



Evaluation of the formal quality of data/service documentation

Report by Task 3.2: European NAPs data quality

Status: final
Version: 1.0
Date: 18 March 2025

Legal disclaimer

The sole responsibility for the content of this document lies with the authors. It does not necessarily reflect the opinion of the European Union. The European Commission is not responsible for any use that may be made of the information contained therein. All images are provided by the respective partners (unless otherwise noted) and are approved for reproduction in this publication.

Document information

Project acronym	NAPCORE
Full project title	National Access Point Coordination Organisation for Europe
Grant Agreement No.	MOVE/B4/SUB/2020-123/SI2.852232
Activity no. and title	WG 3 NAP content and accessibility
Author(s)	Petr Bureš (CZ/TTR)
Co-author(s)	Jan Vlčinský (CZ/TTR), Chrysostomos Mylonas (GR/CERT), Peter Lubrich (DE/BASt), Jasper Beernaerts (BE/NGI)
Related to Milestone no.	M3.8
External Milestone	yes

Document history

Version	Date	created/ modified by	Comments
0.2	06.04.2022	Petr Bureš (CZ/TTR)	The initial version from the project InQMS Machine translated
0.3	27.04.2022	Petr Bureš (CZ/TTR)	Initial Revision of the text for the project NAPCORE and comment resolution
0.4	12.05.2022	Petr Bureš (CZ/TTR)	Revision based on received comments
0.5	16.02.2024	Petr Bureš (CZ/TTR)	Further revisions and additions to the document.
0.6	24.6.2024	Petr Bureš (CZ/TTR)	Added dataset type list, introduction list, and reference to the CMC to the evaluation form and reflected this change in the methodology here. Few textual changes.
0.7	2.10.2024	Petr Bureš (CZ/TTR)	Conversion to the NAPCORE template.
0.8	21.10.2024	Petr Bureš, Chrysostomos Mylonas	Extensive addition of evaluation details after methodology trial of 33 datasets.
0.9	3.3.2025	Petr Bureš	Addition of the evaluation options to all parameters and clarification of parameters definitions and Resolution of the comments of the methodology and setup of issue tracker for next revision.
0.95	13.3.2025	Jasper Beernaerts, Rodolfo da Silva	Commenting of the methodology
1.0	18.3.2025	Petr Bureš, Peter Lubrich	Final revision based on received comments, incorporation of introductory text.

Action requested

- ☐ To be revised by partners involved in the preparation of the document
- ☐ For review/ approval by the Core Alignment Team
- ☐ For approval by the NAPCORE Steering Committee



Abstract

This deliverable by NAPCORE subWG 3.2 introduces the cross-domain quality (CDQ) assessment, which focuses on evaluating the quality of service/dataset documentation, based on metadata provided by the content providers to the National Access Points (NAPs). This includes any additional materials such as documentation, samples, and schemas. The method combines manual evaluation of quality of each assessed parameter into an overall grade (A-F) and is intended for use by competent authorities during the compliance assessments. This document is complemented by an MS Excel template where the evaluation results are recorded, and the overall quality score is calculated.



Table of contents

1	Introduction	6
2	Scope.....	8
3	Terms and abbreviations	9
3.1	Terms and definitions.....	9
3.2	Symbols and abbreviations.....	10
4	Methodology.....	12
4.1	Inputs.....	12
4.2	Outputs.....	12
4.3	Filling in a parameter value	13
4.4	Evaluation of parameter (grading).....	15
5	Evaluated Parameters	17
5.1	Evaluation metadata	18
5.2	Metadata information.....	18
5.3	Content information	20
5.4	Temporal information	25
5.5	Geographical coverage.....	25
5.6	Transport modes information	27
5.7	Responsibilities	27
5.8	Conditions for use	29
5.9	Access information.....	29
5.10	Quality information.....	35
6	Overall evaluation	39
6.1	Overall commentary.....	39
6.2	Overall grade computation.....	39
6.3	Method of presentation	41
7	Future work and conclusions.....	43
8	References.....	44

List of Figures

Figure 1	Cross domain quality framework and its place among the possible NAPCORE Quality framework structure.....	6
Figure 2	Specific evaluations sorted by data/service subscription phases (based on InQMS)	7
Figure 3	Screenshot of the evaluation sheet of the Evaluation Form	13



Figure 4 Filling in the parameter value from the multiselect dropdown menu	13
Figure 5 Fill in the parameter “category” explicitly stated in the collected evidence in the evaluation form.	14
Figure 6 Using the “lists” sheet in the evaluation form to see correct parameter values, filled in via multiselect.....	15
Figure 7 Definition of the parameter for evaluation	17
Figure 8 The “lists” sheet of the evaluation form with predefined parameter values.....	17
Figure 10 Final overall grade and commentary.....	39
Figure 11 Presentation of the resulting evaluation in a FORM, (clipping)	42

List of Tables

Table 1 Generalization of parameter evaluation (more detail provided in each parameter).....	16
Table 2 Evaluation metadata.....	18
Table 3 Metadata information	18
Table 4 Content information	20
Table 5 Temporal information.....	25
Table 6 Geographical coverage	25
Table 7 Transportation system.....	27
Table 8 Responsibilities.....	27
Table 9 Conditions for use.....	29
Table 10 Access information	29
Table 11 Quality information	35
Table 12 Mapping from the mark A-F to a numeric value	39
Table 13 Mapping of mandatory appearance to a value	39
Table 14 Mapping of a weight (importance) to a value.....	40
Table 15 Classification of parameters.....	40
Table 16 Mapping scoring 1-4 to the grade A-F	41



1 Introduction

The Cross-Domain Quality (CDQ) assessment concept encompasses a universal set of evaluation methods that can be applied to any data and services accessible via National Access Points (NAPs), regardless of the domain. The CDQ concept allows for the assessment of common quality features related to data processing and data acquisition, as experienced from the data user's perspective. It defines various evaluation methods to be applied at different stages of data or service subscription.

Overall, the set of assessment methods from the CDQ methodology creates a horizontal layer above the other domain-related Quality Frameworks (QFs), as illustrated in the next figure. This methodology defines agreed-upon features of data and service provision that can be observed and evaluated towards a common goal: providing services uniformly according to agreed rules and best practices. Following this methodology will also highlight which features need enhancement in a NAP system update.

Horizontal formal and technical quality framework														
Horizontal common features framework														
STTP		SRTI		RTTI		MMTIS					emerging services			
Data category quality package	Data category quality package	Data category quality package	Data category quality package	Data category quality package	Data category quality package	Data category quality package	Data category quality package	Data category quality package	Data category quality package	Data category quality package	Data category quality package	Data category quality package	Data category quality package	Data category quality package

Figure 1 Cross domain quality framework and its place among the possible NAPCORE Quality framework structure

The presented CDQ concept builds on a previous approach first introduced in the InQMS¹ project. Which recognizes 10 different evaluations, each documented by a separate report and, if possible, evaluation forms. From the data/service subscription setup perspective, the tests can be sorted according to their different phase into:

- Manual evaluation of the dataset/service documentation before its consumption (planning phase), e.g. based only on provided information not on real data, consists of:
 - Evaluation of the **formal** quality of web service documentation
 - Evaluation of the **objective** quality of web service documentation
- Manual evaluation of the distribution method and data structure at the time of the dataset/service consumption staging, e.g. based on hands on experience with how to get the data and how they initially look like compared to the expectation, consist of:
 - Manual evaluation of the distribution quality of a web service
 - Manual evaluation of the structural quality of a web service

¹ <https://inqms.tamtamresearch.com/>

- Automatic evaluation of distribution, structure, and content of the data/service during the time of its active consumption (production phase), e.g. how the data and distribution service behave over time, consist of:
 - Automatic evaluation of the distribution quality of a web service
 - Automatic evaluation of the structural quality of a web service
 - Automatic evaluation of web service content quality
- Ad hoc manual evaluation of selected aspects of the data/service content that is being actively consumed. In case of detected problems with web service, manual data evaluation can be performed focusing on detailed aspects that are not normally monitored.
 - Manual evaluation of web service content quality
- Finally, it is also possible to manually evaluate the service according to its overall quality:
 - Evaluation of the overall quality of the event-driven service
 - Evaluation of the overall quality of the status-oriented service

The following figure summarises the 10 different evaluations, separated by the above phases “planning”, “staging”, “production” and “ad hoc”.

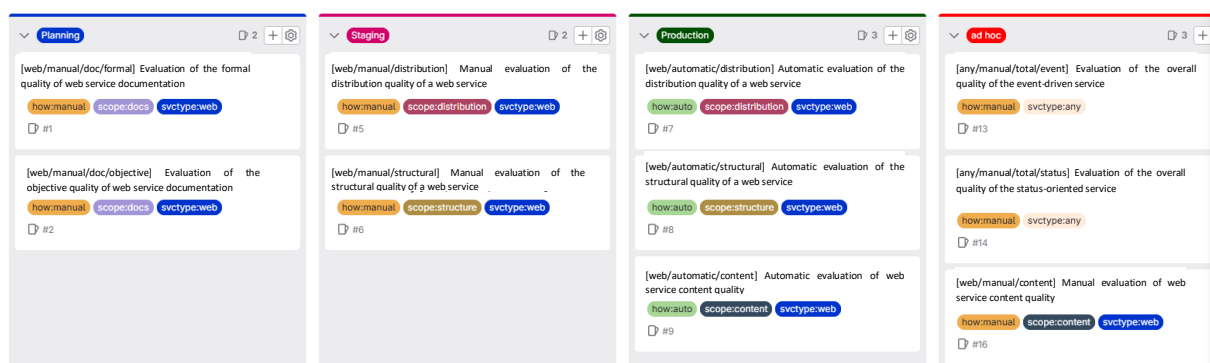


Figure 2 Specific evaluations sorted by data/service subscription phases (based on InQMS)

These tests are intended to be used by data consumer to fully cover his/her needs when evaluating quality of the consumed dataset/service and by the competent authority when performing compliance check of the dataset/service with the requirements of the delegated regulation and by the data holder when preparing for the compliance assessment.

