence via telephone to the Regulator.		
	> Initially Report an occurrence within 72[2] hours via the ONRSR Portal.	
> Complete the Data Submission via the ONRSR Portal within 7 days.		
	> Complete the Data Submission via the ONRSR Portal within 14 days.	
		> Summary total of occurrences data submitted within six months of the end of the operator's reporting period[3]
> Submitted data updated at any time following investigation or validation activity to correct an error or omission.	> Submitted data updated at any time following investigation or validation activity to correct an error or omission.	> Submitted data updated at any time following investigation or validation activity to correct an error or omission.

4.2 Category A and B occurrences – data content and concepts

Notifiable occurrence data is collected for use in:

> informing post incident response by ONRSR and the ATSB;

> informing ONRSR's assessment of operator's safety performance and risk-based decision making regarding regulatory response and effort;

[2] A change to 7 days is proposed that will require amendments to the *Transport Safety Investigation Act 2003*, which will be considered and progressed by the ATSB subject to approval by the Minister.

[3] reporting period refers to the reporting period applicable to the operator's Safety Performance Report.

> safety performance monitoring and reporting;

> input into the risk profiling tool used for regulatory cost recovery (fee) calculations; and

> input into the Australian Rail Risk Model (managed by RISSB).

The data generated from the reporting of notifiable occurrences comprises:

Data type	Reported by	Description	Primary Use(s) Comments
Occurrence type descriptor	Established by the RSNL	Identification of notifiable occurrences for which data is created, as defined in the ***Notifiable Occurrence Reporting Requirements*** document.	Establish a national dataset for notifiable occurrences.
Who, where and when data	Involved rail transport operators	The facts of when and where an occurrence occurred, and the rail transport operators involved.	Provide location and time-based information for occurrence data. Data identifying the operators involved is only shared between the directly involved operators, ONRSR, ATSB and any other authorities with the legal right to access. Data available through the National Rail Safety Dataset is de-identified.
Occurrence description	Involved rail transport operators	Provided when immediately/initially reporting a Category A or B occurrence. Free text description of the incident and any known (at the time) contributing and consequence information.	Shared only with directly involved operators, ONRSR, ATSB and any other authorities with the legal right to access. Provides contextual information to guide post-incident regulatory response (ONRSR) and investigation response (ATSB).
Occurrence specific data - descriptor	Involved rail transport operators	Coded data that provides additional contextual information on the nature of an occurrence or the network or railway operations associated with the occurrence.	Generate contextual (nature of occurrence and/or operations) information for occurrences.

Data type	Reported by	Description	Primary Use(s) Comments
Occurrence specific data - likely cause / precursor data	Involved rail transport operators	Coded data that provides information on the precursor or likely cause event that led to the occurrence. This is not comprehensive causal factor data.	Generate causal and/or contributing information for occurrences. Generate sequence of event information for an occurrence.

Key concepts that apply to the management of occurrence data are described in the following.

4.2.1 Occurrence location data

To generate accurate location data for occurrences in a form that can be used in analysis, the concept of a line section and associated kilometre point is adopted as the standard method of reporting location. The line sections are defined by the relevant rail infrastructure manager through submission of their ontology data (section 5).

Line section and associated kilometre point is required to be reported for any location-specific occurrence, however, if the relevant ontology data is provided by the rail infrastructure manager an option to report locations based on key locations will be available.

