



A Quality Framework for Alternative Fuels Data

Report by Task 3.2: European NAPs data quality

Status: Final
Version: 1.0
Date: 03/06/2025

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)	João Montenegro, ARMIS (PT); Lígia Conceição, ARMIS (PT);
Co-author(s)	Jasper Beernaerts; (NGI-BE), Sara Linhas, ARMIS (PT)
Related to Milestone no.	M3.8
External Milestone	Yes

Document history

Version	Date	created/ modified by	Comments
0.1	18/07/2024	João Montenegro	First draft
0.2	03/01/2025	João Montenegro	Second draft, EC feedback
0.3	20/02/2025	João Montenegro	Third draft
0.4	17/04/2025	João Montenegro	Fourth draft, partner review
1.0	12/05/2025	João Montenegro	Final draft
1.1	03/06/2025	Peter Lubrich	Activity Leader Review

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 is part of a set of quality frameworks being developed under the scope of NAPCORE WG 3. It comprises a Quality Framework dedicated to Alternative Fuels (AF) data. The document contains a detailed description of quality criteria and metrics aligned with the characteristics and elements of Alternative Fuels.

Abbreviations

Abbreviation	Meaning
AF	Alternative Fuels
AFIR	Alternative Fuels Infrastructure Regulation
DR	Delegated Regulation
CEAP	Common European Access Point
EAFO	European Alternative Fuels Observatory
EC	European Commission
EU	European Union
IR	Implementing Regulation
IDACS	ID and Data Collection for Sustainable Fuels in Europe
MS	Member State
NAP	National Access Point
NAPCORE	National Access Point Coordination Organisation for Europe
QF	Quality Framework
RTTI	Real-Time Traffic Information Services



A Quality Framework for Alternative Fuels data

STF	Sustainable Transport Forum
WG	Working Group
UVAR	Urban Vehicle Access Restrictions



Table of contents

1. Introduction	5
1.1. Scoping.....	5
1.2. Methodology and structure	6
2. Background	8
2.1. Literature review on data quality dimensions	8
2.2. Previous and ongoing EU initiatives that relate to AF data quality.....	9
IDACS project.....	9
Sustainable Transport Forum	10
AFIR and the Implementing Regulation (IR) on data.....	10
3. Proposal for Quality Criteria	11
4. Proposal for Quality Assessment Methods.....	14
4.1. Generic approach	14
4.2. Aggregated view on quality criteria and metrics	14
5. Conclusions	16
References	18

List of tables

Table 1 - RTTI data types referring to AF (based on [2])	5
Table 2 - Description of resulting quality dimensions (adapted from [5])	8
Table 3 - Description of quality criteria and metrics.....	12



1. Introduction

The provision of high-quality Alternative Fuels (AF) data is critical to supporting the growing number of initiatives designed to advance sustainable transport across Europe. However, the lack of clear data quality definitions and standards has hindered stakeholders from effectively providing AF data through the National Access Points (NAPs). As highlighted by different AF-related projects and other tasks within NAPCORE, many NAPs currently lack any AF data. While this situation is expected to change, as the Alternative Fuels Infrastructure Regulation (AFIR) Article 20 on Data provision starts to apply on 14 April 2025, this gap is further exacerbated by technical challenges, including varying NAP architectures, cybersecurity concerns, and the absence of comprehensive data quality guidelines. Recognizing these barriers, the European Commission (EC) has launched several initiatives aimed at improving the discoverability and accessibility of AF data. In line with these efforts, NAPCORE's Working Group (WG) 3 has been tasked with developing a dedicated quality framework for AF data, which is essential for ensuring data reliability and fostering wider AF adoption. This report outlines the development of that framework, detailing the methodology and structure behind it.