The first two evaluations focus on the quality of the description of the dataset/service at a NAP. The intention is that better and more complete description of datasets and services at a NAP allows content consumer to easily navigate the many datasets that are available on an NAP, assess their usability, and prepare technically and organizationally for its consumption. It is therefore useful to evaluate the quality of the description provided for such a dataset or service registered on the NAP.

2 Scope

This document defines the manual evaluation of quality of metadata descriptions of datasets or services, aka catalogue records, registered on a National Access Point. This method evaluates the availability, completeness, and usability of metadata elements, further on referred to as parameters, such as:

- Identification and a brief description of the content (data/service)
- Identification of the content provider and owner
- Description of the main concepts, structure, and content
- Samples and validation schemas
- Technical and organizational details, licensing, and subscription description

The parameters assessed in this evaluation are derived from metadata defined by the Coordinated Metadata Catalogue (CMC) [1], findings of the InQMS² project [2], the NAPCORE project [3], and quality recommendations by W3C and open data portals [6]. The automatic methods of metadata structural components evaluation such as MQA³ methodology [5] have not been, due to its nature, used for the current in-depth manual metadata evaluation.

The parameters are classified into eight categories. Each category defines specific parameters to be assessed, along with guidance on how to grade them (A-F). The parameter definition is complemented by references to the Coordinated Metadata Catalogue (CMC) and mobilityDCAT-AP, using the available mapping⁴.

This document also defines the importance/weight of the assessed parameters, their combination into an overall grade, and the documentation required during the assessment.

² <https://inqms.tamtamresearch.com/>

³ <https://data.europa.eu/mqa/?locale=en>.

⁴ <https://mobilitydcat-ap.github.io/mobilityDCAT-AP/releases/index.html#Mapping-to-Coordinated-Metadata-Catalogue>



3 Terms and abbreviations

3.1 Terms and definitions

For the purposes of this document, the terms and definitions given in “Quality of Service Assessment Reports – introduction” and in the following list shall apply:

3.1.1

data portal

A web-based system that contains a catalogue with descriptions of datasets and their distributions.

3.1.2

national access point

A data portal facilitating access to transport-related data according to the EC ITS Directive.

3.1.3

(data) catalogue

A collection of catalogue records on a data portal.

3.1.4

catalogue record

A description of a dataset and its distributions published on the portal.

3.1.5

dataset

A collection of data, published by a single publisher; sharing some common metadata (e.g., theme, rights holder, space); and provided via one or more distributions.

3.1.6

(data) distribution

A dataset served using a single format and protocol.

3.1.7

rights holder; data holder; data owner

An entity holding rights to the dataset.

3.1.8

(data) publisher

An entity responsible for publishing the dataset distributions (making the dataset distributions technically accessible).

3.1.9

content model

A class model of possible data distribution content describing its logical structure, while abstracting specific serialization format.



3.1.10**(content model) profile**

A simplified class model created from the complete content model by removing structures that are not needed and possibly enforcing the presence of some attributes and structures.

3.1.11**schema**

A description of structural constraints for data distribution content using a specific format (e.g., XML, JSON, etc.).

3.1.12**format**

A specification that allows for the interpretation of serialized dataset distribution content. It should include a schema and may include samples and other accompanying documentation.

3.1.13**sample**

An instance of distribution content.

3.1.14**protocol**

A specification of rules and procedures allowing a consumer to get a sample of distribution content.

3.1.15**parameter**

A metadata entry, part of the catalogue record that is being individually assessed.

3.1.16**service**

A collection of operations that provides access to one or more datasets or data processing functions.

3.1.17**metadata**

Data that provides information about other data. It helps describe the content, context, and structure of data, making it easier to find, use, and manage.

3.2 Symbols and abbreviations

CMC	Coordinated Metadata Catalogue
CDQ	Cross-Domain Quality
InQMS	Integrated Quality Management System
NAP	National Access Point
NAPCORE	National Access Point Coordination Organisation for Europe
DCAT-AP	Data Catalogue Vocabulary - Application Profile for Data Portals in Europe
MQA	Metadata Quality Assurance



Evaluation of the formal quality of data/service documentation

GUID	Globally Unique Identifier
URI	Uniform Resource Identifier
URL	Uniform Resource Locator
GML	Geographical Markup Language



4 Methodology

This chapter describes the process of the manual evaluation of the metadata quality, including information about the collected evidence, the inputs, and the evaluation (grading) of individual parameters and method of presentation of evaluation results.

4.1 Inputs

The inputs, required for the evaluation include all documents that the evaluator can retrieve about the dataset or service under test **via the NAP** (documents, information is either part of the metadata at the NAP or are linked in metadata to other data portals). These later can serve as a proof.

Example of input:

The evaluator receives a link to a catalogue record at NAP (e.g., https://registr.dopravniinfo.cz/en/sources/cz-ndic_d2-common/), retrieves and stores all related pages and linked documents for manual inspection.

4.1.1 Collected evidence

The collected evidence includes (if available):

- Textual (Word, PDF) materials describing the data/service, its format, and its protocol
- Textual (Word, PDF) materials describing how to obtain the data, including licensing terms
- Printouts of the webpages with the relevant information about the dataset / service
- Technical materials (RDF XML, json-ml ...) containing the catalogue record
- Technical materials in the form of samples, schemas, and descriptions of WSDL services
- Other relevant materials

4.2 Outputs

The evaluation is recorded into the “Evaluation form” (Report).

4.2.1 Report (Evaluation Form)

The evaluator fills in the “data and evaluation” sheet for the Evaluation Form (MS Excel) with his/her findings, the parameter values, the assessment and comment. Instructions on how to fill the parameter value and how to grade it is in following chapters.

Evaluation of formal documentation quality							
[FRM_web-manual-doc-formal]				version 2025030			
overall quality				B	Dataset described in great detail with schemas samples and separate documentation available both in English and local language. Quality parameters are mostly missing. The consolidation of the documentation of the dataset could be beneficial.		
category	parameter	ref.	status	grade	value	indire	comment
	evaluated catalogue record URL			-	https://registr.dopravniinfo.cz/en/sources/cz-ndic_d2-common/		Reference evaluation for the purpose of the presenting a typical evaluation according to the methodology. Meant to be used by other evaluators as a guidance.
	evaluation date			-	02.03.2025		
	evaluator contact			-	Petr Bureš		
Metadata information	Metadata date	2.2.1.1	mandatory	C	01.03.2024	yes	at the bottom of the page, same for all records. Unclear when this specific record had changed
	Metadata language	2.2.1.2	mandatory	C	cs: Czech, en: English	yes	inferred from the UI
	Contact point	2.2.1.3	partial	B	name: Ctirad Weissmann title: Manager of traffic information distribution department phone: +420-596663578 email: ctirad.weissmann@rsd.cz		data derived from the web, provider information section, with all sufficient details; it is however unclear if this contact is also metadata provider (the person who put metadata together).
	Dataset identifier		recommended	A	x-source:cz-ndic_d2-common		dataset uri
	Name of the dataset	2.2.2.1	mandatory	A	DATEX II - Common Traffic Information		within a dataset specification [cs, en]
< > introduction data_and_evaluation support lists +							

Figure 3 Screenshot of the evaluation sheet of the Evaluation Form

The “data and evaluation” sheet contains:

- Overall grade and comment
- Information about the evaluation itself (category: evaluation metadata)
- A section for each evaluated metadata category, with individual evaluated parameters:
 - **Parameter:** Contains individual parameters to be evaluated
 - **Reference:** A reference to the section of the CMC 2019 [1]
 - **Status:** Information on whether the parameter is mandatory, optional etc.
 - **Grade:** A, B, C, F, and n/a Grading (F = fail, n/a = not applicable)
 - **Value:** Information (parameter value) from the collected evidence
 - **Indirect:** Indicates whether the parameter value is explicitly stated in the expected format in the collected evidence or derived/extrapolated from the evidence
 - **Comment:** Evaluator’s comment (complementing the grade)

4.3 Filling in a parameter value

During the evaluation process the evaluator fills in into Evaluation Form the observed or expected value of the evaluated parameter.

Free text values are written (copy pasted) by evaluator while predefined list values are selected from the multiselect dropdown menu. The cells restricted to predefined lists of values are highlighted in light blue background colour, see next figure.

Metadata information	Metadata date	2.2.1.1	mandatory	C	01.03.2024	yes
	Metadata language	2.2.1.2	mandatory	B		yes
	Contact point	2.2.1.3	partially*	B	Clear All sq: Albanian eu: Basque be: Belarusian bg: Bulgarian hr: Croatian cs: Czech	

Figure 4 Filling in the parameter value from the multiselect dropdown menu

The parameter values could be in the collected evidence, stated either directly or indirectly, for explanation and the related process description see next subchapters.



4.3.1 Directly stated information

If the information is:

- explicitly stated in the collected evidence (see Figure 3) **AND**
- in the expected format,

the evaluator fills in the **observed** parameter value (into the “value” column).

Example: The expected parameter “category” with predefined values (Road work; Unexpected road events and conditions; Real-time traffic data) is in the evaluated catalogue record.