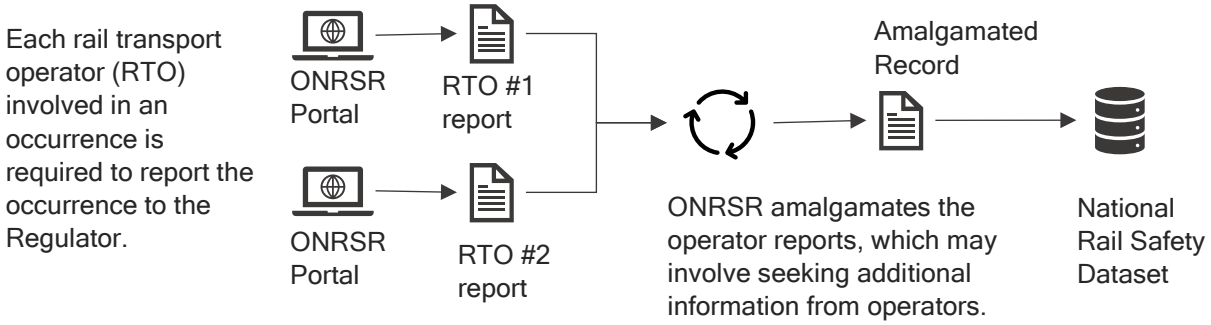
4.2.2 Sequence of event information

This scheme aims to generate rail safety data that can be used to draw insights into why incidents are happening or estimating the likelihood of serious incidents. To achieve this, it is recognised that occurrences are not singular events but are usually a series of sequential events that lead to the accident or incident that is the notifiable occurrence. Therefore, the data required for each type of notifiable occurrence includes precursor or likely cause data and consequence data in order to build a sequence of events for each occurrence. It is not intended to generate comprehensive contributing factor data but capture enough of the event sequence to make the data meaningful for generating insights beyond just a count of occurrence types.

The precursor or likely cause and consequence events relevant to each notifiable occurrence type are set out in the ***Notifiable Occurrence Data Submission Requirements*** document.

4.2.3 Amalgamated Record

Notifiable occurrence data is generated by the operators directly involved in the incident meaning that where multiple operators are involved, multiple reports of the same occurrence are submitted to the Regulator. When this occurs, in addition to receiving these reports, ONRSR manages a process to combine these separate reports into a single record of the occurrence, reflecting the information provided by each operator, known as the **Amalgamated Record**.



Where multiple operators are involved, some of the basic facts at the time of the initial report (for example, operators involved, the location and date/time) are not considered contentious. The Amalgamated Record process will make these basic facts received in the first report of an occurrence available to operators making subsequent reports, to streamline the process (refer to section 4.2.4).

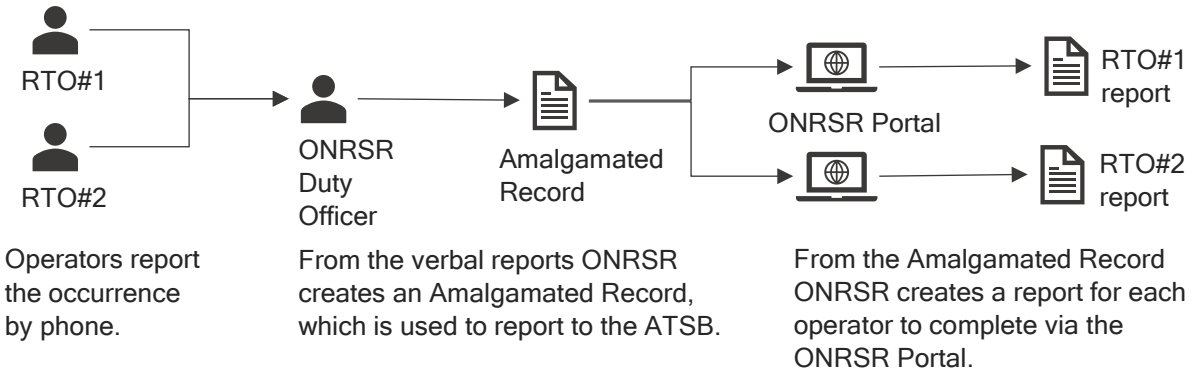
The original reports received from operators are retained by ONRSR, but only the Amalgamated Record is published into the National Rail Safety Dataset, ensuring that there is only a single record for a single occurrence. Each operator retains access to its original report via the ONRSR Portal, which can be amended by the operator if needed.

4.2.4 Immediate/Initial Reporting

The purpose of the initial report, or immediate report for a Category A occurrence, is to inform ONRSR and the ATSB of the occurrence so as to enable any regulatory (ONRSR) or investigative (ATSB) response to be initiated within a reasonable time.