1.1. Scoping

Alternative Fuels are defined by AFIR, EU 2023/1804 [1], as *“fuels or power sources which serve, at least partly, as a substitute for fossil oil sources in the energy supply used for transport and which have the potential to contribute to its decarbonisation and enhance the environmental performance of the transport sector”*. According to Article 7 of the Delegated Regulation (DR) n° 2022/670 [2], which refers to Real-Time Traffic Information (RTTI) and is applicable from 2025 on, recharging and refuelling-related stakeholders must publish AF-related data on the National Access Points (NAPs) in DATEX II format. In its Annex, the regulation provides the list of applicable data types, among which the ones highlighted in Table 1 refer specifically to AF:

Table 1 - RTTI data types referring to AF (based on [2])

Category	Data type
The types of data on infrastructure	location of recharging points for electric vehicles and conditions for their use
	location of compressed natural gas, liquefied natural gas, liquefied petroleum gas stations
	location of refuelling points and stations for all other fuel types
The types of data on the real-time use of the network	availability of recharging points and stations for electric vehicles
	availability of refuelling points and stations for alternative fuel types
	price of ad hoc recharging/refuelling

However, since the RTTI DR has a broad scope, it does not delve deeply into AF-related aspects. Therefore, this list of data types has been revised and expanded by the AFIR [1]. Article 20 of AFIR provides a more detailed list of data types for AF infrastructure, avoiding future overlaps with RTTI and, thus, representing the legal framework for AF data. However, the second paragraph of Article 20, containing information on data types, has been concretised by the Implementing Regulation laying down rules for the application of AFIR [3], resulting in a more comprehensive list of 7 categories comprising 44 data types, as listed in the Annex of this Regulation.

Although these regulatory frameworks require stakeholders to provide such data (or make it available via metadata) through the NAPs, some issues prevent them from accomplishing this task. Findings from the ID and Data Collection for Sustainable Fuels (IDACS) project, as well as internal research coming from other NAPCORE activities (e.g., WG2's AF demonstrator), pointed out that nearly half of the existing NAPs do not possess any AF (meta)data, and numbers are even lower when assessing fuels



other than electric. Research conducted by the Sustainable Transport Forum (STF) group (Activity 1 of the “**Sub-group on a common data approach for electromobility and other alternative fuels**”[4]) identified several challenges that prevent AF data provision, such as the existence of different NAP architectures, cyber security issues, lack of technical implementation guidelines, and the inexistence of clear data quality rules across different actors in the data value chain.

In that sense, different initiatives are supported by the EC with the view to overcoming these challenges. One example is the creation of a Common European Access Point (CEAP) for Alternative Fuels data, settled by the EC to be in operation by 2027 [1]. The development of different sub-groups (under the umbrella of the STF group) dedicated to fomenting the adoption of AF and the discoverability of AF data is another important endeavour. NAPCORE has also taken up relevant activities in this regard, one of them being the AF demonstrator, which is under development by WG2. Additionally, to tackle one of the key challenges highlighted by the STF group – related to the lack of clear data quality rules – NAPCORE WG3 partners are entitled to develop a QF dedicated to AF data, a task which is described throughout this report, which follows the method and structure explained in section 1.2.

1.2. Methodology and structure

The first step was to carry out a comprehensive literature review to select the quality dimensions that would need to be considered while developing the QF. In that sense, information gathered on data quality was analysed through an extensive review of journal articles, conference papers, master's theses, book chapters, and standards. This comprehensive literature review identified a large set of quality dimensions, which were organised into 9 major categories. Each dimension comprises a set of sub-dimensions that share commonalities but also exhibit distinct characteristics.

While the first phase was more theoretical and grounded in research, the next step involves defining quality criteria and use case metrics, bringing us closer to the practical formulation of an AF Data Quality Framework. This process began with formulating generic questions applicable across various data categories within the mobility and transport domain. These questions served as the baseline, offering insights for later producing the use case metrics specific to AFs. Additionally, they were designed with the potential for applicability in other data quality frameworks, as developed by NAPCORE (see the NAPCORE Online Quality Repository¹).

To understand the particularities of AF data, we examined other ongoing and previous initiatives related to the digitalisation and exchange of such data types. In this phase, the quality sub-dimensions and generic questions identified earlier are translated into a series of actionable questions, which can subsequently be processed to determine a data quality level.