NDIC

Sources

Overview

NeTEx - Timetable Information

DATEX II - Common Traffic Information (snapshot)

DATEX II - Common Traffic Information

Source: DATEX II - Common Traffic Information

General Parameters

Content

Coverage

Access, used formats and protocols

Quality information

DATEX II - FCD data v2

DATEX II - FCD data

DATEX II - Truck Rest Areas (predefined locations)

DATEX II - FCD data (predefined locations) v1.1

DATEX II - Traffic Status (predefined locations) v1.1

DATEX II - Weather (predefined locations) v1.1

DATEX II - Restrictions

DATEX II - Traffic Status

DATEX II - Variable Message Signs

General Parameters

Parameter	Value
uri	x-source:cz-ndic_d2-common
provider	NDIC
status	production ¹ , publication start: 2018-01-01
licence ²	Contract and Free of charge
subscription process	Free of Charge Traffic Data Subscription
model	DATEX II v2.3/SituationPublication
used pattern ³	Common Traffic Information extended profile (Czech only)

Content

Parameter ²	Value
resource type	data set
category	Road work Unexpected road events and conditions Real-time traffic data
detailed type	Long-term road works Short-term road works Accidents and incidents Poor road conditions Weather conditions affecting road surface and visibility Location and length of queues
data set language	cs

Related location data sources

Parameter	Value
-----------	-------

Evaluation of formal documentation quality

category	parameter	ref.	status	grade	value	indirect	comment
	Description of dataset	2.2.2.2	mandatory	A	Common traffic information (accidents etc.), provided by NDIC for the whole Czechia, pushed on occurrence, in a common traffic information profile of DATEX II model version 2.3.		within a dataset specification [cs, en]
	Resource type	2.2.2.3	mandatory	A	DataSet		Directly in metadata as literals (listed in a table)
	Dataset type category	2.2.2.4	mandatory	A	Road work information, Unexpected road events and conditions, Real-time traffic data		Directly in metadata as literals (listed in a table)
Content information	Dataset detailed type	2.2.2.5	optional	A	Long-term road works , Short-term road works, Accidents and incidents , Poor road conditions, Weather conditions affecting road surface and visibility, Location and length of queues		Directly in metadata as literals (listed in a table)

Figure 5 Fill in the parameter “category” explicitly stated in the collected evidence in the evaluation form.

4.3.2 Indirectly stated metadata

If the information is:

- **NOT** explicitly stated in the collected evidence **AND/OR**



- **NOT** having expected values of the parameter according to predefined list or expected format (expected values are stated in the list supporting data (see Figure 4) **BUT**
- **COULD** be unanimously inferred from either metadata on the website or provided documentation and samples,

the evaluator fills in the **expected value of the parameter** (into the “value” column), checks the column “**indirect**” with “yes” and fills in the note how was the value retrieved/extrapolated.

The grade is based on the ease of the extrapolation of the parameter, see the next chapter.

Used enumerations								
1								
35	pl: Polish		Closures and dynamic access conditions of lanes	Access interface – application layer protocol				Air
36	pt: Portuguese		Closures and dynamic access conditions of bridges	Clear All				rail (including high speed rail)
37	ro: Romanian		Closures and dynamic access conditions of tunnels	SOAP				conventional rail
38	ru: Russian		Other access restrictions and traffic regulations	OTS2				light rail
39	gd: Scottish Gaelic		Dynamic overtaking bans on heavy goods vehicles	HTTP/HTTPS				long-distance coach
40	sr: Serbian		Dynamic speed limits	FTP				maritime (including ferry)
41	sk: Slovak		Direction of travel on reversible lanes	RSS				metro
42	sl: Slovenian			AMQP				tram
43	es: Spanish	Road work information	Long-term road works	MQTT			Schedule	bus
44	sv: Swedish	Unexpected road events and	Short-term road works	gRPC			d	trolley-bus
45	tr: Turkish		Accidents and incidents	other				shuttle bus
46	uk: Ukrainian		Poor road conditions					shuttle ferry
47	cy: Welsh	Road weather conditions	Weather conditions affecting road surface and visibility	Update frequency				taxi
48		Real-time traffic data	Traffic volume	Clear All				car-sharing
49	Security mechanisms		Speed	On occurrence				car-pooling
50	Clear All		Location and length of queues	Up to 1min			Demand-	car-hire
51	none		Current travel times	Up to 5min			responsiv	bike-sharing
52	IP filter		Predicted travel times	Up to 10 min			e	bike-hire
53	basic		Expected delays	Up to 15 min				car
54	digest		Waiting time at border crossings to non-EU Member States	Up to 30 min				truck
55	certificate		Traffic data at border crossings to third countries	Up to 1h				motorcycle
56	URL parameters			Up to 2h				cycle
57	other	Information for trip-	Topographic places	Up to 3h			Personal	pedestrian

Figure 6 Using the “lists” sheet in the evaluation form to see correct parameter values, filled in via multiselect

4.4 Evaluation of parameter (grading)

After entering the parameter value, in the “data and evaluation” sheet of the Evaluation Form, the evaluator assesses the ease of finding the parameter by

- Providing a **mandatory grade** (A, B, C, F, or not relevant (n/a)⁵) **AND**
- Adding a **mandatory comment** on how and where the parameter was found, including any useful information, such as issues with completeness or relevance.

The evaluation grade focuses on the correctness of the metadata. It checks if the parameter is present and if its value is within the expected range. In some cases, the grade reflects also the quality of how the parameter is provided, as a simple “THERE” or “NOT THERE” may not capture the evaluator’s experience. The expected values and formats for each parameter are specified in this document. Each parameter is assessed separately.

4.4.1 When n/a can be applied

Grading a parameter as not relevant (by means of applying the value n/a) is allowed only in special cases. Mostly in conditional cases, when parameter is exclusive to another parameter. Most of optional or recommended parameters must be graded as F if missing (the option n/a is not available in dropdown menu for the parameter evaluation). This allows catalogue records with optional and recommended parameters filled in to be rated slightly better than those without.

⁵ Only possible for some evaluated parameters, this must be explicitly allowed.

4.4.2 Where to look for the parameter

The technical formats, in which metadata are internally stored (and distributed), could be visualized differently on the NAP. For example, combining a few parameters together to enhance user experience.

Therefore, during the visual evaluation, the focus is not on whether the parameters have same name or follow the exact expected structure. Instead, the emphasis is on the values of the parameters, ensuring they meet expectations and are easily identifiable among other parameters.

4.4.3 General grading information

Due to individual nature of the parameters the grading slightly varies from parameter to parameter, e.g. sometimes some grades are not applicable, therefore we provide grading information for each parameter separately in chapter 5 Evaluated Parameters. The Table 1 provides a **generalization** of the grading principle.

Table 1 Generalization of parameter evaluation (more detail provided in each parameter)

grade	Value	Description
A	1	Metadata is provided in the expected format with prescribed values as defined by the relevant standard (i.e., CMC). = Parameter value(s) in the machine-readable metadata exactly as expected.
B	2	Metadata is provided as part of easily identifiable substructures. It does not have to be directly in the expected format (e.g., contact data in different forms, different wordings of the keywords), but it is semantically compatible. = Parameter value(s) in the machine-readable metadata, in a different form than expected but semantically compatible.
C	3	Metadata is not provided directly but is inferable from the collected evidence. It does not “fit” in terms of format and structure and/or not provided in full. = Parameter value(s) are not part of the machine-readable metadata, but findable elsewhere and semantically compatible.
F	4	Metadata is not provided and cannot be “easily” derived from the collected evidence (including samples and schemas provided). = Parameter value(s) are not found/filled in. Assumptions and guessing would be needed to fill in the value, and correctness is not assured.
n/a	0	(Not applicable) The item is not evaluated because it makes no sense in the context. The possibility of not evaluation parameter is indicated per parameter. It CAN NOT be assumed by the evaluator for any given parameter, e.g. if evaluator does not know.

5 Evaluated Parameters

This chapter contains parameters to be assessed within this methodology. Each parameter is described with following information (see Figure 7):

1. Parameter name
2. Sample(s) of parameter value
3. Short definition of parameter
4. Expected format (e.g. *dateTime*, *enumeration*) and full list of expected values, to differentiate the list of values in in *different font*.
5. Parameter status (*mandatory*, *optional*, *recommended*, *partial*) and if the parameter could contain more than one value (multiple choice), for example more values of the list.
6. Reference to parameter definition in Coordinated Metadata Catalogue or mobilityDCAT-AP
7. Evaluation guideline with all possible grades with detailed explanation how to grade based on the observation and
8. notes.

Resource type	Dataset	3 Classifies the resource of the publication, depending on if it is a “dataset” or a “service”. Dataset is a downloadable resource; service is a map or API. 4 Format: <code>enumeration (dataset; service);</code> Mandatory. 6 reference: CMC 2.2.2.3; MD <i>dcat:Dataset/dcat:DataService/</i> 7 Grading • A = specific field in expected format / expected structure. • B = specific field, not filled in exactly using list item, but identifiable or misclassified by data holder (dataset as data service) • C = inferred from the name, description, documentation, sample or download URL. • F = not provided or found via previous options and it is not clear if the record describes dataset or service 8 NOTE: There is often misclassification by the data holder.
---------------	---------	--

Figure 7 Definition of the parameter for evaluation

Most parameters and values are derived from the Coordinated Metadata Catalogue (including Appendices 1 and 2) [1]. All possible values for each parameter are listed in the evaluation form’s “lists” sheet (see Figure 8). Parameters highlighted in orange are based on other inputs.

A	B	C	D	E	F	G	H	I
Used enumerations								
1								
35	pl: Polish	and regulations	Closures and dynamic access conditions of lanes	Access Interface – application layer protocol				Air
36	pt: Portuguese		Closures and dynamic access conditions of bridges	Clear All				rail (including high speed rail)
37	ro: Romanian		Closures and dynamic access conditions of tunnels	SOAP				conventional rail
38	ru: Russian		Other access restrictions and traffic regulations	OTS2				light rail
39	gd: Scottish Gaelic		Dynamic overtaking bans on heavy goods vehicles	HTTP/HTTPS				long-distance coach
40	sr: Serbian		Dynamic speed limits	FTP				maritime (including ferry)
41	sk: Slovak		Direction of travel on reversible lanes	RSS				metro
42	st: Slovenian		Other temporary traffic management measures or plans	AMQP				tram
43	es: Spanish	Road work	Long-term road works	MQTT				bus
44	sv: Swedish	information	Short-term road works	gRPC				trolley-bus
45	tr: Turkish	Unexpected road	Accidents and incidents	other				shuttle bus
46	uk: Ukrainian	events and	Poor road conditions					shuttle ferry
47	cy: Welsh	Road weather	Weather conditions affecting road surface and visibility	Update frequency				taxi
48		conditions		Clear All				car-sharing
49		Real-time traffic	Traffic volume					

Figure 8 The “lists” sheet of the evaluation form with predefined parameter values

Parameters are grouped into categories, presented below in individual subchapters.



