



Handbook for TIS Data Quality Management

Delivered by	WG TIS Data Quality Framework- 24/03/2014
Approved by	

RailNetEurope
Oelzeltgasse 3/8
AT-1030 Vienna

Phone: +43 1 907 62 72 00

Fax: +43 1 907 62 72 90

mailbox@rne.eu

www.rne.eu

Content

Change history.....	4
Abbreviations - names	5
Foreword	6
1 Introduction.....	7
2 Current situation	8
3 Definitions.....	8
4 Indicators & Reports	9
4.1 Basic data quality.....	11
4.1.1 Fulfillment of the SLA - Downtime (web application)	11
4.1.2 Fulfillment of the SLA - Downtime (data exchange) and performance (processed messages)	11
4.1.3 Performance (Messages-Total).....	11
4.1.4 Performance (Messages-TAF/TAP TSI)	11
4.1.5 Performance (Completeness of data)	12
4.1.6 Key information.....	12
4.2 Performance Analyses – completeness	12
4.2.1 Compare running advice with CTT	12
4.2.2 Completeness of Running Advices	12
4.2.3 Completeness of CTTs	12
4.2.4 Delay causes	13
4.3 Performance Analyses – correctness.....	13
4.3.1 Missing responsible company	13
4.3.2 Consistency of CTTs at border/hand-over point.....	13
4.3.3 Consistency of RAs	14
4.4 Train Performance Management Report	14
5 Data quality management process.....	15
5.1 Actors and responsibilities	15
5.1.1 RNE.....	17
5.1.2 Infrastructure Managers members of TIS.....	17
5.1.3 TIS Technical Board (TB) and Change Control Board (CCB).....	17
5.1.4 Train Performance Management WGs (RFCs)	17
5.1.5 Railway Undertakings	18
5.2 Processes.....	18
5.2.1 Monthly procedure	19
5.2.2 Quarterly procedure	20
6 Update procedure	23

7	Short Glossary	23
8	Annexes.....	24
9	Appendix.....	24
9.1	Appendix 1 – List of Data Quality Contacts	25
9.2	Appendix 2 – List of Data Defect Indicators and short explanation.....	26

Change 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 14 th 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 24 th 2014

Abbreviations - names

Explanations/definitions of some items can be found in the rest of the document or in section 7 "Short Glossary"

CCT: Contracted Timetable
DDI: Data Defect Indicator
DQH: Data quality Management Handbook
EPR: European Performance Regime
EPR OWG: EPR Operation Working Group
IM: Infrastructure Manager
OAS WG: Operations & After-Sale Working Group
OBI: Oracle Business Intelligence
RA: Running Advice
RFC: Rail Freight Corridor
RU: Railway Undertaking
SLA: Service Level Agreement
TAP/TAF: Telematic Applications for Passenger/Freight
TIS: Train Information System
TIS CCB: TIS Change Control Board
TIS TB : TIS Technical Board
TPM: Train Performance Management
TPM PMs : TPM Performance Managers
TSI: Technical Specifications for Interoperability
UIC: Union Internationale des Chemins de Fer

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 quality problems may occur data at different stages of the TIS process:

- » 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 have to 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 1, 2, 3 and 6 at the moment): “TIS data quality” is mentioned in different frameworks, with different meanings. This causes a lot of confusion and an inefficient use of resources. In a future perspective, with more Rail Freight Corridors carrying out TPM tasks using TIS data, these inefficiencies risk to increase.

In order to decrease workload and to ensure a high data quality level RNE has proposed to the different involved partners (TIS TB & CCB, Train Performance Management and OAS experts) 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. It was steered by RNE and composed by the following persons:

- » RFI – Roberto Caruso
- » SBB – Karl Guntern
- » ÖBB – Georg Mayer
- » DBNetz – Siegfried Nierichlo
- » Prorail – Pieter Vanklaveren
- » RFF – Frédéric Richard
- » BLS – Alexander Paulus
- » INFRABEL – Ann Verstraelen
- » Trafikverket - Per-Anders Winge

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

- » Description of the actors involved in TIS data quality management with a proposal for improvement
- » 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

1 Introduction

As mentioned in the foreword, the expression “TIS data quality” can have different meanings according to the context where it is used.

For example, in the European Performance Regime project (joint RNE/UIC project carried out between 2003 and 2012) traffic data were used to calculate penalties, therefore data had to be fully correct and complete (100% data quality needed). When data are used for statistics on punctuality and delays, they must have a high quality (to grant reliability of the statistics) but it is not necessary that each single piece of information is verified and its completeness/correctness ensured. If the data quality checks are performed to verify the general performance and efficiency of TIS, then these checks are less strict and demanding.

Also the actors involved in the phases of the data quality management (definition, check, problems identification, corrective actions) are different.

To grant a clear understanding, the Data Quality Framework working group (afterwards WG) has described the different fields of action where the management of TIS data quality is an important task.

