



# **TIS Data Quality Management**

## **HANDBOOK**

RailNetEurope  
Oelzeltgasse 3/8  
AT-1030 Vienna

Phone: +43 1 907 62 72 00

Fax: +43 1 907 62 72 90

[mailbox@rne.eu](mailto:mailbox@rne.eu)

[www.rne.eu](http://www.rne.eu)

## Content

1	Document history .....	3
2	Abbreviations Used in this Handbook.....	4
3	Foreword .....	5
4	Introduction.....	6
5	Data Quality Management Process.....	7
5.1	Actors and responsibilities .....	7
5.1.1	RNE.....	8
5.1.2	Infrastructure Managers members of TIS.....	8
5.1.3	TIS Technical Board (TB) and Change Control Board (CCB).....	9
5.1.4	RNE Performance Management WG (TPM WGs and Reporting CBB) .....	9
5.1.5	Railway Undertakings .....	9
5.2	Procedure.....	9
5.2.1	Overview.....	10
5.2.2	Details .....	11
5.3	Performance monitoring.....	11
6	Data Defect Indicators .....	12
7	Reports.....	13
8	Annex 1: Data Quality Contacts list.....	14

## 1 Document history

**Table 1: Document history**

Version	Author	Date	Changes
00.1	Simona Di Loreto	02/10/2013	First draft
01.0	RNE	04/10/2013	Revised by RNE JO
01.1	RNE	14/10/2013	Text added on page 4 according to SBB) ´s (Karl Guntern) proposal
01.2	RNE	04/11/2013	Comments received before the OAS meeting on 14-11-2013 (AP, SN,PvK) + additional items by RNE IT staff
02.0	RNE	12/11/2013	Modification according to decision of the OAS on November 14th 2013
02.1	RNE	13/01/2014	Revision before the meeting
3	WG	21/01/2014	Revision during the meeting
4	RNE	13/02/2014	Final draft
4.1	RNE	24/03/2014	Comments by sub-group
4.2	RNE	28/07/2014	Minor modifications; update of annexes and appendixes; agreed by the DQCs in the meeting of July 24th 2014
5	RNE	10/02/2015	Overall revision
5.1	RNE & DQCs	06/03/2015	Modifications after remarks and DQCs meeting (25/02/2015)
5.2	RNE & DQCs	28/03/2015	Approved version with fine tuning after written remarks by DQCs. Fine-tuning of the glossary
5.3	RNE & DQCs	28/05/2015	Revision during the DQCs meeting
6.0	RNE	23/10/2017	Complete revision of Handbook including introduction to new setup and TIS DQ Management procedures

## 2 Abbreviations Used in this Handbook

**Table 2: Abbreviations used in Handbook**

Abbreviation	Explanation
CMS	Content Management System
CTT	Contracted Timetable
DDI	Data Defect Indicator
DQC	Data Quality Contact
DQH	Data Quality Handbook
IM	Infrastructure Manager
OBI	Oracle Business Intelligence
RA	Running Advice
RFC	Rail Freight Corridor
RNE	RailNetEurope
RNE BM	RNE Board Members
RNE GA	RNE General Assembly
RU	Railway Undertaking
TIS	Train Information System
TIS CCB	TIS Change Control Board
TIS TB	TIS Technical Board
TPM	Train Performance Management

### 3 Foreword

RNE manages the Train Information System (TIS) a web-based tool which provides information about international train runs. Besides the use as self-standing system (using a graphical interface) the available (raw) data-sets are also used by some IMs and RUs as well to feed their own company systems. Furthermore, data extracted from the database are used to create reports through additional tools, such as Oracle Business Intelligence (OBI). As TIS is fed by different national systems data quality problems may occur at different stages:

- » Delivery of incorrect data by national systems
- » Wrong or missing data transfer from national systems into TIS
- » Incompatible data provision from different national systems concerning the same train
- » Incorrect data procedure or data loss within TIS
- » Incorrect data-transfer to subsystems (like OBI)
- » Wrong data processing or data loss in subsystems

To provide reliable information at all these stages data quality checks must be performed.

TIS data are used by several bodies (RNE groups and related working groups, for example the Train Performance Management group of Rail Freight Corridors) and “TIS data quality” is mentioned in different frameworks, with different meanings. This causes a lot of confusion and an inefficient use of resources.

To decrease workload and to ensure a high data quality level RNE has proposed to the different involved partners to draw up a structured framework to define the different meaning of data quality and the related procedures (reports, evaluation, corrective actions, responsible, KPIs).

At this aim, a Working Group has been set up and steered by RNE.