5.1 Evaluation metadata

Information about the evaluation itself. This part is not evaluated.

Table 2 Evaluation metadata

Parameter	Example of a value	Definition, comments
evaluated catalogue record URL	https://data.gov.eu/dataset1	An URL to the catalogue record of the data under the test at the NAP. It is the link to metadata at the NAP, (not the data access URL).
evaluation date	18.5.2024	Date when the evaluation was completed. (ideally in standardised format)
evaluator contact	John Evaluator, john@evaluate.it	Name, surname, and email contact to the evaluator.
comment	This is a test	Any information about the evaluation (not about its result) that the evaluator wants to share with the content provider. If there is no direct URL of the dataset, identify the correct dataset to be evaluated. Identify the purpose of the evaluation.

5.2 Metadata information

Information about the metadata includes its age, the provider, and the language(s) in which it is available.

Table 3 Metadata information

Parameter	Example of a value	Definition, evaluation, comments
Metadata date	2023-10-23T09:00:00+01:00	The time when the metadata was created. Format: DateTime; Mandatory reference: CMC 2.2.1.1; MD <i>dcat:CatalogRecord</i>-> <i>dct:created</i> Grading • A = specific field in expected format. • B = specific field, other than expected format (e.g. 21.3.2024). • C = inferred from the webpage (change of the page), when only the evaluated dataset is present on the single webpage (so it is not mixed); • F = if not provided or part of documentation only (date in documentation is not specific since it usually suggests when documentation was updated not when the catalogue record was created on the NAP)
Metadata language	cs; en	Language in which the metadata is described. Format: 2 code of metadata language; Mandatory; multiple choice reference: CMC 2.2.1.2; MD <i>dcat:CatalogRecord</i>-> <i>dct:language</i> Grading • A = specific field with list of language codes (2 letter language code accepted, e.g. cs not cz or cze); • B = specific field with "human" language descriptions or other nonpreferred language codes (czech or cz or cze instead of cs); • C = inferred from user interface or textual content of the metadata. • F = not provided or found via previous options (in case of language unambiguity)



Table 3 Metadata information

Parameter	Example of a value	Definition, evaluation, comments
		NOTE: Could be multiple languages, either part of the metadata content or implied by having a language switch present and working on the webpage). The actual language used vs metadata indication is not evaluated, only noted in comments. In the comment provide info about other translations (language code, existence of the translation).
Contact point	Jane Metadatist, jane.metadatist@metacreator.com,	The organization, or person, who is responsible for the creation and maintenance of the metadata. Format: vCard-Textfields with the expected level of details (name, email, telephone, organization, position, address); Partial reference: CMC 2.2.1.3; MD <i>dcat:CatalogRecord</i>-> <i>dct:publisher</i> OR <i>dcat:Dataset</i> -> <i>dcat:contactPoint</i> Grading • A = dedicated contact information field, full information is provided (or at least name, position, email, telephone), responsibility clear as metadata publisher • B = a contact information field, full information is provided (or at least name, position, email, telephone), responsibility between data-owner and data-publisher and metadata publisher unclear, • C = inferred from documentation, responsibility is not clear as metadata contact, partial information is provided (only name and/or generic email). • F = not there, provision of a website URL or just a contact form without any contact information. NOTE: Could be part of the organization data or directly in the dataset description, could be contact info not in vCard format. Contact point, for the metadata could be data publisher, data owner or any other entity. If just one contact is filled in, the responsibility is unclear.
Dataset identifier	x-source:en-ndic_d2-common; D834B1C7; 249ddb42-3b9a-44fc-b965-8e7dc8325fb2	The identifier different from the URL that could be used to refer to the dataset. Format: free text; Recommended Reference: CMC NONE; MD <i>dcat:dataSet</i>-> <i>dct:identifier</i> (not present for service) Grading • A = specific field, some form of id that is unique out of the NAP context. • B = not an option here. • C = inferred from “access URL” that points directly to the dataset, not the URL of the catalogue record • F = all other situations, (access links to the portal) NOTE: Could be a URI or GUID, usually it is part of the URL specific to the data source. Dataset URI allows to link and identify dataset on the distribution interface, especially in case of user specific generated access URLs.

5.3 Content information

The dataset information includes its name, short description, type, category, detailed category, language(s), location referencing, documentation, and referenced datasets.

Table 4 Content information

Parameter	Example of a value	Definition, comments
Name of the dataset	DATEX II Situation Publication - Common Traffic Information; Rest areas static information; etc	Very short dataset description. Format: Free text; Mandatory. Reference: CMC 2.2.2.1; MD (<i>dcat:Dataset/dcat:DataService</i>) -> <i>dct:title</i> Grading - A = specific field, the heading, the web page title, in English and optionally other languages. B = not used - C = specific field, the heading, the web page title, BUT provided only in local language (non-English) - F = if there is just an URL instead on name, if it is unclear what the name is. NOTE: If available in more languages, give only English, but in the comment provide info about other translations (language code, translation ok?).
description of dataset	Common traffic information (accidents etc.), provided by NDIC for the whole Czechia, pushed on occurrence, in a common traffic information profile of DATEX II format.	Gives the user more information about the content of the dataset or service, a brief description just below the name of the dataset. Format: Free text. Mandatory Reference: CMC 2.2.2.2; MD (<i>dcat:Dataset/dcat:DataService</i>) -> <i>dct:description</i> Grading: - A = specific field, distinct part of the web page, in English and optionally other languages and extensive enough. B = not used - C = specific field, distinct part of the web page, BUT provided only in local language (non-English) - F = not present OR copy of the dataset name NOTE: If available in more languages, give only English, but in the comment provide info about other translations (language code, translation ok?). Provide in comment how extensive and descriptive the description is.
Resource type	Dataset	Classifies the resource of the publication, depending on if it is a “dataset” or a “service”. Dataset is a downloadable resource; service is a map or API. Format: enumeration (dataset; service); Mandatory. reference: CMC 2.2.2.3; MD <i>dcat:Dataset/dcat:DataService</i> Grading A = specific field in expected format / expected structure.

Table 4 Content information

Parameter	Example of a value	Definition, comments
		- B = specific field, not filled in exactly using list item, but identifiable or misclassified by data holder (dataset as data service) - C = inferred from the name, description, documentation, sample or download URL. - F = not provided or found via previous options and it is not clear if the record describes dataset or service NOTE: There is often misclassification by the data holder.
Dataset type category	Road work, Unexpected road events and conditions, Real-time traffic data	A description of a dataset type as a category is important for data seekers who are interested in a particular type of data. As per CMC <u>only one category</u> per dataset shall be selected. Format: enumeration (Static road network data; Static traffic signs and regulations; Toll information; Parking and rest area information; Filling and charging stations; Freight and logistics; Dynamic traffic signs and regulations; Road work; Unexpected road events and conditions; Real-time traffic data; General information for trip-planning; Public transport: location information; Public transport: operational information; Public transport: fare and purchase information; Cycle network data; Pedestrian network data; Demand-responsive modes); Mandatory; Multiple choice. reference: CMC 2.2.2.4; MD <i>dcat:Dataset</i> -> <i>mobilitydcatap:mobilityTheme</i> Grading - A = specific field, filled in with items from the list. - B = specific field but the items are not exactly same, but they are identifiable OR more than one type of category is used. - C = inferred from name, description or the documentation. - F = if it is not clear from anything above what dataset type categories are. - n/a = If the resource type = service, see note. NOTE: the service and dataset are mutually exclusive. In CMC the service does not have these categories. This is different in mobilityDCAT-AP, where service could be described by these categories. At least one of the parameters dataset or service detailed category must be present. Also, in the grading we ignore CMC requirement for just one category per record, allowing for dataset describe more categories.
Dataset detailed type	Long-term road works, Short-term road works, Accidents and incidents, Poor road conditions, Weather conditions	Describes the classification of the dataset content on a detailed level. It is used to concretize the element "Dataset type category" Format: enumeration (82 enumerations sorted into categories above); Mandatory; Multiple choice reference: CMC 2.2.2.5; MD <i>dcat:Dataset</i> -> <i>mobilitydcatap:mobilityTheme</i>

Table 4 Content information

Parameter	Example of a value	Definition, comments
	affecting road surface and visibility, Location and length of queues	Grading • A = specific field, filled in with items from the list. • B = specific field but the items are not exactly same, but they are identifiable, or some listed detailed types does not belong to the type category/ies listed above. • C = inferred from name, description or the documentation. • F = if it is not clear from anything above what dataset type it is. • n/a = If the resource type = service see note in the previous parameter.
Service type category	Location search, Information service, Trip plans	If the resource type is "service". Describes the classification of a service, following possible services listed in EU Delegated Regulation 2017/1926 Format: enumeration (Location search; Information service; Trip plans; Trip plans, auxiliary information, availability check; Trip plan computation - scheduled modes transport; Dynamic Passing times, trip plans and auxiliary information; Dynamic Information service; Dynamic availability check); Mandatory; multiple choice reference: CMC 2.2.2.6; MD <i>dcat:Dataset</i> -> <i>mobilitydcatap:intendedInformationService</i> Grading • A = specific field, filled in with items from the list. • B = specific field, but the items are not exactly same, but they are identifiable. • C = inferred from name, description or the documentation. • F = if it is not clear from anything above what dataset type it is. • n/a = If the resource type = dataset. NOTE: the service and dataset are mutually exclusive. In CMC the dataset does not have these categories. This is different in mobilityDCAT-AP, where service could be described by dataset categories. At least one of the parameters dataset or service detailed category must be present.
Dataset language	cs	Describes the language of the dataset contents (free text fields, enumerations, etc.). Format: 2 or 3-letter code of metadata language. (2-letter is preferred); Mandatory; multiple choice reference: CMC 2.2.2.7; MD <i>dcat:Dataset</i> -> <i>dct:language</i> Grading • A = specific field with list of language codes (2 letter language code accepted, e.g. cs not cz or cze); • B = specific field with "human" language descriptions or other nonpreferred codes (czech or cz or cze instead of cs); • C = inferred from sample or if possible, documentation.