1. Basic data quality
2. Performance analyses
3. TPM reports

For each of the items above, the WG has identified:

- » Definitions (section 3)
- » Indicators and reports (section 4)
- » Actors and responsibilities (section 5.1)
- » Procedures for data quality management (section 5.2)

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.

If so decided by the stakeholders, in the future, the procedure described here could be integrated with an additional phase or step where the existing data quality standards are revised: in any case the goal would be to define the procedure to revise the standards not to set the standards themselves.

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.

2 Current situation

Table 1 summarizes the situation regarding the data quality management in TIS until today.

FIELD	USE	What	Checked by
TIS technical checks	TIS maintenance	Monthly reports	TIS TB & RNE
TPM	Keep TPM reliable	Formula calculating a “data quality indicator” for certain points	TPM PMs & RNE
EPR	Deeper analysis of data quality	“DDIs”= data defect indicators (formerly: exclusion rules)	EPR OWG ¹ & RNE

Table 1 – TIS data quality status quo

The following sections will describe the way the framework for data quality in TIS shall be changed.

3 Definitions

The definitions of data quality are described in Table 2.

FIELD	Definition
BASIC QUALITY	Completeness of data sent by each IMs related to certain points and measured on all TIS trains; it includes system availability and performance
PERFORMANCE ANALYSIS	Completeness of data for each train run related to certain points and measured on sample trains Correctness and consistency data for each train run related to certain points and measured on sample trains
TPM REPORTS	Minimum level of reliability of the reports granted applying a formula which calculates a “data quality indicator” (%) on certain points

Table 2 – definition of data quality according to the fields

¹ The EPR OWG was closed in December 2012, when the project was finalized
Version 04.2

4 Indicators & Reports

Table 3 summarizes the indicators & reports that will be used to measure data quality as defined in table 2.

FIELD		INDICATORS	REPORTS
BASIC QUALITY		Keeping the SLA (defined KPIs in TIS member contract - it includes system availability and performance) <ul style="list-style-type: none"> • Downtime (web application) • Downtime (data exchange) • Performance (Messages-total) • Performance (TAP/TAF TSI messages) Completeness of data sent by each IMs related to certain points and measured on all TIS trains	Using the existing Quarterly Report – based on the TB and CCB procedure.
		Key information of the messages content are correct (different point time at the transfer point)	Formula developed by RNE– details after
PERFORMANCE ANALYSIS	COMPLETENESS	Compare Running Advices (RA) with CTTs	Formula developed by RNE – details after
		Completeness of RAs	EPR tool
		Completeness of CTTs	OBI SE
	CORRECTNESS	Completeness of delay causes	EPR tool
		No responsible company	EPR tool
		Consistency of CTTs at the border/handover points	EPR tool
	Consistency of RAs at defined points	EPR tool	
TPM REPORTS		<i>Not relevant</i>	Formula newly developed by the TPM

Table 3 - summary of data quality indicators and reports

The following sections give details about each of the above listed indicators. In particular, they contain:

- » Detailed explanation of the indicator
- » Report:
 - Tool used
 - Formula (if relevant)
 - Points/statuses where the check are performed
 - Sample
 - Time frame
 - View

4.1 Basic data quality

4.1.1 Fulfillment of the SLA - Downtime (web application)

- » Explanation of the indicator: ; the maximum downtime (when not available) of the web application is 1.5% (not including regular maintenance time needed to secure or improve normal system operation and availability, which is part of the general Service Level); the indicator shows whether this percentage is respected
- » Report:
 - Tool: RNE Monitoring tools (+ manual final check)
 - Formula: shows the share of the downtime during a certain period
 - Points/statuses: n.a
 - Sample: n. a.
 - Timeframe: *per* month
 - View: n. a.

4.1.2 Fulfillment of the SLA - Downtime (data exchange) and performance (processed messages)

- » Explanation of the indicator: the maximum downtime (when not available) of the data exchange is 1.5% (not including regular maintenance time needed to secure or improve normal system operation and availability, which is part of the general Service Level); in addition, if more than 4000 messages are queued for more than 10 minutes the SLA is not reached; the indicator shows whether this percentage is respected
- » Report:
 - Tool: RNE Monitoring tools (+ manual final check)
 - Formula: shows the % of SLA fulfilled
 - Points/statuses: n.a
 - Sample: n.a
 - Timeframe: *per* quarter
 - View: *per* partner (IMs and RUs receiving data) .