The Working Group has delivered this handbook (from now DQH) containing the following topics:

- » Description of the actors involved in TIS data quality management
- » Definition of the different aspects of data quality according to the use of the data
- » List of indicators that should be used to measure data quality and related reports
- » Procedure for the TIS data quality management
- » Additional remarks

#### Note:

The Handbook has been completely revised by the internal RNE Staff in October 2017, based on new IT development and of the experience and lessons learnt. The Data Quality Contacts have discussed the changes and proposed new modifications. The final version of the handbook needs to be approved by the end of 2017.

## 4 Introduction

The framework proposed by this document is aimed to provide a structured procedure to check the quality of the data, according to the needs emerged in the past.

The setting of new/additional targets or standard for data quality is not an aim of this project and therefore they are not mentioned in this document.

In addition, it must be underlined that this document does not attribute additional requirements or additional obligations to the partners. As said above, the aim is to provide rules and tools to check whether the existing requirements have been met and the existing obligations have been fulfilled. If this is not currently or not fully the case in some IMs, there could be implications on processes and/or additional resources could be required from the national side. These issues, however, cannot be investigated at a general level and the needed estimation should be performed internally and singularly by the involved IMs.

## 5 Data Quality Management Process

DQ WG is composed of experts on data quality involved in action of deep analyses of TIS data and relations between them. Group analyses problems and forwards the tasks/corrective actions to the correct responsible partner.

The basic scheme for the data quality management procedure.

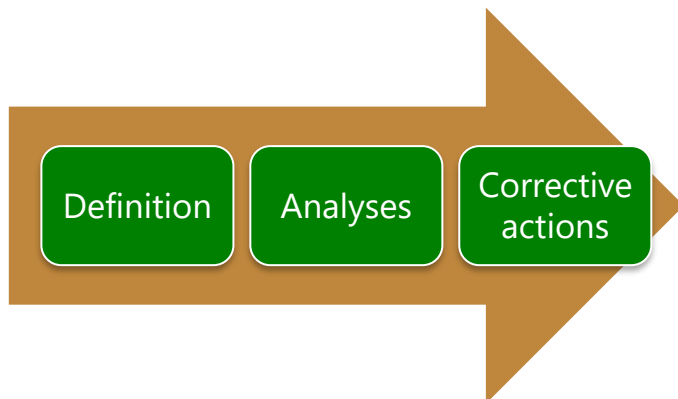


Figure 1: Basic scheme for Data Quality Management Procedure

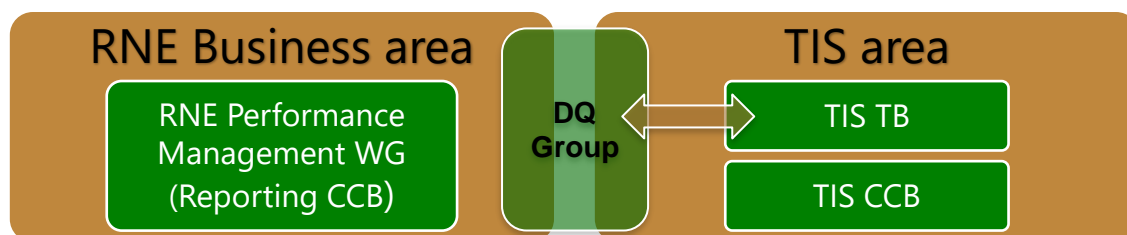


Figure 2: Relation of data quality with other areas

Performance Management working group is fully responsible for TPM and any OBI developments. TIS CCB is fully responsible for TIS development. DQ Group is intermediate step before possible full integration of DQ responsibility to TIS TB. Merging of tasks of DQ and TIS TB to be considered in the future.

### 5.1 Actors and responsibilities

TIS is a system managed by RNE, therefore the main responsibility to grant that TIS provides high data quality rely on RNE. In this context, RNE must take care that the system is always running in an efficient and correct way. To do so, RNE performs checks and keeps contacts with the IT supplier.

However, RNE needs the input and the support of the interested partners in all phases of the process.

In the definition phase, the users of TIS should define together with RNE which indicators and which reports should be used to measure data quality, according to their needs and the activities in which they use TIS data.

The minimum level of data quality to be provided by TIS must be decided by the users of the tool and this should be done in the definition phase as well.

In the analyses and corrective actions phases, RNE has the main role. However, the fact that data are provided by the IMs which directly send them from their national systems, must be considered. When data sent by an IM are not complete or correct, on one hand, RNE has limited possibilities to find out the cause(s). In addition, most of the times, RNE is not able to solve problems internally caused by the IM's national systems or procedures.