Recognising that data quality is context-dependent, our approach emphasises flexibility, enabling the framework to be tailored to specific needs. Thus, each sub-dimension identified in the literature was assessed to evaluate its suitability for application to AF datasets. This assessment focused on tailoring the identified quality dimensions and sub-dimensions to the context of AFs. Drawing on insights from previous initiatives, such as the IDACS project, ongoing efforts like the STF Group, as well as legislative developments, we formulated AF-specific criteria tailored to the domain's characteristics. It is important to note that not all dimensions are included, as some were either not applicable to AF data or considered not feasible to measure. For each criterion, one or more metrics were designed.

After consolidating a set of criteria and metrics, we developed a simple assessment method that allows us to obtain quality indicators for each quality dimension, as well as for the dataset as a whole. The

¹ <https://napcore.eu/quality/>



method is a basic scaled weighted average that takes the evaluator's value as input and provides quality indicators as output.

This report consolidates the outcomes of the several steps needed to achieve this QF and provides a concrete example of how to apply it. It was organised into 5 sections:

- 1. Introduction – including crucial AF-related definitions and explaining the process beneath the formulation of the quality assessment methods;
- 2. Background – exploring the main outcomes of the literature review on data quality dimensions, as well as gathering important information from other AF-related initiatives;
- 3. Quality criteria – demonstrating the results of the criteria and metrics developed, along with important clarifications;
- 4. Assessment methods and explanation regarding the assessment method proposed;
- 5. Conclusions – including the main conclusions, lessons learnt, limitations and suggestions for future work.



2. Background

2.1. Literature review on data quality dimensions

In today's data-driven landscape, the importance of accessible, reliable, and accurate data cannot be neglected. High-quality data is fundamental to informed decision-making and efficient operations across various sectors. However, ensuring data quality is complex due to differing opinions on what constitutes data quality dimensions and the lack of systematic methodologies for developing quality frameworks. Despite extensive research over the decades, there is no consensus regarding the precise definition of data quality and its constituent dimensions.

To support a robust data quality framework, existing literature has been revisited, including scientific articles, book chapters, and standards that explore data quality dimensions. This comprehensive literature review identified 66 quality dimensions, which were then organised into a coherent structure contributing to developing effective quality frameworks. This effort resulted in a conference paper presented at the 6th International Conference on Industry 4.0 and Smart Manufacturing [5].

The outcome of this review is a consolidation of 37 quality dimensions into the following 9 major categories, which are described in Table 2. Other dimensions were also outlined in the original paper, but they were deemed to have either the same or opposite meaning as the included ones and, therefore, were combined with them.

Table 2 - Description of resulting quality dimensions (adapted from [5])

No.	Category/Group	Description/dimensions included
1	Openness	Comprises Accessibility ("how easily and quickly accessible data is") and Availability ("extent to which information is available for download and in which formats").
2	Exactness	Includes Accuracy ("how well the data values stored in a database align with real-world values"), Precision ("degree of detail and fitness of the measurement"), and Correctness ("the characteristic of data to be free of error or mistakes").
3	Comprehensiveness	Deals with the wholeness and extensiveness of data. It includes Completeness ("the capacity of data to represent every meaningful state of the corresponding real-world environment"), the Amount of Data ("the quantity or volume of available data being appropriate and able to significantly represent reality"), Granularity ("degree to which the data elements can be divided"), Data Coverage ("the availability and comprehensiveness of data compared to the total universe of real-world states"), and the Variety of Data and Data Sources ("the variety of data and data sources that are available for the intended task").
4	Coherence	Includes Consistency in data itself ("equivalence of information across data stores, applications, and systems") and the Consistent Representation of data ("presentation of information in a uniform format that aligns with previous data, semantic rules and format"), Data Specification ("compliance with any existing legal or regulatory obligations, data standards, or business rules"), and Reliability ("the capacity of data to maintain a specified level of performance when used on specified condition").
5	Veracity	Includes Believability, which is defined as the "extent to which information is regarded as true and credible by the user", Traceability ("ability to track all access and changes to the data"), Objectivity ("extent to which data are unbiased unprejudiced and impartial"), and Reputation ("reputation, trustworthiness, or credibility of the source, provider or content").
6	Promptness	Gathers dimensions focused on the frequency and speed of data exchanges to ensure it is up to date. It joins Timeliness ("time between the data becoming into existence and its availability for use"), Currency ("fit between