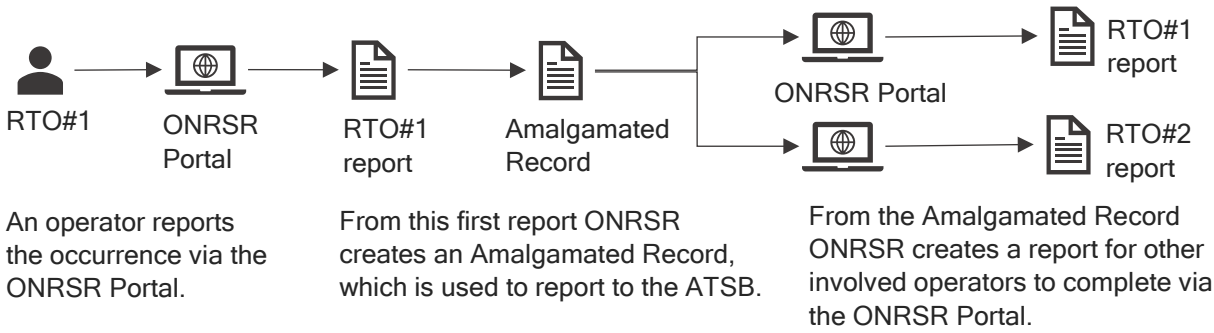
ONRSR manages the information at this stage of reporting and informs the ATSB of the occurrence, as required under the *Transport Safety Investigation Act 2003*, and shares information between directly involved rail transport operators for the purposes of validating factual information. This also assists with the completion of the written reporting and data submission of occurrences, including the linking of the separate operator reports.

For **Category A occurrences**, ONRSR will initiate the creation of an ONRSR Portal occurrence report for each operator to complete in the following way:



The description of the occurrence created by ONRSR from the verbal operator reports is provided to the involved operators in the ONRSR Portal record and can be amended by the operator as required.

For **Category B occurrences**, the first reporting operator will initiate the creation of an ONRSR Portal occurrence report for themselves and any other operator involved in the occurrence (where a report would be expected from the operator), in the following way:



The description of the occurrence provided by the first reporting operator is not included in the ONRSR Portal occurrence report created for the other involved operators.

Rail transport operators can add to or amend the information contained in the ONRSR Portal occurrence reports. These reports are only considered to be the operator's once the operator has formally submitted the report. Once submitted the operator retains responsibility for the accuracy of the information provided.

4.2.5 Data submission

The purpose of the data submission is to capture the balance of required data not already provided with the initial report. Additional time is provided for this submission, to enable operators to collect additional information and confirm the accuracy of the data.

The data submission is completed via the ONRSR Portal, using the operator's occurrence report that was created when initially reported. The operator may make this data submission at the same time as the initial report, where they have the required information available, or at any subsequent time within the relevant reporting period.

The operator has the option of identifying that the occurrence is still under investigation when submitting the likely cause data. Where this occurs, the ONRSR Portal will identify this report as requiring an update once the investigation has been completed.

4.2.6 Updating following investigation

Occurrence data remains open for amendment with additional information becoming available through investigation or other data validation activities.

Following any occurrence investigation:

> rail transport operators are expected (by the Regulator) to correct their occurrence record via the ONRSR Portal in response to any investigations undertaken into their occurrences; and

> ONRSR may amend the Amalgamated Record in response to investigation report reviews associated with ONRSR investigations, ATSB investigations, investigations required under s122 of the RSNL and operator investigations provided to the Regulator.

4.2.7 ONRSR Portal – occurrence reporting

The initial reporting of Category B occurrences and the submission of data for all categories of notifiable occurrences is via the ONRSR Portal featuring:

> tailored presentation of data field selections that draw from relevant ontology data where appropriate;

- guided reporting of the sequence of events for an occurrence - allowing an operator to commence reporting of an occurrence with any of the notifiable occurrence types associated with the occurrence;
- validation assistance ensuring occurrence reports are complete by the due date, including alerting the operator when the information is not complete;
- automated monitoring of occurrences identified as “under investigation” to ensure that updated data is provided once an investigation is complete; and
- ability to access and amend previously submitted data.