Therefore, both TIS users and data providers (IMs) must be deeply involved not only in the definition phase (i.e. the setting up of the data quality framework) but also in the procedures for checking and improving the data quality.

RNE shall maintain the leading and coordinating role and it shall involve the concerned partners according to the phase of the above-mentioned procedures and to the type of problems found out through the quality checks.

The “concerned partners” are the representatives of the IT departments of the IMs in the TIS Technical Board and Change Control Board (TIS TB and CCB), the IMs contact partners in the operational and planning departments, the responsible persons within the TPM organisations and the RUs contacts.

However, to make this process as efficient as possible, it is preferable that RNE is mainly dealing with only one partner, a contact-point ensuring that the national partners having the possibility to check or solve the identified data problems are duly informed and, in turn, take the necessary actions to fix the causes of insufficient data quality.

The Data Quality working group proposes that such partner is network of “single supervising contact person” for TIS data quality issue (afterwards “DQ Contact person”- afterwards DQC), one for each IM.

### 5.1.1 RNE

In general, RNE is asked to steer the process of data quality management. Specifically, RNE must:

- » Develop and maintain the pre-defined reports in OBI
- » Produce the reports or make them available to users
- » Perform a first check of the data quality
- » Involve the responsible partners when problems occur and support them in the analysis to find out the causes
- » Keep contacts with the IT supplier and ensure that the technical problems are duly solved

Most of the items listed above, are part of the services portfolio offered to IMs organisations dealing with the Train Performance Management (within or outside RFCs organisations) described in the document “TPM Cooperation Manual”.

Therefore, the process described here must consider the procedures described in the TPM Manual.

### 5.1.2 Infrastructure Managers members of TIS

Data quality could be caused by IT problems, by not suitable procedure in the planning or in the operation phases, therefore involvement of all the related national company departments is needed. In these cases, each IM must grant that both analysis and corrective actions are performed in due time and efficiently. At this aim, each IM nominates a “single supervising contact person” for TIS data quality issue (DQC) whose tasks will be:



- » Being informed by RNE about all TIS data quality issues
- » When necessary, forwarding the request of action/clarification to internal responsible persons and forwarding to RNE their feedback
- » Monitoring the internal problem-solving process, helping RNE to keep contact with the interested persons

Of course, the IMs, according to their internal organisation, are free to nominate several “TIS department responsible persons”, one for each concerned internal departments to be involved according to the type of problem, but this is a company choice and it is not dealt with in this document.

The effects on internal processes of the application of this procedure and the evaluation of the related work-load cannot be included in this paper, because they depend on the internal organisation of the single IMs and it is not possible to make any general estimation.

### 5.1.3 TIS Technical Board (TB) and Change Control Board (CCB)

The TIS TB and the TIS CCB will continue to carry out their normal tasks in the management of TIS, therefore they will be involved in the data quality management according to their defined responsibilities, specifically in those cases when the data problem identified is caused by a technical issue.

- » The TIS TB role is to support RNE in identification of the causes of the problems, the definition of the possible solution. The TIS TB members are also responsible for the application of the corrective actions when the cause of the problem originates from an IM system or interface. The TIS TB shall be kept informed about the status of data quality in TIS by RNE
- » The TIS CCB has normally no active role in the data quality management, but it shall be kept informed by RNE. It might be required to approve specific corrective actions when these have an influence on TIS budget.

### 5.1.4 RNE Performance Management WG (TPM WGs and Reporting CBB)

The RNE Performance Management WGs do not have any responsibilities in checking data quality in TIS, but currently being the main users of train run data coming from TIS, they:

- » Can require or take care of the inclusion of train numbers they are monitoring in their Train Performance Management tasks;
- » Will be regularly informed by RNE about data quality issues and will use the information for their own analyses
- » May point out data quality problems if they find some that are not detected
- » Can take part in the DQCs

### 5.1.5 Railway Undertakings

The Railway Undertakings are involved in the data quality management within the TPM WGs, where existing.

## 5.2 Procedure

Data quality management procedure is applied monthly. Once the process proves to go smoother (no reporting problem, communication flows up and running, updates procedures fully working) this

section will be re-discussed and a quarterly procedure will be proposed. The aim of the procedure is to keep data quality in TIS under constant control.