4.1.3 Performance (Messages-Total)

- » Explanation of the indicator: count of messages per IM (RA, Delay causes, CTT)
- » Report(s):
 - Tool: RNE Monitoring tools
 - Formula: Count of Messages
 - Points/statuses: n.a
 - Sample: n. a.
 - Timeframe: *per* quarter
 - View: *per* IM, compared to average of the previous two quarters

4.1.4 Performance (Messages-TAF/TAP TSI)

- » Explanation of the indicator: count of TAF/TAP Messages between partners (inbound/outbound)
- » Report:
 - Tool: RNE Monitoring tools
 - Formula: Count of Messages
 - Points/statuses: n.a
 - Sample: n. a.
 - Timeframe: *per* quarter
 - View: *per* partner , compared to average of the previous two quarters

4.1.5 Performance (Completeness of data)

- » Explanation of the indicator: completeness of data on the train run sent by each IMs
- » Report:
 - Tool: RNE Monitoring tools
 - Formula: number of RA messages and of CTTs messages and ratio between them, excluding train runs where no RA has been sent
 - Points/statuses: all points (upon decision of the IM, some points can be skipped)
 - Sample: all TIS trains
 - Timeframe: per quarter
 - View: *per IM*

4.1.6 Key information

- » Explanation of the indicator: list of cases where the mandatory information (UIC/TAF TSI messages) as basis for merging of different messages types (CTT, Running advices, delay codes ...) is not correct (if the transfer-point time in the CTT message and in the RA message are not identical, the information cannot be merged).
- » Report:
 - Tool: OBI SE
 - Formula: specific formula developed by RNE
 - Points/statuses: defined points
 - Sample: all TIS trains
 - Timeframe: monthly
 - View: *per worst point*

4.2 Performance Analyses – completeness

4.2.1 Compare running advice with CTT

- » Explanation of the indicator: completeness of data on the train run sent by each IMs (same as 4.1.5, with different sample, timeframe and view)
- » Report:
 - Tool: OBI SE
 - Formula: number of RA messages and of CTTs messages and ratio between them, excluding train runs where no RA has been sent
 - Points/statuses: defined points
 - Sample: specific train numbers decided by the user (currently: TPM trains)
 - Timeframe: monthly
 - View: by points/most critical cases

4.2.2 Completeness of Running Advices

- » Explanation of the indicator: number of train runs where one or more RAs are missing and related ratio
- » Report:
 - Tool: EPR tool and OBI SE
 - Formula: EPR DDI 6, count of trains where, at one or more defined points, the RA is missing but a CTT exists; Points/statuses: defined points
 - Sample: specific train numbers decided by the user (currently: TPM trains)
 - Timeframe: monthly
 - View: by corridor line

4.2.3 Completeness of CTTs

- » Explanation of the indicator: number of train runs where one or more CTTs are missing
- » Report:
 - Tool: OBI SE
 - Formula: % of the CTTs sent for defined points (n° of CTTs sent/n° of expected CTTs according to registered train runs) – only points where % is <80% are displayed
 - Points/statuses: defined points
 - Sample: specific train numbers decided by the user (currently: TPM trains)
 - Timeframe: monthly
 - View: by line

4.2.4 Delay causes

- » Explanation of the indicator: % of the undocumented delay minutes compared to total sum of additional delay minutes in defined segments
- » Report:
 - Tool: EPR tool and OBI SE
 - Formula: EPR undocumented delays formula (difference between lateness at two consecutive points represents the additional delay, if it is >0. The difference between this additional delay and documented caused delays represents the undocumented minutes if documented delays < additional delay); both number of minutes and *ratio* with total delay minutes are shown
 - Points/statuses: not relevant
 - Sample: specific train numbers decided by the user (currently: TPM trains)
 - Timeframe: monthly
 - View: by line and IM

4.3 Performance Analyses – correctness

4.3.1 Missing responsible company

- » Explanation of the indicator: number of train runs affected by a data mistake which prevents the EPR tool to identify the company, the IM or the RU, responsible for a delay; train runs affected by such problems might be anyway usable for statistics, but the figure provided by this check might be interesting for data quality improvement, because the cause could be:
 - Wrong use of the international delay codes
 - Wrong timestamp (actual time of occurrence)
 - Wrong point/status
 - Missing RU code
 - Delay codes actual time before/after original departure/final destination
- » Report:
 - Tool: EPR tool and OBI SE
 - Formula: EPR DDI 14; train runs where the field “responsible company” (for a specific delay) is empty
 - Points/statuses: defined points
 - Sample: specific train numbers decided by the user (currently: TPM trains)
 - Timeframe: monthly
 - View: by IM

4.3.2 Consistency of CTTs at border/hand-over point

- » Explanation of the indicator: number of train runs where one or more CTTs at border (or hand-over point) is not consistent
- » Report:
 - Tool: EPR tool and OBI SE

- Formula: EPR DDI 1; train runs where the CTTs in the two consecutive border points of neighbouring IMs are not consistent
- Points/statuses: defined border and hand-over points
- Sample: specific train numbers decided by the user (currently: TPM trains)
- Timeframe: monthly
- View: by line, border points

4.3.3 Consistency of RAs

- » Explanation of the indicator: number of train runs where one or more RAs at specific points is not consistent
- » Report:
 - Tool: EPR tool and OBI SE
 - Formula: EPR DDI 5; train runs where the two RAs (arrival/departure) at specific points are inconsistent (arrival after departure) or where arrival to specific point is earlier than the departure from the previous point
 - Points/statuses: defined points
 - Sample: specific train numbers decided by the user (currently: TPM trains)
 - Timeframe: monthly
 - View: by IM, critical points