4.3 Category C notifiable occurrences

Category C notifiable occurrence data comprises:

Data type	Reported by	Description	Primary use(s)
Numbers of occurrences of a defined type	Rail infrastructure managers and/or Rolling stock operators	Summary count of the number of occurrences of a defined Category C occurrence type.	Input into ONRSR’s safety performance monitoring and risk-based regulatory decision making.

As this data is captured in a summary form, to mitigate the potential for duplication of occurrences that are reported, Category C occurrences are only required to be reported by one of the operators involved in an occurrence. This requirement is set out in the ***Notifiable Occurrence Data Submissions Requirements*** document.

The submission of Category C notifiable occurrences data is via the ONRSR Portal featuring:

- tailored presentation of data field selections that draw from relevant ontology data where appropriate;
- data upload capability for the submission of Category C occurrence data[4];
- ability to progressively report Category C occurrence data prior to the reporting period submission date;
- validation assistance ensuring occurrence reports are complete by the due date, including alerting the operator when the information is not complete;
- ability to access and amend previously submitted data.

While the reporting period and due date of the Category C data submission aligns with those of the Safety Performance Report, the Category C occurrence data submission can be made separately to the submission of the Safety Performance Report.

[4] Subject to final ONRSR Portal data submission design.

5 Network and railway operations (ontology) data

Network and railway operations data comprises data that describes the geographical nature of a railway network and network and train operational characteristics. It effectively establishes a description of the manner and form of railway operations in a digital data form.

The data comprises:

Data type	Reported by	Description	Primary use(s)
Track managed	Rail infrastructure managers	Measure of the total length of track that is managed by the rail infrastructure manager in both an operational and non-operational state.	Normalising factor. Input into ONRSR operator relative risk prioritisation. Input into risk profiling tool used for regulatory cost recovery (fee) calculations.
Network description – line sections	Rail infrastructure managers	Description of the railway in terms of one or more-line sections, as defined by the rail infrastructure manager, that is used to describe operational portions of the railway network.	Reference information to input into occurrence reporting.
Network description - private sidings	Registered rail infrastructure managers	Description and location (in reference to the relevant adjoining rail infrastructure manager's line sections).	Reference information to input into occurrence reporting.
Network description - characteristics	Rail infrastructure managers	Description of characteristics of the network that assist in describing the nature of operations that occur on the network, e.g. track speeds, traffic density, interfaces with other railways etc.	Reference information to input into occurrence reporting. Input into ONRSR operator relative risk prioritisation. Input into the National Level Crossing Portal (default characteristic data).
Network description – level crossings	Rail infrastructure managers	Location (in reference to the network description – line sections) and characteristics of level crossings on the railway, e.g. traffic control types, operational status, road manager etc.	Reference information to input into occurrence reporting. Input into ONRSR operator relative risk prioritisation. Input into the National Level Crossing Portal. Input into regulatory cost recovery (fee) calculations.

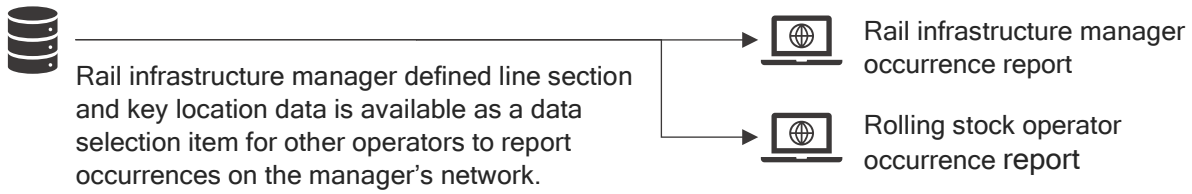
Data type	Reported by	Description	Primary use(s)
Passenger operations interfaces	Rolling stock operators	Count of the number of stations and stops that are used for passenger services.	Input into ONRSR operator relative risk prioritisation. Input into regulatory cost recovery (fee) calculations.
Network description – key locations (optional data, see section 5.2)	Rail infrastructure managers; and Rolling stock operators	Description and location (in reference to the network description – line sections) of key locations, such as stations, siding, yards and loops on a railway that are commonly referred to in describing locations on the railway.	Reference information to input into occurrence reporting, enables a reporting option of providing location description in place of providing line section and track km data.
Railway (train) operations (optional data, see section 5.2)	Rolling stock operators	Description of the nature and scope of railway (train) operations undertaken by a rolling stock operator.	Reference information to input into occurrence reporting. Input into ONRSR operator relative risk prioritisation.