Table 4 Content information

Parameter	Example of a value	Definition, comments
		F = if it is not clear, sample is not provided, or in the sample there is a mixture of different languages used interchangeably or if the language could not be discerned because of language ambiguity. NOTE: The actual language used vs metadata indication is not evaluated, if there is an indication of language in the metadata the sample is not inspected.
Georeferencing method	ALERT-C, OpenLR	Describes what location referencing methods are used by the dataset. The names are derived from DATEX II. Format: Enumeration (ALERT-C, Geocoordinates, georeferencing method, GML, ISO 19148, OpenLR, TPEG-LOC, Other); Recommended; Multiple Choice. Reference: CMC NONE; InQMS best practices; MD <i>dcat:Dataset -> mobilitydcatap:georeferencingMethod</i> Grading - A = specific field, items from the list. - B = described within the metadata or explicitly stated in the documentation (separate file). - C = inferred from the sample or unexplicit statement in the documentation - F = could not be found, there is predefined location set which is not <u>clearly</u> identified and not available NOTE: the values have been taken from mobilityDCAT-AP. It does not concern coordinates or reference system, just a location referencing method. Provide in the note info about explicit georeferencing method used.
Dataset documentation	../cs/sources/en-ndic_d2-common/	A separate file describing context, concepts, usage, guidelines, and examples, provides clarification of how the dataset is handled, information about architectural decisions, profiling, extensions, etc. Recommended. Format: URL to a document. Recommended. Reference: CMC NONE; InQMS best practices; <i>dcat:Dataset -> foaf:documentation (removed in MD)</i> Grading - A = separate file or webpage is present, contains useful info about the contents of the dataset/service. mainly: description, time concepts, location concepts, subscription, etc. - B = generic file (with combined info) or webpage, that contains all important information but not declared as dataset documentation or published for all datasets but identifiable. - C = separate or generic file or webpage, contains all important information, BUT provided only in local language (non-English) OR in English but not providing all required details.

Table 4 Content information

Parameter	Example of a value	Definition, comments
		- F = not provided or provided as promo material, leaflet, not really covering the documentation topics NOTE: Shall be present as an external file, but some useful info can be scattered across the catalogue record (or as a part of organization metadata). Evaluator downloads the file and screens its content for the expected information. There is ONLY evaluation of the documentation content NOT the comparing against reality. Minimum information is: - Identification, name, description, - Architectural decisions - Concepts of identity, time, location, message lifecycle, descriptions - Subscription information - (Profile of extension information – could be part of format docs) Documentation could be covered by one file containing documentation of the content, format and protocol.
Related linked data sources	x-source:en-ndic_d2-predefined-location-set	A dataset might for location referencing purposes refer to predefined location set (using identifiers from that dataset). Such predefined location set locations shall be accessible and shall be identified as a requirement for the successful decoding of the primary dataset. Format: URL or URI of the location dataset. Recommended Reference: CMC NONE; MD <i>dcat:Dataset</i> -> <i>dct:relation</i> Grading - A = specific field in expected format, an URL, or identifier to an accessible predefined location set. - B = information with URL or identifier to an accessible predefined location set available elsewhere, in provided documentation in samples. - C= inferred from documentation that “something” is needed, or if the sample uses some references to other datasets (e.g. location code list = important is that <u>no link is provided</u>). - F = from documentation, description (sample not provided) it is not clear if this is necessary or not. - n/a = if the locations are provided inline in the data / service and this is evident from documentation or sample or readily available data. NOTE: Key is if there is a catalogue record identical to the related predefined location dataset.

5.4 Temporal information

Information about the start and stop of the publication (when it becomes available to consumers and when it is removed).

Table 5 Temporal information

Parameter	Example of a value	Definition, comments
Start date of publication	2023-10-23T09:00:00+01:00	Describes from which date on, the dataset is available for download or service is active with the real / up to date content. Format: DateTime. Mandatory reference: CMC 2.2.3.1; MD <i>dcat:Dataset</i> -> <i>dct:temporal</i> Grading: - A = specific field with a timestamp in right format. - B = specific field, not filled in exactly using list items, but identifiable - C = inferred from description or documentation. - F = if not found or if found in documentation but it is not clear what the date means. NOTE: Either in the metadata or in case of missing parameter, inferred from the metadata date (grade = F).
End date of publication		Describes the date when data delivery of this publication terminates. Format: DateTime; Optional reference: CMC 2.2.3.2; MD <i>dcat:Dataset</i> -> <i>dct:temporal</i> - A = specific field with a timestamp in right format. - B = specific field with different than required datetime format. - C = inferred from description or documentation. F = not used. - n/a = when is date not filled in or not found in the documentation, meaning it is not set and therefore unknown to the data holder. NOTE: In most cases the end date of publication value will NOT be filled in, since the date is not known, therefore graded as n/a (not applicable). If it is filled in then it could be evaluated A-C, (F is not used). Usually, this value is known where there is a replacement dataset planned or already provided in parallel.

5.5 Geographical coverage

Information about the area and network coverage of the dataset.

Table 6 Geographical coverage

Parameter	Example of a value	Definition, comments
Area covered by publication	CZ0	Describes the geographic area covered by a dataset or a service. Multiple choice since they can be valid for more than one region. Format: NUTS codes http://data.europa.eu/nuts/, geonames, bounding box or GML polygon. Mandatory; multiple choice reference: CMC 2.2.4.1; MD <i>dcat:Dataset</i>-> <i>dct:spatial</i> Grading



Table 6 Geographical coverage

Parameter	Example of a value	Definition, comments
		- A = specific field in expected format is NUTS, geonames, bounding box or polygon. - B = specific field but not right format, e.g. referring to country name without the direct reference to code dictionary OR mixed hierarchical level of geo / NUTS codes is used, e.g. codes that include other codes. - C = inferred from the title or the description, or sample visualization available on the NAP. - F = not provided as previous options NOTE: If NUTS or geonames identifiers are used, only the lowest possible levels should be utilized. For example, for country coverage, only the country code is provided. For region coverage, only the region codes are provided, not the respective subregion codes. Mixed hierarchy codes e.g. Region and Subregion codes are provided together only if they do not intersect (the subregions do not belong to the referred to regions). Some datasets have coverage areas expressed in GML (polygons, rectangles) using ETRS (GPS) coordinates, which are equivalent to the NUTS codes.
Network coverage	Motorways, Arterial road network, Regional roads, Urban and local roads	Describes the hierarchical part of the transport network and modality that is covered by datasets. Format: enumeration: (Motorways; Arterial Road network (in the meaning of state roads or federal roads); Regional roads; Urban and local roads; Rail (long-distance or heavy-rail); Metro or light-rail network; Other public transport network; Waterways; Air network; other) . Mandatory, multiple choice reference: CMC 2.2.4.2; MD dcat:Dataset->mobilitydcatap:networkCoverage Grading: - A = specific field in expected format (a list items). - B = specific field, not filled in exactly using list items, but identifiable - C = inferred from name, description, from network coverage description (if filled in), or too generic e.g. road network - F = not provided or found via previous options NOTE: this is usually supplemented by textual description.
Network coverage description	the whole territory of the Czech Republic	Describes details of the transport network in addition to the element "Network coverage", e.g. specific roads covered. Format: Free text; optional reference: CMC 2.2.4.3; MD not used, could be part of description. Grading: - A = specific field filled, complementing network coverage item with identifiable description of roads or sub-areas covered - B = inferred from name, description (must complement the network list with additional information) or from the dataset/service documentation where is detailed map of the coverage.

Table 6 Geographical coverage

Parameter	Example of a value	Definition, comments
		- C = specific field filled in, but the information does not provide sufficient detail. E.g. copy network coverage field. OR in the documentation (as text) OR using Non-English language. - F = not provided in any of the form mentioned above NOTE: Could be in detail provided in documentation together with a map. If a textual description is given, it shall be in English. Provide in comment how extensive and descriptive the description is. Does not have to be provided but lowers the score if not.

5.6 Transport modes information

Information about transport mode that is covered by the dataset/service.

Table 7 Transportation system

Parameter	Example of a value	Definition, comments
Transport modes covered	Personal / car, truck	Describes covered transportation modes. Format: enumeration: (Air, rail (including high speed rail), conventional rail, light rail, long-distance coach, maritime (including ferry), metro, tram, bus, trolley-bus, shuttle bus, shuttle ferry, taxi, car-sharing, car-pooling, car-hire, bike-sharing, bike-hire, car, truck, motorcycle, cycle, pedestrian); Mandatory; Multiple choice. reference: CMC 2.2.5.1; MD dcat:Dataset -> mobilitydcatap:transportMode Grading: - A = specific field in expected format (a list items). - B = specific field, not filled in exactly using list items, but identifiable - C = inferred from name, description, from network coverage description (if filled in inferred from a road type) or documentation - F = not provided or found via previous options

5.7 Responsibilities

Information about data owner and publisher.