		the data's age or currency with its intended use”), and Volatility (“the period of validity in the real world, which the data aims at representing”).
7	Usability	Includes Understandability (“how clear, without ambiguity, and easily comprehended data is”), Conciseness (“how well information is compactly represented”), Interpretability (“standardizing language and representations, facilitating clear communication”), Ease of Manipulation (“assesses how data can be manipulated and applied consistently”), Ease of Operation (“evaluates the ease with which the data can be updated, maintained, and managed”), and Navigation (“assesses how easily data can be found and linked across databases”).
8	Safekeeping	Comprises Security (“extent to which access to information is restricted appropriately to maintain its security”), followed by Safety (“capability of the function to achieve acceptable levels of risk of harm”), and Integrity (“a measure of the fundamental/basic characteristics of the data”).
9	Relevance	Includes Relevance - described as “extent to which information is applicable and helpful for the desired use”-, Effectiveness (“ability to achieve specified goals with accuracy, completeness, and at a viable cost in a specified context of use”), Efficiency (“ability to quickly meet the information needs for a specific task”), Value added (“evaluates the extent which data provides advantages through its use”), Transactability (“degree to which data will produce the desired business transaction or outcome”), Usefulness (“fitness of data for a specific use”), and Uniqueness (“lack of duplicates records for the same event”).

It's crucial to recognise that not all dimensions must be applied universally when developing a data quality framework. The concept of data quality as "fitness for use" reinforces the importance of context and the specific needs of data consumers. Different applications and domains may prioritise certain dimensions over others based on their specific requirements and operational environments. In this sense, the presented AF QF has concretised and tailored the above-mentioned categories in the context of AF data exchange. This means that some categories have been translated into one or several quality criteria, and a few have been neglected. This means that this QF allows organisations to further tailor their definitions to their specific needs.

This approach provided guidelines for defining this Alternative Fuels Data Quality Framework. Nevertheless, the approach is flexible enough to be also applicable to other quality frameworks and contexts. This means that organisations across different sectors can leverage the same principles to develop data quality frameworks that address their specific challenges and goals (as done for, e.g., the UVAR QF developed within NAPCORE).

2.2. Previous and ongoing EU initiatives that relate to AF data quality

IDACS project

The IDACS (ID and Data Collection for Sustainable Fuels in Europe) project, initiated in 2019 and funded by the Connecting Europe Facility under the supervision of the European Commission's Directorate-General for Mobility and Transport, aimed to enhance the adoption of alternative fuels – such as electricity, hydrogen, and biofuels by improving consumer access to information. Fifteen Member States (MSs) collaborated to collect and standardise data on the location and availability of alternative fuel recharging and refuelling points, making this information accessible through the NAPs. The project's outcomes have been integrated into the AFIR, contributing to a more cohesive and user-friendly alternative fuel infrastructure across Europe.



Sustainable Transport Forum

The EU's Sustainable Transport Forum (STF) focuses on enhancing data quality and interoperability in AF infrastructure. It specifically aids in the collection, sharing, and refinement of critical data regarding electric mobility and alternative fuels. This initiative ensures that data around fuel station locations, availability, pricing, and technical standards meet high standards, which is essential for developing the infrastructure to support the widespread adoption of these fuels across Europe.

Key efforts include overseeing data provision and quality control through the European Alternative Fuels Observatory (EAFO), where the STF also facilitates data standardisation and publishes recommendations to guide policy and technical standards for alternative fuel usage across sectors like road, rail, and aviation. EAFO acts as a central repository that supports government agencies and stakeholders in tracking the deployment of alternative fuel stations, measuring data reliability, and promoting public access to high-quality data across EU countries. This improves both transparency and accessibility for consumers and stakeholders in the alternative fuels ecosystem.