This data is captured at a jurisdictional level to generate normalisation data at an acceptable granularity when used against location-based occurrence data to account for operational differences that occur across Australia. Noting that ONRSR needs to be able to provide jurisdictional specific information if requested.

The detailed dataset description for the above is set out in the ***Network and Railway Operations Data Submission Requirements*** document.

5.1 Location data accuracy - sharing network description data

Rail infrastructure manager line section and, if provided, optional key location data is made available to rolling stock operators as reference data when reporting the location of an occurrence.



The presentation of available line section and key location data to a rolling stock operator can be further tailored if the operator voluntarily provides optional data to identify the railway networks where the operator typically operates.

5.2 Data capture and management

The ontology data is maintained as an ‘as of now’ description of the operator’s network and railway operations characteristics and is:

> submitted on commencement of an operator’s accreditation. The manner and format of the data at the time of initial submission can be agreed between the operator and ONRSR;

> amended as changes occur to the operator’s nature of operations as described by the data. This will occur as a result of variations of accreditation/registration or notification of change processes of the RSNL as appropriate; and

> reviewed annually by a rail transport operator prior to the due date of the operator’s Safety Performance Report submission.

The ontology data under this scheme involves data that is submitted:

> as required by the RSNL to establish data that is needed as input into regulatory cost recovery (fee) calculations and is used for other industry purposes (e.g. the National Level Crossing Portal); and

> voluntarily to establish network and railway operations data that is directly used to provide a tailored notifiable occurrence reporting solution for operators. Submission of this data is optional for an operator, however, without this information operators will be required to provide network and operations information each time an occurrence is reported.

Once the initial data has been submitted by the operator, ongoing review and updates to the data are managed via the ONRSR Portal, which provides for:

> accessing the data to enable its review and amendment, as required;

> adding new data to be entered through a smart-form type interface; and

> clearly differentiating between required data and optional data submission.

6 Rail operations activity (monthly) data

Rail operations activity data provides information on operational and management activities associated with an operator's railway operations for which they are accredited.

The data comprises:

Data type	Reported by	Description	Primary use(s)
Number of rail safety workers	Rail infrastructure managers; and Rolling stock operators	Measure of the size of the rail safety worker workforce for the operator and across Australia. Measured in full-time equivalent or total hours of worker engagement each month.	Normalising factor. For the purpose of generating normalisation data, a national value of 160 hours = 1 FTE worker per month to be applied to generate comparable data.
Drug & alcohol testing activity	Rail infrastructure managers; and Rolling stock operators	Measure of the scale of testing and results of rail safety workers, broken down by worker type; pre- and post-sign on testing and test type.	Normalising factor. Input into ONRSR drug and alcohol testing program. Input into regulatory cost recovery (fee) calculations.
Train operations undertaken	Rolling stock operators	Measure of the total kilometres travelled by passenger, freight and maintenance trains.	Normalising factor. Input into ONRSR operator relative risk prioritisation. Input into regulatory cost recovery (fee) calculations.
Train operations undertaken on a network	Rail infrastructure managers	Measure of the total kilometres travelled by passenger and freight trains on an operator's network.	
Passenger journeys undertaken	Rolling stock operators (passenger)	Measure of the scale of passenger operations being undertaken. Measured as passenger journeys and, where possible, the passenger kilometres (providing a better measure of the extent of passenger travel).	Normalising factor. Input into ONRSR operator relative risk prioritisation. Input into regulatory cost recovery (fee) calculations.
Interfaces – rolling stock operators on a network	Rail infrastructure managers	Identification of the rolling stock operators that operate on the manager's network.	Input into ONRSR operator relative risk prioritisation.