4.4 Train Performance Management Report

Weighted data quality: the completeness of data should be shown as a total for each reporting point as a percentage (see formula below). If necessary, this formula is applied in sections defined for TPM purposes.

$$Weighted\ DQ = \frac{\sum n^{\circ}\ of\ trains\ \left[\left(\frac{Max\ CTTs * 100}{RA_DQ\ SUM} \right) \right] * Max\ RA_DQ\ SUM}{\sum RA_DQ\ SUM}$$

5 Data quality management process

The basic scheme for the data quality management procedure is described in figure 1:

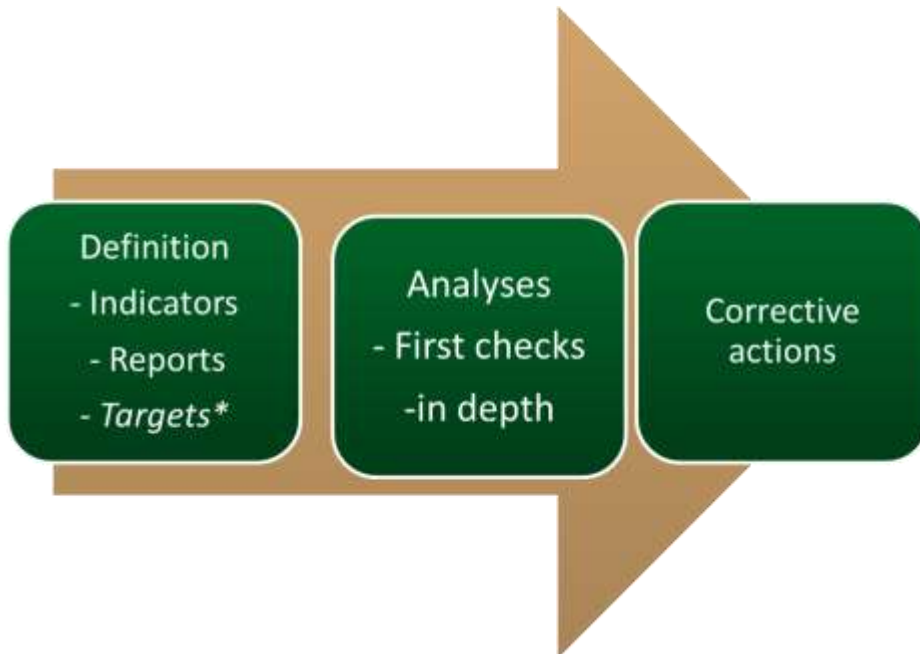


Figure 1 – general data quality management process

*not dealt with in this document (open point for the future)

These steps must be performed in different ways and timeframe according to the type of data quality (see “field of actions” described above).

After the sub-section dedicated to actors and responsibilities (5.1), the processes (monthly and quarterly) for data quality management will be described in more detail (5.2).

The definition phase corresponds to the update of this document and is described in section 6.

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. This task was performed by the sub-group drafting this Handbook and shall be periodically repeated according to the stakeholders’ decision (see section 6).

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. This part is not treated in this document.

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 taken into account. 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 Framework 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. Tasks are described in details in section

The new definition of roles and responsibilities proposed here and summarized in Table 4. Details in the next sub-sections

FIELD	Involved partners			
	Definition	Check	Problem identification	Corrective actions
BASIC QUALITY	RNE, TIS TB and CCB	RNE	RNE, DQCs, TIS TB and CCB	RNE, Concerned IM(s) partners, RNE, TIS TB and CCB
PERFORMANCE ANALYSIS	RNE, users	RNE, users	RNE, DQCs , Concerned IM(s) partners	RNE, Concerned IM(s) partners, DQCs
TPM REPORTS	RNE, TPM PMs	RNE, TPM group	RNE, DQCs , Concerned IM(s) partners, TPM P	RNE, Concerned IM(s) partners, DQCs, TPM PMs

Table 4 – involved actors – summary

5.1.1 RNE

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

- » Create and maintain the reports in OBI SE
- » Produce the reports or in any case 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
- » Coordinate the corrective actions taken to solve the identified problems
- » Keep contacts with the IT supplier and ensure that the technical problems are duly dealt with by it

5.1.2 Infrastructure Managers members of TIS

Since a problem of data quality could be caused by IT problems, by not suitable procedure in the planning or in the operation phases, 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 (afterwards “**DQC Contact person**”) 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 TB and the 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 TB role is to support RNE in identification of the causes of the problems, the definition the possible solution. The TB members also are responsible for the application of the corrective actions when the cause of the problem originate from an IM system or interface. The TB shall be kept informed about the status of data quality in TIS by RNE
- » The 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 Train Performance Management WGs (RFCs)

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

- » 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

5.1.5 Railway Undertakings

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

5.2 Processes

Indicators for both basic data quality and performance analysis will be monitored using the same procedure (see below).