Table 8 Responsibilities

Parameter	Example of a value	Definition, comments
Publisher	John Doe, john.doe@data.eu, http://www.data.eu	The organization or person that publishes the dataset / service and is responsible for the content and concluding of a contract if applicable. Format: vCard-Textfields with the expected level of details (name, email, telephone, organization, position, address). Mandatory reference: CMC 2.2.6.1; MD dcat:Dataset -> dct:publisher Grading

Table 8 Responsibilities

Parameter	Example of a value	Definition, comments
		• A = dedicated contact information field, full information is provided (or at least name, position, email, telephone), responsibility clear as data publisher • B = a contact information field, full information is provided (or at least name, position, email, telephone), responsibility between data-owner and data-publisher and metadata publisher unclear, • C = inferred from documentation, responsibility is not clear as publisher contact, partial information is provided (only name and/or <u>generic</u> email). • F = not there, provision of a website URL or just a contact form without any contact information. NOTE: Could be part of the organization data or directly in the dataset description, could be contact info not in vCard format. Contact point, data / metadata publisher and owner are distinct roles that could be performed by different people and / or organizations. If just one contact is filled in, the responsibility is unclear. If they are same, they shall be filled in with same data.
Data owner	(copy from "publisher")	Describes the organization that owns the dataset and is responsible for the content and quality of the dataset. Could be the same as the publisher Format: vCard-Textfields with the expected level of details (name, email, telephone, organization, position, address). Mandatory reference: CMC 2.2.6.2; MD dcat:Dataset -> dct:rightsHolder Grading • A = dedicated contact information field, full information is provided (or at least name, position, email, telephone), responsibility clear as data owner • B = a contact information field, full information is provided (or at least name, position, email, telephone), responsibility between data-owner and data-publisher and metadata publisher unclear, • C = inferred from documentation, responsibility is not clear as data owner contact, partial information is provided (only name and/or <u>generic</u> email). • F = not there, provision of a website URL or just a contact form without any contact information. NOTE: Could be part of the organization data or directly in the dataset description, could be contact info not in vCard format. Contact point, data / metadata publisher and owner are distinct roles that could be performed by different people and / or organizations. If just one contact is filled in, the responsibility is unclear. If they are same, they shall be filled in with same data.

5.8 Conditions for use

Information about contract or license and conditions for use.

Table 9 Conditions for use

Parameter	Example of a value	Definition, comments
Contract or licence	License and Free of charge	Describes the condition of use: whether an unrestricted use is possible, a contract shall be concluded, or a licence shall be agreed on to use a dataset. Format: Enumeration: (No licence - No contract; Licence and Free of charge; Licence and Fee; Contract and Free of charge; Contract and Fee; Not relevant) . Mandatory reference: CMC 2.2.7.1; MD dc:Distribution -> dct:rights Grading: • A = specific field in expected format (a list item). • B = specific field, not filled in exactly using list item, but identifiable • C = inferred from name, description, documentation or conditions of use • F = not provided or found via previous options
Conditions for use	.../cs/providers/indic/#process-ddrcontract	If the option "Licence" or "Contract" in element "Contract or licence" is selected, the condition of use shall be clarified. Here a sample contract or the terms of use need shall be provided Format: an URL to a licence document, or standard licence reference. Conditionally reference: CMC 2.2.7.2; MD dc:Distribution -> dct:license Grading: • A = specific field in with link to a licence/contract and subscription process (could be separate), or name of the standardized licence e.g. CC-BY-4.0. • B = specific field with a link descriptive information in the text of the web page or to a file, but without subscription process description. • C = inferred from name, description, documentation, the licence is not explicitly stated, subscription process is either not defined or generic. • F = not provided or found via previous options • n/a = If the option "Licence" or "Contract" in NOT selected in the element "Contract or licence" then this parameter is not relevant

5.9 Access information

Information about the format and encoding of the provided dataset, as well as about the sample and validation schema and protocol information

Table 10 Access information

Parameter	Example of a value	Definition, comments
Data format - encoding	UTF-8	This describes the atomic element of the transfer syntax description. Systems that can go down to single bits are called "binary". Systems that are specified based on character standards where a single character has

Table 10 Access information

Parameter	Example of a value	Definition, comments
		more than one bit are specified by the name of the character standard used. Format: Enumeration: (ASCII; UTF-8; UTF-16; ISO-8859-1; ISO-8859-15; other) . Optional reference: CMC 2.2.8.1; MD dcat:Distribution -> cnt:characterEncoding Grading: • A = specific field in expected format (a list items). • B = specific field, not filled in exactly using list items, but identifiable • C = inferred from sample encoding • F = not provided or found via previous options
Data format - syntax	XML	This describes the base standard that specifies syntactically correct documents. On this level, only base elements of building documents properly are specified and can be proved by syntax checks. Enumeration: (XML; JSON; CSV; ASN.1 encoding rules; Protocol buffers; Other) . Mandatory reference: CMC 2.2.8.2; MD dcat:Distribution -> dct:format Grading: • A = specific field in expected format (a list items). • B = specific field, not filled in exactly using list items, but identifiable • C = inferred from sample / schema or live data (if available) • F = not provided or found via previous options
Data format - grammar	XSD	This describes standards on top of the elementary syntax that describe data structures in the dataset used for validation purposes. Format: Enumeration: (XSD; JSON Schema; ASN.1; Protocol buffers; other); Optional reference: CMC 2.2.8.3; MD dcat:Distribution -> mobilitydcatap:grammar Grading: • A = specific field in expected format (a list items). • B = specific field, not filled in exactly using list items, but identifiable • C = inferred from sample / schema or live data (if available) • F = not provided or found via previous options
Data format - Data Model	DATEX II profile	This describes concrete data models that use the specification elements so far to specify specific data models. The NAP should keep references to the concrete data model used for a dataset, e.g. including model versions. Format: Enumeration (many possible values including DATEX, NETEX, OCIT, GTFS, KML, OpenAPI, and others) . Mandatory reference: CMC 2.2.8.4; MD dcat:Distribution -> mobilitydcatap:mobilityDataStandard Grading: • A = specific field in expected format (a list items). • B = specific field, not filled in exactly using list items, but identifiable

Table 10 Access information

Parameter	Example of a value	Definition, comments
		- C = inferred from sample / schema or live data (if available) - F = not provided or found via previous options If not specified in metadata it could be inferred from the provided sample, schema or live data (if available).
Data format description	DATEX II v2.3/SituationPublication Extends: networkLocation	Used to provide additional information on the data format, e.g. are extensions used? On which class is the profile based, additional version information. Format: Free text, but in a short and structured way e.g. "MODEL – version/publication, extends: extended classes"; Optional reference: CMC 2.2.8.5; MD dcat:Distribution -> mobilitydcatap:dataFormatNotes Grading: - A = in specific field, version and, for DATEX: publication type and extension, - B = in specific field only version. - C = inferred from sample / schema, documentation - F = not provided or found via previous options NOTE: This is very important that the information here is structured, it is only field where the version of model could be specified.
Format documentation		A separate file describing context, concepts, usage, guidelines, and examples, information about architectural decisions, profiling, extensions, etc. Format: URL to a document. Recommended. Reference: NONE; dcat:Distribution -> foaf:page (documentation) // not in MD) Grading - A = separate file or webpage is present, contains useful info about the format of the dataset/service. mainly: schema and commented sample(s), profile extension information, format specifics. In English - B = generic file (with combined info) or webpage, that contains all important information but not declared as format documentation or published for all datasets but identifiable. In English - C = separate or generic file or webpage, contains all important information, BUT provided only in local language (non-English) OR in English but not providing all required details. - F = not provided or provided as promo material, leaflet, not really covering the documentation topics NOTE: Shall be present as a link to an external file or web document. Evaluator downloads the file and screens its content for the expected information. Ideally in English, directly usable descriptive well-structured text. There is ONLY evaluation of the documentation content NOT the comparing against reality.

Table 10 Access information

Parameter	Example of a value	Definition, comments
		Minimum information is: - Identification, name, description, - Basic concepts - Profile of extension information - Commented sample, schema Documentation could be covered by one file containing documentation of the content, format and protocol.
Data sample	.../cs/formats/enndic_d2-common-v1.1/samples/xxx	A link to a file with a sample content of the dataset, that allows consumers to evaluate how the real data would look like. Format: URL to an existing file; Recommended reference: NONE; MD dcat:Distribution -> adms:sample Grading: • A = separate file, sample on URL. • B = live data if available, • C = in documentation included in text (pdf, word), • F = not there NOTE: A carefully curated sample is preferred over live data because it clarifies the data owner's intentions and provides a complete set of expected information. This is not always the case with live data. Additionally, a sample allows for evaluating how closely the provider meets their intended goals.
Data schema / validation	.../cs/formats/enndic_d2-common-v1.1/schemas/xxx	A link to a file with a validation schema that is used to validate the structure of the dataset (and the sample). This is very important since the schema provides information about all potential values of the data elements and informs the consumer about pruning the model has undergone. Format: URL to an existing file. Mandatory reference: NONE; dcat:Distribution -> dc:conformsTo, in mobilityDCAT-AP as sub property dcat:Distribution -> mobilitydcatap:MobilityDataStandard -> mobilitydcatap:schema Grading: • A = separate file, sample on URL. • B = reference to schema is in live data, • C = in documentation included in text (pdf, word), • F = not there of reconstructed for the sample // we do not evaluate if this is full or trimmed (profiled) schema. NOTE: Could be replaced by a validation schema of the whole model, BUT a real specific schema is much better because it provides the intention of the data owner. If the schema must be inferred by the tools from the sample, then it is the same as if it is missing (grade F), such inferred schema is only a guess that real data will be the same and it is by no means a commitment by the data holder/owner.