Data type	Reported by	Description	Primary use(s)
Interfaces – rail networks operated on	Rolling stock operators	Identification of the rail networks where the operator has operated trains.	Input into regulatory cost recovery (fee) calculations.

This data is captured:

> at a monthly level to generate data at a granularity that can be used to normalise date-based occurrence data and account for seasonal or other operational variations in railway operations.

> at a jurisdictional level to generate data at granularity that can be used to normalise location-based occurrence data and account for operational differences that occur across Australia. Noting that ONRSR needs to be able to provide jurisdictional specific information if requested.

The detailed dataset description for monthly reporting is set out in the ***Monthly Activity Reporting Requirements*** document.

6.1 Data capture and management

The monthly reported data is due by the 21st of the following month.

Submission of the data is via the ONRSR Portal, which provides:

> entry of data through a smart-form type interface, which includes the ability to use information that has been submitted for previous reporting periods;

> upload capability for drug and alcohol testing activity data[5];

> validation assistance in seeking further information on unexcepted deviations from previously submitted data (where appropriate);

> automated checking for completeness at the due date (in reference to what information is expected and previously provided by the operator), including alerting the operator when the information is not complete; and

> ability to access and amend data that has been previously submitted.

[5] Subject to final ONRSR Portal data submission design.

7 Data Use (National Rail Safety Dataset)

The National Rail Safety Dataset is maintained by ONRSR to provide rail transport operator and stakeholder access to national rail safety data.

7.1 National Rail Safety Dataset

The National Rail Safety Dataset[6] provides stakeholder access to rail safety data at a level of detail that can facilitate the reporting and analysis of rail safety performance by stakeholders at an industry sector level.

The data available in the National Rail Safety Database includes for example:

Notifiable occurrence data	Provided in a de-identified format (described in section 8.2). Contains the majority of the codified data fields that are collected from operator reporting of notifiable occurrences. This means that the National Rail Safety Dataset can be used to access rail safety data associated with an industry sector, occurrence type or network/operational characteristic (where available). However, the National Rail Safety Dataset is not able to be used to obtain data on: > specific rail transport operators; or > specific railway networks or locations.
Network and railway operations (ontology) data	Selected data that can be provided at an industry sector type level, which includes: > total track managed across Australia
Rail operations activity (monthly) data	Selected data that can be provided at an industry sector type level, which includes: > total train kilometres operated; > number of rail safety workers; > extent of drug & alcohol testing that is occurring; > umbers of passengers carried.

7.1.1 Access

The National Rail Safety Dataset is accessed via the ONRSR website: www.onrsr.com.au.

[6] The content of this section has been prepared in anticipation of the National Rail Safety Dataset product.

8 Data governance and disclosure

8.1 Governance and responsibilities

This scheme is jointly owned by ONRSR and the ARA. It is given legal and operational effect through the RSNL, which represents the endorsement and sets the requirements and expectations of ministers and parliaments under the RSNL.

Key responsibilities are:

Role	Responsibilities
ONRSR	> establishing the manner and form for data reporting, as per the RSNL. > decisions relevant to performing the role and functions of the Regulator under the RSNL, including the monitoring and enforcement of legislated data reporting requirements. > provision and maintenance of the primary data capture technology solutions. > administration of the national rail safety data scheme. > maintain access to relevant documentation via www.onrsr.com.au.
ARA	> representing the rail industry data requirements and decisions relevant to the operation of the national rail safety data scheme.
Rail transport operators	> provision of accurate data in accordance with legislated data reporting requirements. > operator level use and analysis of data for safety performance monitoring and risk management decision making.
Rail industry stakeholders	> stakeholder level use and analysis of the data for their specific application. > responsible use and reporting of the data.