As far as the basic data quality is concerned, and especially the indicators concerning the system availability, the checks will also be done in the day-to-day business and corrective actions immediately taken to restore the availability of the system as soon as possible. This will be done in addition to the processes described below by RNE, supported by TIS TB and CCB.

According to the time frame specified for each indicator in section 4, the reports will have a different content in terms of set of indicators shown, as displayed in detail in table 5:

Field	Indicator number	Description	In report
Basic Data quality	1.1	Fulfilment of the SLA - Downtime (web application)	Quarterly
	1.2	Fulfilment of the SLA - Downtime (data exchange) and performance (processed messages)	Quarterly
	1.3	Performance (messages-total)	Quarterly
	1.4	Performance (TAP/TAF TSI)	Quarterly
	1.5	Performance (completeness of data)	Quarterly
	1.6	Key-information	Quarterly
Performance Analyses-Completeness	2.1	Comparison CTTs/RAs	Monthly
	2.2	Completeness of RAs	Monthly
	2.3	Completeness of CTTs	Monthly
	2.4	Completeness of delay causes	Monthly
Performance Analyses-Correctness	3.1	Missing responsible company	Monthly
	3.2	Consistency of CTTs	Monthly
	3.3	Consistency of RAs	Monthly

Table 5 – content of the reports

5.2.1 Monthly procedure

Figure 2 illustrates the monthly procedure for data quality management. The aim of the monthly procedure is to keep data quality in TIS under constant control. It should deal with urgent problems or problems which can be solved quickly.

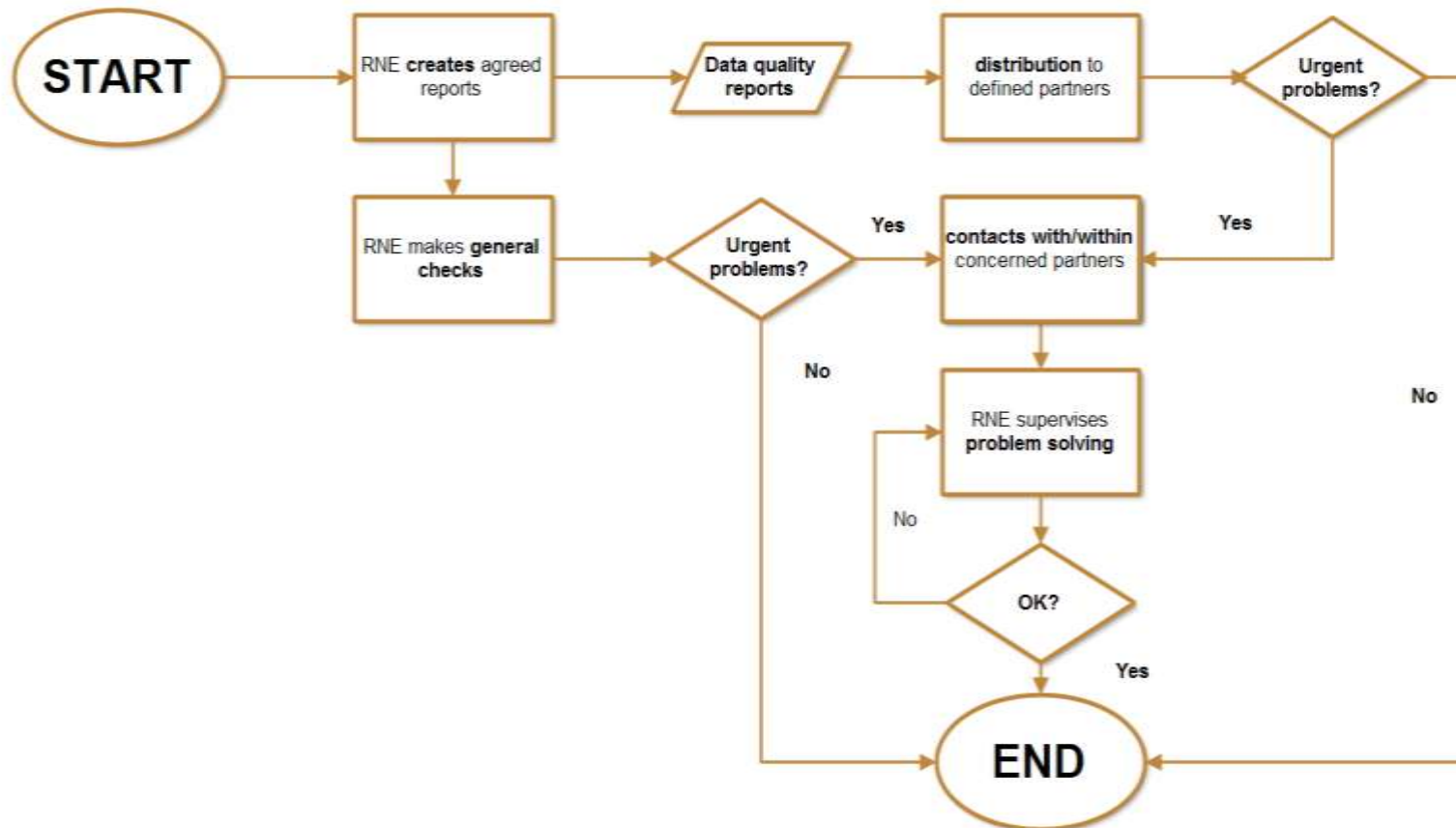


Figure 2 – Monthly data quality management

According to the outcomes of the first checks, RNE contacts the concerned partners on the basis of the principles described in table 6, in order to define the corrective actions to be taken. RNE can contact additional persons if so suggested by the analysis of the problems.