### 5.2.1 Overview

The procedure consists of the following steps:

- » RNE makes general checks on the data available at the beginning of each month regarding the previous month's train runs
- » RNE produces a data quality report and sends it to the DQCs (and other recipients if so decided by the DQCs group) by seventh (7<sup>th</sup>) day of the beginning of each month
- » More detailed analyses will be done upon request by the DQCs to target the problem more efficiently
- » All partners check the report and react to urgent problems
- » RNE populates the Shortcoming List based on data quality reports created
- » All partners check the Shortcoming List and provide feedback on problem indicated by twentieth (20<sup>th</sup>) day of the beginning of each month
- » Regular meetings are organised where data quality and reporting problems are discussed

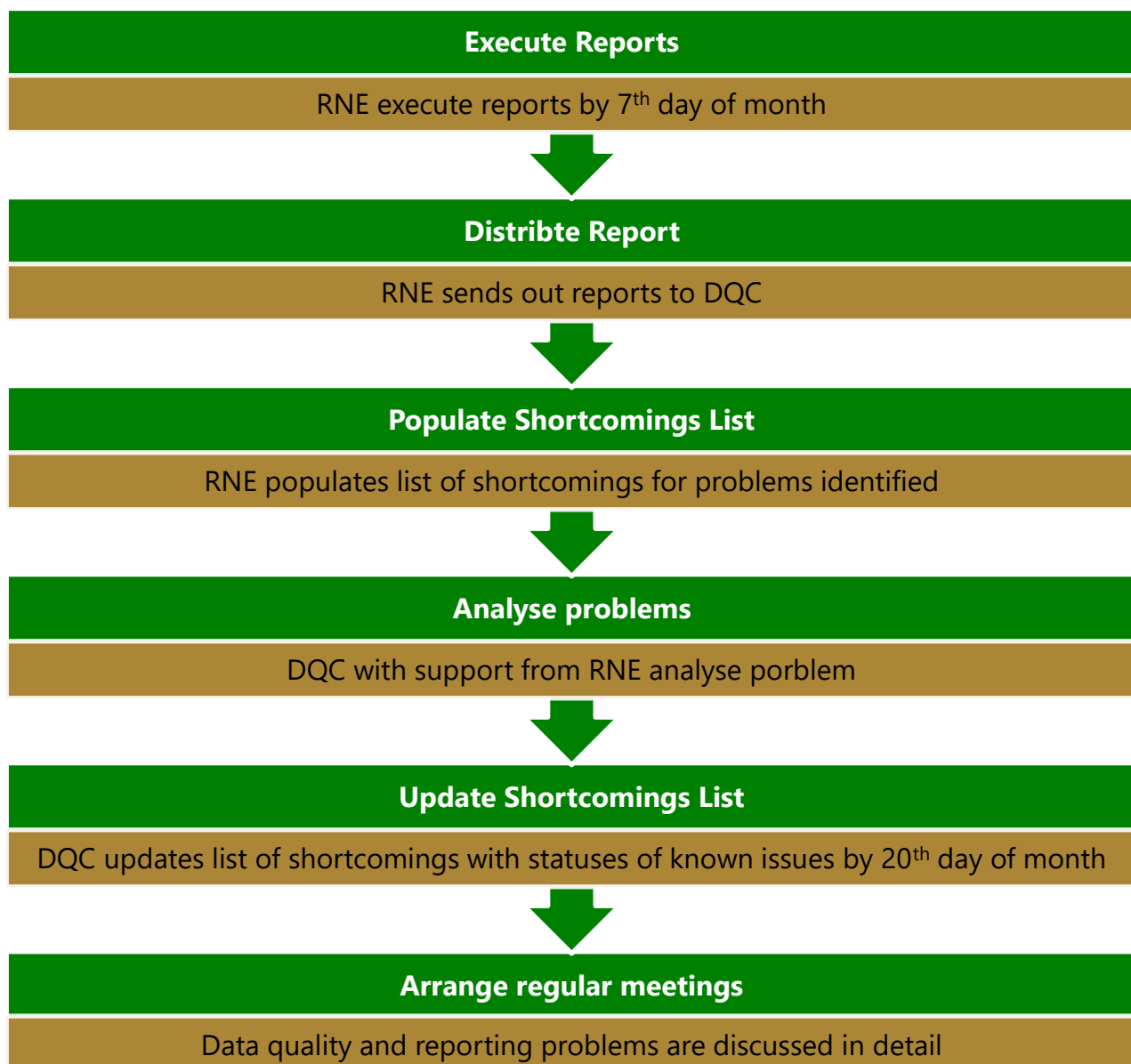


Figure 3: Data Quality Management Procedure

### 5.2.2 Details

Regular meetings will be held twice a year.