Amendment to this scheme will be administratively managed by ONRSR in consultation with the ARA and other stakeholders as required.

8.2 Data management

The data collected under this scheme is reported and supplied by rail transport operators or sourced from publicly available data and is used to:

> provide the Regulator with information necessary to fulfil the RSNL regulatory functions; and

> populate the National Rail Safety Dataset.

To achieve this outcome ONRSR undertakes the data management, validation, quality control, readiness for publication and publication (through the National Rail Safety Dataset). ONRSR manages this data in accordance with its data management principles, confidentiality requirements under s.244 of the RSNL and privacy obligations under the Commonwealth *Privacy Act 1988.*

ONRSR's functions in making the data ready for publication include:

> de-identifying the data by removing organisation names and specific location details and replacing this detail with operator or location type descriptors;

> where multiple rail transport operators are involved in an occurrence, creating an Amalgamated Record, merging the separate reports provided by the operators or other valid sources of information (e.g. ATSB investigation);

> restricting access to commentary type data where this data may contain personal information (privacy) or organisational operational details;

> managing the data quality and accuracy, the latter to the extent practicable as this is largely dictated by the accuracy of reporting by operators when submitting the data; and

> amending the data (principally the Amalgamated Record) to address errors or omissions or where further information becomes available through investigation or other regulatory activity, in order to maintain a current and accurate dataset.

8.2.1 Application of s.244 of the Rail Safety National Law (confidentiality)

Section 244 of the RSNL has the effect of limiting the circumstances where ONRSR can disclose or provide data that it has obtained in exercising any power or function under the RSNL.

The Regulator has functions (in context with this scheme) under the RSNL to work with rail transport operators and others involved in railway operations to improve safety nationally, to collect and publish information relating to rail safety and promote and coordinate the sharing of information to achieve the safety objectives of the RSNL. The publication of rail safety data that has been collected by ONRSR in the form of a National Rail Safety Dataset is done to enable the Regulator to discharge these functions.

In balancing the confidentiality provisions of the RSNL (which is important for maintaining effective communication between operators and the Regulator) and making meaningful data available for rail transport operators and stakeholders, the dataset forming the National Rail Safety Dataset will contain:

De-identified data	Operator names will be replaced with an industry sector type descriptor, which will allow data on similar type operators to be grouped but will not directly identify the operator. Industry sector type descriptors will be maintained and managed by ONRSR. Network locations will be replaced with a generic location descriptor, which will allow data on similar network types/locations to be grouped but will not directly identify the operator. Generic location type descriptors will be maintained and managed by ONRSR. Data fields collected by ONRSR that can readily be used to identify operators, such as train identification fields will not be included in the national rail safety dataset.

Codified data	The dataset will principally consist of codified data and will not include commentary type information that has the potential to have information relating to actions or omissions of workers or the operator in relation to the effectiveness of a safety management system. Note: Occurrence descriptions that are provided at the time of the initial reporting of Category A and B notifiable occurrences (section 4.2.4) will not form part of the published national rail safety dataset. Sharing of the description information is limited to directly involved rail transport operators to enable consistency of reporting, ONRSR and the ATSB to performing their regulatory or investigative role and functions respectively.

8.2.2 Implications of the Privacy Act 1998

The *Privacy Act 1998* applies to information and data held and managed by ONRSR. This ensures that personal information is not unnecessarily held, or inappropriately used or published.

The nature of the data collected under this scheme means personal information on individuals is not required. Operators are asked to avoid submitting any data or information that is personal or identifies an individual.

Appendix A - Why the change, for those that know the history

Implementation of this scheme marks a fundamental change to the national approach for the collection of - and access to - rail safety data. Effectively a generational change, it adopts a green field view to what data is needed at a national level and how it should be collected and managed.

Rail safety data (with a focus on occurrences) has been collected by rail safety regulators since the establishment of rail safety regulatory regimes across Australia from the late 1990s. While having a common objective, the collection of this data was established independently by each jurisdiction. Through the period leading to the establishment of ONRSR, regulators, with industry involvement, had collaborated to progressively align and improve (quality and content) the data to establish a dataset that could be used at a national level.