Specifically related to alternative fuels data quality, STF provides the following:

- **Guidance on Data Standards and Quality:** STF helps establish standardized data formats and protocols to ensure consistent and high-quality data on alternative fuels infrastructure. This ensures that data is reliable and interoperable across EU member states.
- **Collaboration Platform:** STF serves as a collaborative forum for stakeholders, including national authorities, industry players, and infrastructure providers. This cooperation fosters sharing best practices and supports consistent alternative fuel infrastructure expansion.

AFIR and the Implementing Regulation (IR) on data

The AFIR is an important regulation of the European Union (EU) that sets milestones for the deployment of AF infrastructure across the EU. The IR, a further specification of the AFIR framework, focuses on the standardisation and management of data related to alternative fuel infrastructure, supporting the establishment of a coherent framework for the collection, management, and sharing of data on AF [3].

The IR introduces a set of requirements concerning quality-related aspects on data. These are: 'Completeness,' 'Correctness,' 'Consistency,' 'Timeliness,' and 'Reliability.' In addition, the IR has a dedicated annex where the whole set of data types is further described, including a detailed description of each data type, as well as information such as data level and data format.

The recommendations outlined in the IR served as a major input for building the quality criteria in this study. However, some labels used in this work for quality dimensions and criteria differ from those in the legislation. This divergence reflects the findings from the literature review in chapter 2.1, which demonstrated that certain academic definitions of data quality terms were more appropriate to describe the content of these requirements. Detailed examples of these adaptations are provided in chapter 3, where the criteria, metrics, and their respective quality dimensions are fully described.



3. Proposal for Quality Criteria

The quality criteria for Alternative Fuels have been defined following the process described in Chapter 1.2. As highlighted in that chapter, not all quality dimensions defined by Carvalho et al. [5] were included. In particular, we could not identify applicable or feasible criteria for the dimensions included in the categories “Veracity”, “Safekeeping”, and “Usability”. It is worth mentioning that, although the STF group report highlighted the importance of avoiding overwhelming or redundant data in AF datasets (which could fit into the “Conciseness”, a dimension part of the “Usability” group), we could not find any measurable criteria in this regard. A similar situation was observed for the “safekeeping” group, as the STF group indicated that AF datasets should be subjected to some degree of control and monitoring, but no applicable metrics to assess this “monitoring” were identified.

On the other hand, for the dimensions included in the other 6 dimension groups, we were able to identify 8 criteria, encompassing 14 metrics. The summary description of the criteria and respective metrics is provided in Table 3, which compiles the proposed quality dimensions, criteria and metrics, including their semantical definitions, mathematical formulations and value ranges.

For the proper usage of the criteria and metrics, some additional explanations are given in the following paragraphs.

For the dimensions included in the first group (“Openness”), we elaborated on three criteria considered adequate for the “Accessibility” dimension in the context of AFs. One refers to a broader view of “how accessible” the dataset is, being translated into a multiple-choice metric which assesses the existing access conditions for the general public. Another criterion contains a single binary metric to verify if the dataset is made available through its respective NAP; as recommended by AFIR and reinforced with the IR. AF data availability through the NAP is crucial to ensure its discoverability by Service Providers, for example. The third criterion verifies the provision of an API by the recharging or refuelling point operator, as recommended by the IR.

Regarding the second category – “Exactness” – a single criterion was identified for the “correctness” dimension. This criterion contains a metric that determines – among the data elements covered by a AF dataset – the proportion that adheres to the description provided in the Annex of the IR [3].

As for the “Comprehensiveness” group, a single criterion was considered appropriate to assess the “Completeness” dimension. The criterion evaluates how the dataset covers the required data elements according to AFIR. It is comprised of seven metrics, which separately address each data category exposed in the IR.

Two criteria were designed for the “Promptness” of the AF dataset. The IR, in Article 2, provides specifications regarding the update frequency for static and dynamic data. There are specific timeframes mentioned in such an article, which we used to define the metrics belonging to the “update frequency” criterion as part of the “currency” dimension.

With regard to “Coherence”, we included a metric to verify if the dataset contains a standardised format, which fits into the “Consistency” dimension. Regarding such format, usage of the DATEX II data standard is mandated by the IR in Article1 for AFIR data deliveries from April 14, 2026.