Table 10 Access information

Parameter	Example of a value	Definition, comments
Access interface – application layer protocol	HTTP/HTTPS	Describes the IT protocol of the data interface that will be used to transfer data. Format: Enumeration: (SOAP; OTS2; HTTP/HTTPS; FTP; RSS; AMQP; MQTT; gRPC; other). Mandatory reference: CMC 2.2.8.6; MD dcat:Distribution -> mobilitydcatap:applicationLayerProtocol Grading: • A = specific field in expected format (a list items). • B = specific field, not filled in exactly using list items, but identifiable • C = inferred from data access URL or documentation • F = not provided or found via previous options
Security mechanisms	IP filter	Describes security mechanisms in place to protect the integrity and access to the dataset. Recommended; Multiple choice Format: Enumeration: (none; IP filter; basic; digest; certificate; url parameters; other); recommended Reference: CMC NONE; based on the InQMS best practices; MD NONE Grading: • A = specific field in expected format (a list items). • B = specific field, not filled in exactly using list items, but identifiable • C = inferred from documentation • F = not provided or found via previous options
Communication method	Push on occurrence	Describes the transmitting procedure from data provider to data receiver. It differs between push and pull; Mandatory Format: Enumeration: (Push; Push on occurrence; Pull); Mandatory reference: CMC 2.2.8.7; MD dcat:Distribution -> mobilitydcatap:communicationMethod Grading: • A = specific field in expected format (a list items). • B = specific field, not filled in exactly using list items, but identifiable • C = inferred from protocol documentation • F = not provided or found via previous options • n/a = for service (API)
Protocol documentation	../cs/protocols/indic_push-v1.0/xxx	Describes the protocol implementation and complements and extends information about application layer protocol, security mechanism, communication method, and others. Allows data users to program interface for data retrieval, which could be an example code. Format: URL to a document. Recommended. Reference: NONE; dcat:Distribution -> foaf:page (documentation) // not in MD) Grading

Table 10 Access information

Parameter	Example of a value	Definition, comments
		- A = separate file or webpage is present, contains useful information about the protocol, security, set up, organizational requirements example implementation, etc. in English - B = generic file (with combined info) or webpage, that contains all important information but not declared as protocol documentation or published for all datasets but identifiable. In English - C = separate or generic file or webpage, contains all important information, BUT provided only in local language (non-English) OR in English but not providing all required details. - F = not provided or provided as promo material, leaflet, not really covering the documentation topics Shall be present as a link to an external file or web document. Evaluator downloads the file and screens its content for the expected information. Ideally in English, directly usable descriptive well-structured text. There is ONLY evaluation of the documentation content NOT the comparing against reality. Minimum information is: - Identification, name, description, - Basic concepts - Security mechanisms, protocol description - Commented example implementation Documentation could be covered by one file containing documentation of the content, format and protocol.
Access URL	https://datex.rsd.cz	Provides a general link for access to the current dataset or a connection link to a service. If the access URL is unique for every single relation between the data owner and the data receiver, the access URL is linked to a subscription that enables access to the publication Format: url; Mandatory reference: CMC 2.2.8.8; MD dcat:Distribution -> dcat:accessURL (for service or link to distribution / dcat:downloadURL (for downloadable file) Grading: - A = an URL linking to the downloadable data-set file contains downloadable resource with data, or to an API from which the service could be directly consumed. Key issue is direct reference to the dataset/service. - B = link to data distribution platform, not to a directly consumable file or service, but at the platform datasets and directly referenceable (e.g. by same id) to NAP catalogue record. - C = link to data distribution platform, not to a directly consumable file or service, AND without direct linkage to the NAP catalogue record at the platform. - F = not provided or found via previous options, e.g. link to organizations website, to map visualization, to a protected website

Table 10 Access information

Parameter	Example of a value	Definition, comments
		without clear instructions how to get in (so the evaluator cannot assess the content of the URL) NOTE: If the catalogue is not providing the URL to get the data, how it could be consumed? At least leading to a data holders data distribution platform is required.

5.10 Quality information

Information about update frequency, data availability, collection method, verification method, reaction time, quality, national body assessment, and delegated regulations that this dataset partially satisfies.

Table 11 Quality information

Parameter	Example of a value	Definition, comments
Update frequency	On occurrence	Describes the update rate of the dataset. If there is a specific time interval or data only provided on occurrence precise information should be given. Format: Enumeration: (On occurrence; Up to 1min; Up to 5min; Up to 10 min; Up to 15 min; Up to 30 min; Up to 1h; Up to 2h; Up to 3h; Up to 12h; Up to 24h; Up to Weekly; Up to Monthly; Up to every 3month; Up to every 6month; Up to yearly; Less frequent than yearly). Mandatory reference: CMC 2.2.9.1; MD dcat:Dataset -> dct:accrualPeriodicity Grading: • A = specific field in expected format (a list item). • B = specific field, not filled in exactly using list items, but identifiable • C = inferred from name, description or documentation • F = not provided or found via previous options
Quality description	quality information is unknown	Describes the methods and results of a quality assessment. Within such an assessment, individual quality criteria of a dataset are checked and compared with pre-defined quality requirements. Format: free text, containing a general description and possibly a link (URL) of the document with more information; Mandatory reference: CMC 2.2.9.2: CMC NONE; MD dcat:Dataset -> dqv:hasQualityAnnotation • A = link to a file or separate description with evaluation; according to EU-EIP quality packages. • B = Link to a file or separate description that describes how quality is evaluated and contains evaluation some basic directly observable quality parameters, according to EU-EIP quality packages. • C = any description focusing on how quality is evaluated but without actual evaluation commitments / threshold values, partial alignment with quality packages is expected, inferred from documentation or separate quality document.

Table 11 Quality information

Parameter	Example of a value	Definition, comments
		F = in all other situations, e.g. stating "quality is unknown" or "quality is high" or similar is also F NOTE: The quality description could contain also the information on availability, data collection method, cross verification, reporting process and reaction time as stated below. In the comment field, the evaluator shall comment on the languages of the quality information and its details
Availability	24/7	Describes expected time period availability of the dataset / service as expected by the publisher in terms of hours/week. It is not actual availability expressed in % (of the time when was service available / time when the service was expected to be available). Format: Free text (24/7; other; ...). Recommended reference: CMC NONE; MD dcat:Dataset -> dqv:hasQualityAnnotation Grading: - A = specific field, with time information. - B = filled in "<u>quality description</u>" metadata, structurally separated and identifiable - C = inferred from name, description or documentation, time information - F = not provided or found via previous options
Data collection method	Sensor input, police, ...	Indication and description of potential data sources from which this dataset was created. Format: Free text describing data sources.; recommended reference: CMC NONE; MD dcat:Dataset -> dqv:hasQualityAnnotation Grading: - A = specific field, with list of different data sources (e.g. police, induction loops, ...) and their processing method (combined automatically, manually, ...). - B = filled with sources and method in "<u>quality description</u>" metadata, structurally separated and identifiable - C = inferred from description or documentation, with at least sources - F = not provided or found via previous options
Cross verification	Cross verified	Indicates whether data has been cross-verified with one or more additional sources. Format: Enumeration: (Cross verified, not cross-verified; other); Recommended reference: CMC NONE; MD dcat:Dataset -> dqv:hasQualityAnnotation Grading: - A = specific field, in expected format (a list items). - B = filled in "<u>quality description</u>" metadata in expected format, structurally separated and identifiable OR in the specific field but not filled in exactly using list items, but identifiable - C = inferred from description or documentation, with at least status - F = not provided or found via previous options

Table 11 Quality information

Parameter	Example of a value	Definition, comments
Problem reporting process	../cs/providers/en-ndic/xxx	Description of how the problem reporting is handled, individual steps, contact. Format: Free text describing the process or link to a description; recommended reference: CMC NONE; MD dcat:Dataset -> dqv:hasQualityAnnotation Grading: • A = specific field with actionable process description. In English • B = actionable process description filled in “<u>quality description</u>” metadata or as part of dataset or provider documentation in English • C = mentioned in the metadata or documentation but without sufficient quality, e.g. only email is provided explicitly for problem reporting but no process is described or provided only in local language (non-English) • F = not provided or found via previous options NOTE: Describe in a comment if is description in multiple languages
Reaction time to respond	3 working days	If a Problem reporting process is provided, it must not be empty. Describe when, on average, the reporter shall expect the first reaction. Format: Free text = time information expected; Recommended reference: CMC NONE; MD dcat:Dataset -> dqv:hasQualityAnnotation Grading: • A = specific field with reaction time information. In English • B = reaction time information description as part “<u>quality description</u>” metadata. In English • C = reaction time inferred from documentation. or provided only in local language (non-English) • F = not provided or found via previous options • n/a = if “Problem reporting process” is Graded as F NOTE: Describe in a comment if is description in multiple languages
National body assessment status	-	Used to indicate the history and status of such an assessment. It may include the date and the result of the recent assessment procedure. Format: free text (date or a link (URL) of the document with more information); Optional. reference: CMC 2.2.9.3; MD dcat:Dataset -> mobilitydcatap:assessmentResult Grading: • A = specific field with an URL of assessment info in English OR with a status (compliant, non-compliant, compliant with exception) and date OR info that it has not been executed in English. • B = specific field with information (directly or in linked docs) only about the result (compliant, non-compliant, compliant with exception) in English

Table 11 Quality information

Parameter	Example of a value	Definition, comments
		• C = specific field with information (directly or in linked docs) only about the date of assessment • F = not provided or found via previous options
Applicable DRs	SRTI, RTTI	Describe which delegated regulations (DR) are being fulfilled by the dataset, it might be just partial fulfilment and the dataset could, partially, fulfil more than one DR. Format: enumeration: (SSTP; SRTI; RTTI; MMTIS; AFIR); Recommended reference: CMC NONE; MD dcat:Dataset -> dcatap:applicableLegislation Grading: • A = specific field in expected format (a list items). • B = specific field or other metadata such as “tags”, not filled in exactly using list items (e.g. Numbers instead of short names, or actions (885/2013 or Action a?) but identifiable • C = inferred from documentation or provision of self-assessment • F = not provided or found via previous options

6 Overall evaluation

The overall evaluation is provided as

- overall evaluation grade (A, B, C, F, or n/a) and
- overall comments that summarize the main findings of the evaluator about the service.