While further progress to improve the quality and consistency of the data has been made since the establishment of ONRSR, several issues remained with the data framework inherited from the previous regulatory regimes. As the need for high quality, timely data increases, maturity in the rail industry has seen data analytic tools advance and IT platforms enhanced. These developments prompted the National Rail Safety Data Strategy to review and re-shape the approach to meet the requirements for risk-based regulation and good safety decision making into the future.

The following areas of concern were identified, through the work of the *National Rail Safety Data Strategy,* with the historical approach and addressed:

Collecting the right data.	A significant number of occurrences were being reported to ONRSR that did not need to be reported bringing with them a redundant reporting and data validation burden. This data scheme has rationalised the range of occurrences that are reported. This meant that the Regulator moved from collecting lots of detail to collecting more focused data that provides enough information to trigger the need to go to the operator to examine safety issues in detail.
Aggregation of data to present a national picture	While efforts were made to generate a national data the free text nature of key data fields, principally location data, has inhibited the ability to use the data reliably in a location-based sense at a national level due to differences in spelling or the use of different ‘common’ names for the same location. This is addressed with the introduction of the network operations (ontology) data, the sharing of which across operators provide a common and codified dataset that is used for reporting occurrence locations.
Data availability to operators and stakeholders (the risk and safety management decision makers)	The collected data was also principally used by the Regulator for regulatory decision making with limited data made available through reports and publications. ONRSR provided rail safety data to stakeholders via its website for key occurrences and provided reports to level crossing committees, RISSB, TrackSafe, transport agencies as well as several other reports as requested by stakeholders. The establishment of a National Rail Safety Dataset means stakeholders can now access the national data and tailor it for their specific use. This sees ONRSR’s role change from data collection and preparation of structured external reports, to collecting for regulatory purposes and as a facilitator for making industry level data available.

Timing of data reporting	Rail safety data has historically consisted of the data that was provided to the Regulator within a written report required within 72 hours of the occurrence. There was no requirement to update the data if new information became available through investigation or other actions. More time is now provided to gather and provide detailed data associated with the occurrence, and occurrence records may be updated by operators or ONRSR if new information becomes available.
Top event reporting	Historically occurrence data was captured under a top-event principle, meaning rail transport operators had to determine and report what event, in the sequence of events making up an occurrence, presented the greatest adverse outcome in terms of casualties, damage or seriousness. This meant that some detail of an occurrence was lost. For example, a signal passed at danger (SPAD) that results in a collision would be reported as a collision, the fact that a SPAD occurred being lost from the safety data. Data relevant to the sequence of events for a notifiable occurrence is now captured. This enables, in relation to the example above, data analysis of collisions and SPADs within the national rail safety dataset.
Categorisation of occurrences	Operators were asked to categorise occurrences against the classification scheme set out in the *Reporting Requirements for Notifiable Occurrences* document – a 168-page document setting out reporting requirements for the 21 categories and 127 sub-categories. Unfortunately, the classification scheme was often found wanting in supporting data analysis, with the need to return to the free-text description to extract more useful information. Significant time was also spent validating occurrence reports to ensure that the occurrence had been accurately categorised. The national dataset now focuses on the capture of data that can be measured or observed to generate a set of codified data characteristics that describe the occurrence. This enables more flexible data analysis and the generation of data classification systems that are not hard coded into the raw data.

Discontinued Data

Some types of historical rail safety data are no longer being captured under this scheme, of note is data associated with:

> slip, trips or falls on railway premises not associated with train operations;

> railway network security, e.g. alleged assaults, vandalism and trespass; and

> collisions with animals, unless these results in a more significant outcome.

The removal of this data from the national rail safety dataset reflects the lack of historical use of the information by the Regulator (due to the matters being outside of the regulatory reach of the Regulator) or by other stakeholders.

While they are no longer collected at a national level, these events are still monitored by rail transport operators, who are expected to report such events to relevant authorities (i.e. work, health and safety regulators or police) as required.