TYPE OF PROBLEM	Addressed Partners	C/C	Notes
Internal IT problem (TIS)	SIGNON	TIS TB and CCB when necessary	
National IT problem	TB	DQCs TIS CCB when necessary	
National problem ≠ IT	DQCs		Interested DQCs shall forward the request to the responsible national partners
National unknown	DQCs		Interested DQCs shall forward the request to the responsible national partners

Table 6 – overview of partners to be addressed

5.2.2 Quarterly procedure

Figure 3 illustrate the quarterly procedure for data quality management. The aim of the quarterly procedure is to:

- » Monitor the status of the problems identified in the monthly procedure
- » Revise the “shortcomings list” (see also after)
- » Deal with problems which are more difficult to solve or require a budget or happen more often

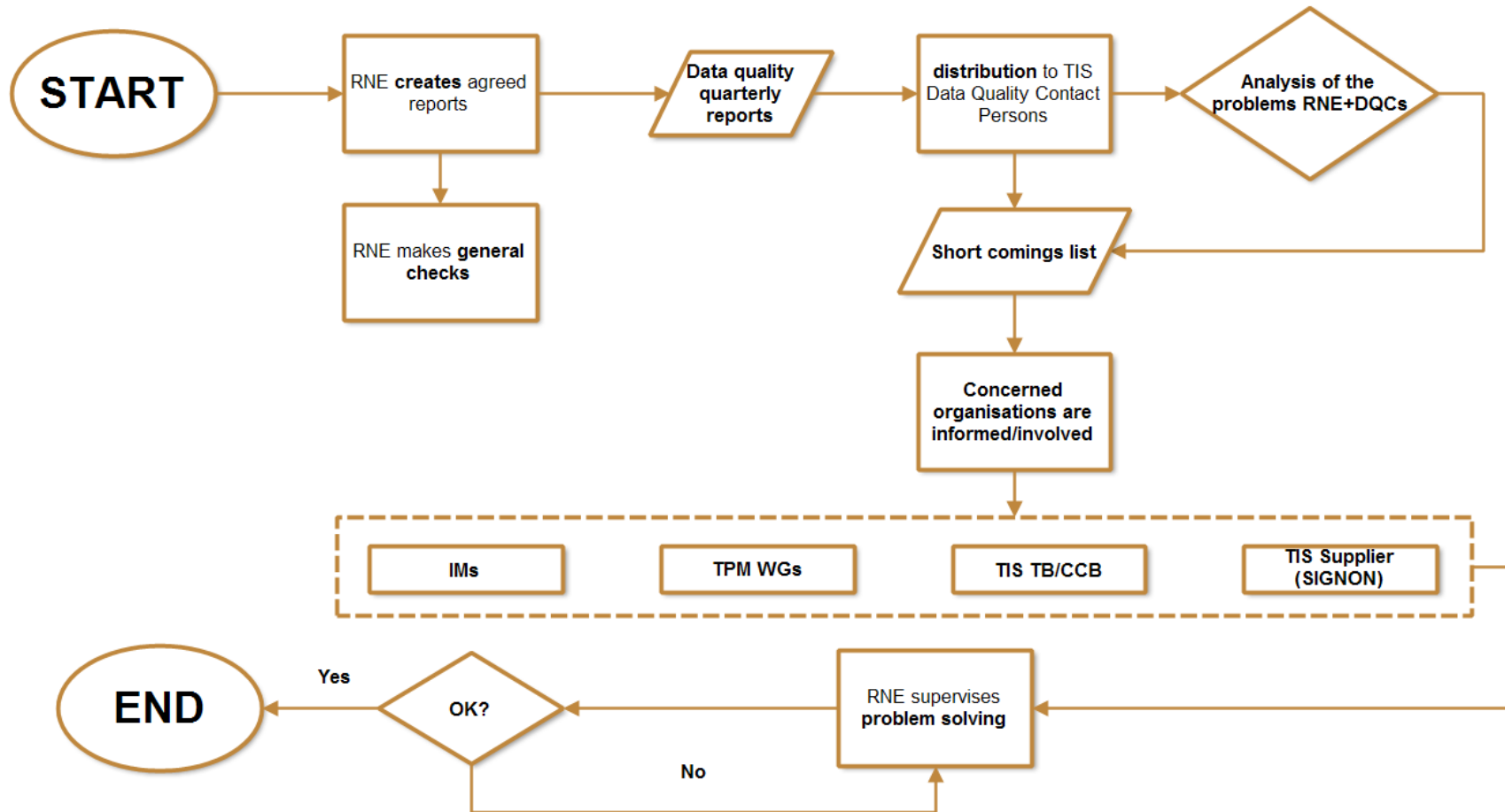


Figure 3 – quarterly data quality management

The principles of choosing the partner to contact are the same as in the monthly procedure (table 6).