The goals of the meetings are:

- » Discuss of data quality and reporting problems
- » Identify the cause of the problems where RNE is not able to do it with single DQC
- » Discuss together the problems that occur more often to be avoided in the future
- » Update of DQH and procedures related to better management of data quality

The DQCs have access to the OBI tool, where they can directly execute and download reports specifying different parameters as wanted.

More detailed analyses can be done upon request of DQC to RNE.

The outcome of the problem identification and analysis shall be a list of problems, shortcomings list. For the list of shortcomings online web based tool will be used, CMS. Every DQC is obligated to use this tool and update status of any know issue listed. Also, if identified by DQC list of shortcomings can be populated with new issue.

### 5.3 Performance monitoring

To improve quality of data more efficiently and to improve problem solving related to data quality, RNE will create additional reports that will indicate performance of solving data quality issues. These reports will be based on data from TIS and on data from CMS tool for List of Shortcomings.

Reports will show increase/decrease of data quality in general and engagement of actors and responsible entities, over a certain period.

Performance monitoring reports will be presented to RNE BM and RNE GA to evaluate work efficiency of Data Quality Working Group.

## 6 Data Defect Indicators

One of main functions implemented in TIS is DDIs processing function. It consists on the processing of data coming from TIS to identify those train runs affected by data quality problems, i.e. “data defects”.

Currently sequence of the points is not known to TIS, there is no information about consecutive points, distance between the points and load shifting calculation function it not working as expected. Therefore, for data quality reporting purposes only these indicators can be used:

- » Completeness of CTTs
- » Completeness of RAs
- » Undocumented delays

In addition, information about linking trains from TIS will be used in data quality reports as well.

## 7 Reports

A monthly report is produced by RNE and sent to the DQC each month. It is a PDF file containing the indicators previously described in this document.

Each report has multiple parameters to be chosen from therefore each user can additionally execute every report using different parameters to identify problem more easily and to help to propose a possible solution.

RNE will send out PDF reports with general parameters, but it is strongly advised that every DQC executes reports on its own in addition using different parameters if it will contribute in identifying and solving a data quality issue.

Currently there are four (4) reports available in OBI that can be found on this location:

### Shared Folder/TIS/Data quality reports

List of currently available reports:

**Table 3: List of data quality reports in OBI**

Report name	Description
DDI Completeness of CTTs	Showing top 10 train numbers with highest number of missing contracted timetables and all points that have missing contracted timetables. Report is customizable by using different available parameters.
DDI Completeness of RAs	Showing top 10 train numbers with highest number of missing running advices and all points that have missing running advices. Report is customizable by using different available parameters.
Linking Trains	Showing all train numbers that are candidates to be linked and train numbers that are manually linked. Report is customizable by using different available parameters.
Undocumented Delays	Showing top 10 points with highest percentage of undocumented delays at arrival and departure. Report is customizable by using different available parameters.

More detailed description of each report can be found within folder in OBI.

## 8 Annex 1: Data Quality Contacts list

Table 4: List of data quality contacts

Name	Surname	Email	Company
Eduardo	Martinez	emmart@adif.es	ADIF
Félix	Bartolome	fbartolomea@adif.es	ADIF
Francisco	de la Vega	fjvega@adif.es	ADIF
			Bane NOR
			Banedanmark
Alexander	Paulus	alexander.paulus@bls.ch	BLS Netz AG
			CFL
			CFR
Siegfried	Nierichlo	siegfried.nierichlo@deutschebahn.com	DB Netz AG
Csaba	Németh	csnemeth@gysev.hu	GYSEV ZRt
Mate	Marić	mate.maric@hzninfra.hr	HŽI
Ann	Verstraelen	ann.verstraelen@infrabel.be	INFRABEL
			IP
			Lietuvos Geležinkeliai AB
			LISEA
László	Fenyves	fenyvesl@mav.hu	MÁV
Balázs	Lőrincz	lorincz.balazs@mav.hu	MÁV
			MŽI
			NRIC
Georg	Mayer	georg.mayer@oebb.at	ÖBB
			PKP-PLK
Peter	Vanwaveren	peter.vanwaveren@prorail.nl	PRORAIL
Roberto	Caruso	r.caruso@rfi.it	RFI
Karl	Guntern	karl.guntern@sbb.ch	SBB AG
			SNCF Réseau
Richard	Svoboda	svobodar@szdc.cz	SŽDC
Boris	Žebalc	boris.zebalc@slo-zeleznice.si	SŽ-I
Per-Anders	Winge	per-anders.winge@trafikverket.se	Trafikverket
Igor	Vojtek	vojtek.igor@zsr.sk	ŽSR