See the figure below.

	A	B	C	D	E	F	G	H
1	Evaluation of formal documentation quality							
2	version 2024-09-02							
3	overall quality							
4	samples, schemas and detailed documentation of format and protocol provided. However many of the parameters required by CMC were either completely missing or had to be derived from textual descriptions.							
5	Category	Parameter Name	ref.	obligation	evaluation	data provided or implied	implied	comment
6		content provider contact			-	https://registr.dopravniinfo.cz/en/providers/cz-ndic/		
7	Evaluation metadata	evaluated dataset url			-	https://registr.dopravniinfo.cz/en/sources/cz-ndic_d2-common/		THIS is a sample evaluation done for updating and evaluation of the assessment. It does not have to be complete nor TRUE!
8		evaluation date			-	01.12.2019		
9		evaluator			-	Petr Bureš		
								data formally not provided - not part of the metadata set

Figure 9 Final overall grade and commentary.

6.1 Overall commentary

The evaluator, in few sentences describes the background for the overall grade. This commentary highlights strongest and weakest point of the evaluated metadata.

Example:

Samples, schemas and detailed documentation of format and protocol provided. However, many of the parameters required by CMC were either completely missing or had to be derived from textual descriptions.

6.2 Overall grade computation

The overall grade (1) in the figure above is computed automatically from grades of individual parameters. The computation is performed in the Evaluation form on the sheet “support”. The mechanics of the computation is described in subchapters below.

6.2.1 Individual grades and their part in overall grade computation

Evaluation grades of each parameter, expressed as grades **A, B, C, and F**, by the evaluator, are converted into numeric values.

Table 12 Mapping from the mark A-F to a numeric value

classification	A	B	C	F	Not relevant
value	1	2	3	4	0

Each parameter contributes to the overall score differently based on its **status** and **importance**, by **weighing** the parameter and **masking** the nonrelevant parameters, see next tables.

Table 13 Mapping of mandatory appearance to a value

“status”	optional	recommended*	partial**	conditional	mandatory
value	1/4	1/2	3/4	1	1



* Recommended is for parameters that are not originally from CMC, except schema which is set as mandatory

** Partial is for contact fields of a vCard type

Table 14 Mapping of a weight (importance) to a value

Importance (weight)	Normal	Important	Very important
Value	1	2	3

Most of the parameters are of normal importance and only a few are tagged “important”.

Computation of the individual **score** = **classification*status*weight**; meaning that nonrelevant parameters (with classification = 0) are not counted, their value is 0.

Table 15 Classification of parameters

Category	Element Name	classification	status	weight	count	score
Metadata information	Metadata date	3	1	1	1	3
	Metadata language	3	1	1	1	3
	Contact point	2	0,75	1	0,75	1,5
	Dataset identifier	1	0,5	1	0,5	0,5
Content information	Name of the dataset	1	1	1	1	1
	Description of dataset	1	1	1	1	1
	Resource type	1	1	1	1	1
	Dataset type category	2	1	1	1	2
	Dataset detailed type	2	0,25	1	0,25	0,5
	Service type category	0	1	1	0	0
	Dataset language	1	1	1	1	1
	georeferencing method	2	0,5	1	0,5	1
	Dataset documentation	3	0,5	1	0,5	1,5
	Related linked data sources	1	0,5	2	1	1
Temporal information	Start date of publication	2	1	1	1	2
	End date of publication	0	0,25	1	0	0
Geographical coverage	Area covered by publication	1	1	1	1	1
	Network coverage	1	1	1	1	1
	Network coverage description	4	0,25	1	0,25	1
Transportation system	Transportation modes covered	1	1	1	1	1
Responsibilities	Publisher	2	0,75	1	0,75	1,5
	Data owner	2	0,75	1	0,75	1,5
Conditions for use	Contract or license	1	1	1	1	1
	Conditions for use	1	1	1	1	1
Access information	Data format - Encoding	1	0,25	1	0,25	0,25
	Data format - Syntax	1	1	1	1	1
	Data format - Grammar	1	0,25	1	0,25	0,25
	Data format - Data Model	1	1	1	1	1
	Data sample	1	0,5	1	0,5	0,5
	Data schema / validation	1	0,5	2	1	1

Table 15 Classification of parameters

Category	Element Name	classification	status	weight	count	score
Access information (continued)	Data format description	1	0,25	1	0,25	0,25
	Data Format Documentation	1	0,5	1	0,5	0,5
	Access interface	1	1	1	1	1
	Security mechanisms	3	0,5	1	0,5	1,5
	Communication method	1	1	1	1	1
	Protocol documentation	2	0,5	1	0,5	1
	Access URL	2	1	2	2	4
Quality information	Update frequency	1	1	1	1	1
	Quality description	4	1	1	1	4
	Availability	4	0,5	1	0,5	2
	Data collection method	4	0,5	1	0,5	2
	Cross verification	3	0,5	1	0,5	1,5
	Problem reporting process	4	0,5	1	0,5	2
	Reaction time to respond	4	0,5	1	0,5	2
	National body assessment status	4	0,25	1	0,25	1
	Applicable DRs	1	0,5	1	0,5	0,5

6.2.2 Overall grade computation

The numerical overall evaluation is computed as the arithmetic average of all the grades awarded, with the possibility of assigning a weight to each grade:

- Numerical value = (sum of scores) / (sum of placement * weight, where classification > 0)

The final evaluation grade, expressed as **A, B, C, and F**, is converted from the numerical value according to the formula in next table.

Table 16 Mapping scoring 1-4 to the grade A-F

type/mark	A	B	C	F
mark	<1,1.5>	(1.5,2.5>	(2.5,3.5>	(3.5,4)

6.3 Method of presentation

The outputs are presented in Report (MS Excel format)

Evaluation of formal documentation quality					
[FRM_web-manual-doc-formal]			v20240616		
overall quality		B	samples, schemas and detailed documentation of format and protocol provided. However many of the parameters required by CMC were either completely missing or had to be derived from textual descriptions.		
Category	Parameter Name	evaluation	data provided or implied	implied	comment
Evaluation metadata	content provider contact	-	https://registr.dopravniinfo.cz/en/providers/cz-ndic/		
	evaluated dataset url	-	https://registr.dopravniinfo.cz/en/sources/cz-ndic_d2-common/		
	evaluation date	-	01.12.2019		
	evaluator	-	Petr Bureš		
Metadata information	Metadata date	F	01.12.2019	yes	data formally not provided, not part of the metadata set. Date was derived from date accessed (web page)
	Metadata language	B	cze;eng	yes	not part of the metadata set. Date was derived from context (web page)
	Contact point	B	name: Ctirad Weissmann title: Manager of traffic information distribution department phone: +420-596663578 email: ctirad.weissmann@rsd.cz		data derived from the web, provider information section, it is however unclear if this contact is also metadata provider (the person who put metadata together).
	Dataset identifier	A	x-source:cz-ndic_d2-common		dataset uri
Content information	Name of the dataset	A	DATEX II Situation Publication - Common		within a dataset specification [cze;eng]
	Description of dataset	A	Common traffic information (accidents etc.), provided by NDIC for the whole Czechia, pushed on occurrence, in a common traffic information profile of DATEX II format.		within a dataset specification [cze;eng]
	Resource type	A	Data set	yes	Directly in metadata as literals (listed in a table)
	Dataset type category	A	Road work, Unexpected road events and conditions, Real-time traffic data	yes	Directly in metadata as literals (listed in a table)
	Dataset detailed type	A	Long-term road works, Short-term road works, Accidents and incidents, Poor road conditions, Weather conditions affecting road surface and visibility, Location and length of queues	yes	Directly in metadata as literals (listed in a table)
					not applicable for dataset only services. The value is however not part of the metadata since

Figure 10 Presentation of the resulting evaluation in a FORM, (clipping)

The report serves as a record of the quality of the dataset information provided and is used for:

- internal checking of the correctness and integrity of the provided metadata,
- Feedback to the data owner/provider (here the ideal form is a table or text file).

7 Future work and conclusions

The presented methodology provides a detail instructions on how to evaluate the quality of metadata that are available on the NAP, describing a particular dataset or service. This methodology should be used as feedback to the data holder, the entity responsible for filling in metadata for the dataset / service for improvement. The main users of the methodology are:

- **Competent authorities:** when performing the compliance assessment of the data holder's dataset.
- **Data holders:** to conduct in-house check
- **Data users:** when assessing the quality of consumed data for their internal purposes

The presented methodology is largely based on the Coordinated Metadata Catalogue, matched where possible with mobilityDCAT-AP. However, the structure and strict nature of the mobilityDCAT-AP will, in the future, necessitate a different type of evaluation where some checks could be done automatically, bringing it closer in some aspects to the Metadata Quality Assessment Methodology (MQA⁶) used by the European Open Data Portal.

⁶ data.europa.eu/mqa/methodology?locale=about



8 References

The evaluation criteria are based on the following project results:

- [1] CMC Coordinated Metadata Catalogue, describing agreed metadata developed by EU-EIP (<https://gitlab.com/tamtamresearch/inqms/evaluation-procedures/-/tree/master/doc>)
- [2] Project InQMS evaluations and methodology (<https://inqms.tamtamresearch.com/>)
- [3] NAPCORE project (<https://napcore.eu>)
- [4] mobilityDCAT-AP, extended and consolidated metadata catalogue developed by NAPCORE (<https://mobilitydcat-ap.github.io/mobilityDCAT-AP/releases/index.html>)
- [5] MQA, methodology for DCAT-AP metadata catalogue assessment, developed by SEMIC (<https://data.europa.eu/mqa/methodology?locale=en>)
- [6] Quality recommendation by W3C and by open data portals (data.gov.cz and others) (<https://opendata.gov.cz/%C5%A1patn%C3%A1-praxe:start>)