The analysis of the problems is normally done by RNE. However, in the quarterly procedure, RNE might need the support of the DQCs to:

- » Check the status of the problems identified in the monthly procedure
- » Identify the cause of the problems where RNE is not able to do it alone
- » Discuss together the problems that occur more often (is there a way to avoid them in the future? Can we identify a way to solve them the works every time and apply it in the future also in the case of new IMs? etc..)

The DQCs have access to the OBI tool, where they can directly download many of the indicators displayed in the reports; these indicators are processed using “OBI analyses” that are contained in commonly accessible folders. The DQCs can also perform more detailed investigation applying additional filters, sorting and selecting, etc. However, they are not allowed to apply these modifications in the commonly used analyses, therefore they have to copy and paste them in the private folders in OBI and apply the needed modifications there.²

This analysis can be done:

- » In a common meeting or telephone conference
- » By bi- or multi-lateral contacts:
 - Telephone conference
 - Meetings
 - E-mail

The decision on which of the above will be commonly taken each time according to the situation. It is anyway advised to plan a meeting at least once a year³.

The outcome of the analysis shall be a list of problems including:

- » Number
- » Date of identification
- » Reported by
- » Addressed partner(s) (if it is a national problem, both the IM and the concerned department shall be indicated, when known)
- » Type of problem
- » Description
- » Consequence
- » Possible solution, if any
- » Ticket number (if any)
- » Estimation of the time for solving the problem and/or patch planned for the release (if any)
- » Cost (if any)
- » Status (when the problem is solved, indication of the date)

The tool to keep track of the above mentioned list will be, for the time being, an xls file called “Shortcoming list” stored in project place (...).

The maintenance of the file will be done by RNE but the DQCs will be access to it and will be responsible to keep it updated for the problems under their competence (i.e. national problems ≠ IT).

RNE will study a more efficient tool (web-based) for the future.

² For additional instructions, RNE IT staff is available

³ It does not necessarily have to be a specific meeting; TIS data quality can also be a point in the agenda of another group’s meeting. What is important is that the DQs participate in this meeting.

6 Update procedure

This document shall be updated at least once a year or more often, according to the needs, if this is asked by several partners and on the basis of urgent modifications of tools or in case of big data quality problems, that cannot be analysed using the agreed procedures and reports.

During the update, the following items **can** be revised:

- » List of indicators and related analyses
- » Formulas or other technical features
- » Procedures

During the yearly revision, the following items **must** be revised:

- » Data quality targets (when existing, see following section)
- » DQCs List

The DQCs list must be anyway kept updated; therefore any change shall be immediately communicated by the concerned IM to RNE.

The actual procedure to update the document depends on the circumstances: huge revisions might require a meeting, minor revision will only need e-mail communication.

RNE is responsible to lead this procedure.

The train list will be updated every 2 months according to the procedure described in Annex 4.

7 Short Glossary

- » Border point (or station): handover point or station coinciding (at least from the operational / administrative point of view) with the geographical border between two countries.
- » Contracted Timetable (CTT): Defines the planned route and planned time of a train run. It is delivered by the IMs to the TIS system with message 2090 and merged into an international timetable by the TIS-tool.
- » Data Defect Indicator (DDI): formerly “exclusion-rule”. Initially, the exclusion rules are functions of the EPR tool to avoid that train runs with incorrect or insufficient data or not agreed delay codes are considered in the EPR calculations. Now, these items are called “DDIs” because they are used as “labels” to indicate when a train run, processed in the EPR tool, contains a data quality problem (see also annex 4)
- » Delay: first or subsequent delay above a pre-defined national threshold; it is delivered by the IMs with message 2005 along with a delay cause code according to UIC leaflet 450-2.
- » Delay code: the “explanation” why a delay occurred in form of a code defined in UIC leaflet 450-2. Delay codes are subject to validation.
- » European Performance Regime (EPR): joint project carried out by RNE and UIC, dealing with international trains performance, concluded in 2012 (<http://www.rne.eu/epr.html>)
- » EPR tools: tools created within the EPR project to monitor data quality of EPR trains, support the validation of delay codes and carry out the calculations for the EPR model.
- » EPR Operation Working Group (EPR OWG): working group dealing with operational aspects within the EPR project
- » Lateness: difference between planned and actual time for a specific point/ status when >0.
- » Operations & After-Sale Working Group (OAS WG): working group dealing with operational and after-sales aspects within RNE organisation
- » Oracle Business Intelligence (OBI SI): tool normally used to elaborate data analysis and to create reports in RNE.
- » Point status: describes the status of a train run at a specific point. Relevant *stati* are:
 - 1 Arrival at final destination
 - 2 Departure from origin
 - 3 Arrival at an intermediate station

- 4 Departure from an intermediate station
 - 5 Run-through
- » Running Advice (RA): delivers the actual time at a specific point/ status and the deviation from the planned time at that point. It is sent by the IMs to the TIS system with message 2002.
 - » Rail Freight Corridor (RFC): a corridor set up and organised in accordance with EU Regulation 913/2010.
 - » Segment: part of a line between two consecutive points/status
 - » TAP/TAF TSI: (TSI means a specification adopted in accordance with EU Directive 2008/57 by which each subsystem or part subsystem is covered in order to meet the essential requirements and ensure the interoperability of the rail system; TAP and TAF are the Technical Specifications for Interoperability relating to Telematic Applications for Passenger and Freight traffic (more info: http://www.rne.eu/index.php/taf-tap_tsi_it.html)
 - » TIS Change Control Board (TIS CCB) : body, within RNE organisation, dealing with functional aspects and budget issues related to TIS. It is composed by one representative *per* IM member of TIS
 - » TIS Technical Board (TIS TB) : body, within RNE organisation, dealing with technical aspects related to TIS. It is composed by one representative *per* IM member of TIS
 - » Train Performance Management (TPM) : tasks carried out formerly within RNE Corridors organisation and now within the RFCs organisation, aimed to monitor and improve the performance of the international traffic (<http://www.rne.eu/train-performance-management.html>)
 - » Undocumented delays: delay minutes to which no cause has been attributed. It is a calculation function in the EPR calculation tool and calculated on segment level as follows:

$$\text{Lateness}_{\text{EPR Point/status 2}} \text{ minus Lateness}_{\text{EPR Point/status 1}} \text{ minus Delay}_{\text{Segment}}, \text{ if } >0$$

8 Annexes

1. Report template
2. Shortcoming list template
3. Train list for TPM analyses
4. Train list update procedure
5. Train List collection template

9 Appendix

1. List of DQCs
2. List and explanation of DDIs

9.1 Appendix 1 – List of Data Quality Contacts

IM	Name	Surname	e-mail address
BDK	Troels	Richter	TRRI@bane.dk
BLS AG	Alexander	Paulus ⁴	Alexander.Paulus@bls.ch
CFL			
DB Netz AG	Siegfried	Nierichlo	siegfried.nierichlo@bahn.de
GySEV- Raaberbahn	Csaba	Nemeth	csnemeth@gysev.hu
HŽ	Mate	Maric	mate.maric@hzinfra.hr
INFRABEL	Ann	Verstraelen	ann.verstraelen@INFRABEL.BE
KEYRAIL			
MÁV	László	Fenyves	fenyvesl@mav.hu
ÖBB	Georg	Meyer	georg.mayer@oebb.at
PKP PLK	Maarten	Guut	Maarten.Gutt@plk-sa.pl
ProRail B.V.	Peter	Vanwaveren	?
RFF	François	Binet	francois.binet@sncf.fr
RFI	Roberto	Caruso	r.caruso@rfi.it
SBB	Karl	Guntern	KARL.GUNTERN@sbb.ch
SŽ	Boris	Žebalc	boris.zebalc@slo-zeleznice.si
SŽDC	Richard	Svoboda	svobodar@szdc.cz
TRV	Per-Anders	Winge	per- anders.winge@trafikverket.se
ŽSR	Ludmilla	Malíková	malikova.ludmila@zsr.sk

⁴ TIS-Data provided by SBB, technical issues to be dealt with SBB

9.2 Appendix 2 – List of Data Defect Indicators and short explanation

Number	Name	Type	Description
0	Not excluded	-	No data quality problems
1	CTT not fitting at borders	Regular	No consecutive (from chronological point of view) points at border
2	Train partly cancelled	Regular	A failure message exists
3	Missing CTT	-	Not implemented
4	Partly excluded	-	Not implemented
5	RA inconsistent	Regular	Order of the RA is not consistent with the order of CTTs
6	Missing RA	Regular	Missing RA on a EPR point
7	Train with 95 at the end of the validation period	End of validation period	EPR train has a delay with code 95 ("further investigation needed")
8	Train with C at the end of the validation period	End of validation period	EPR train has a closed delay (validating partners cannot find an agreement)
9	Manually excluded	Manual	EPR train has been manually excluded from the user interface
10	Standby	Manual	EPR train has been manually marked for re-checking the exclusion rules from the user interface
11	Train with D at the end of the validation period	End of validation period	EPR train has a disputed delay (validating partners cannot find an agreement)
12	Train run not complete	Regular	Train does not match EPR train list
13	Error un calculation	Calculation	It is used in the calculation tool when unexpected calculation mistakes occur
14	Train with no responsible company	End of validation period	It was not possible to identify a company responsible for a delay
15	Train with VI_RU as responsible company	End of validation period	E PR train has delay with virtual RU as responsible Company (not existing anymore from TIS V5 onwards)