Lastly, with respect to the “Suitability” group, we added a single metric criterion evaluating if there is data duplication within the dataset, which refers to the “relevance” dimension. This criterion was based on Article 3 of the IR, which outlines – under the consistency-related requirement – that potential overlapping or duplicate data shall be removed. According to the definitions from Table 2, the uniqueness or duplication of data fits into the “Relevance” dimension rather than “consistency” (label adopted by the IR), which refers, in the data quality domain, mostly to the consistent use of specific data formats. It is noted that uniqueness can be fostered by using unique identifiers for the recharging infrastructure, which are defined as mandatory data elements by the IR.



Table 3 - Description of quality criteria and metrics

Quality dimension	Quality criterion	Criterion description	Metrics	Metric description	Mathematical/theoretical formulation	Possible values
Completeness	AFIR data coverage	Assess the proportion of AF-related data types that are covered by the dataset, according to the AFIR. Note: By “coverage”, it is meant that the information exists in the dataset, even if the information does not adhere to reality or to the description provided in the Annex of the IR on data.	Percentage of Static data for recharging and refuelling points - % Static (general)	Measures the percentage of the 24 data types belonging to the data category (a) of the IR [3] that are covered by the dataset.	$\% \text{ Static} = \frac{\text{count of data types covered}}{24} \times 100\%$	0 to 100%
			Percentage of further static data for recharging points - % Static (Recharging)	Measures the percentage of the 10 data types belonging to the data category (b) of the IR [3] that are covered by the dataset.	$\% \text{ Static (recharging)} = \frac{\text{count of data types covered}}{10} \times 100\%$	N/A: for hydrogen and liquified methane refuelling infrastructure If applicable: 0 to 100%
			Percentage of further static data for hydrogen refuelling infrastructure - % Static (Hydrogen)	Measures the percentage of the data types belonging to the data category (c) of the IR [3] that are covered by the dataset.	$\% \text{ Static (hydrogen)} = \frac{\text{count of data types covered}}{4} \times 100\%$	N/A: for liquified methane refuelling infrastructure and electric recharging points If applicable: 0 to 100%
			Percentage of further static data for refuelling infrastructure for liquified methane - % Static (Methane)	Measures the percentage of the data types belonging to the data category (d) of the IR [3] that are covered by the dataset.	$\% \text{ Static (methane)} = \frac{\text{count of data types covered}}{1} \times 100\%$	N/A: for hydrogen refuelling infrastructure and electric recharging points If applicable: 0 to 100%
			Percentage of further static data for refuelling infrastructure for alternative fuels - % Static (refuelling)	Measures the percentage of the data types belonging to the data category (e) of the IR [3] that are covered by the dataset.	$\% \text{ Static (refuelling)} = \frac{\text{count of data types covered}}{1} \times 100\%$	N/A: for electric recharging points (only)
			Percentage of dynamic data for refuelling and recharging points - % Dynamic	Measures the percentage of the data types belonging to the data category (f) of the Delegated and Implementing Acts (see Fehler! Verweisquelle konnte nicht gefunden werden. of this document for reference) that are covered by the dataset.	$\% \text{ Dynamic (general)} = \frac{\text{count of data types covered}}{3} \times 100\%$	0 to 100%
			Percentage of further dynamic data for hydrogen recharging infrastructure - % Dynamic (Hydrogen)	Measures the percentage of the data types belonging to the data category (g) of the IR [3] that are covered by the dataset.	$\% \text{ Dynamic (hydrogen)} = \frac{\text{count of data types covered}}{1} \times 100\%$	N/A: for hydrogen refuelling infrastructure and electric recharging points
Correctness	Alignment with the IR	Assesses how the elements in a AF dataset correctly reflect the description set out in the Annex of the IR.	Percentage of correct data types - % Correct	The ratio between the number of data types that correctly reflect the data types according to the IR on data and the total number of data types covered by the dataset (sum of the counts performed in the criterion “AFIR data coverage”)	$\% \text{ Correct} = \frac{\text{count of correct data types}}{\text{count of total data types covered}} \times 100\%$	0 to 100%
Accessibility	Level of accessibility	Assesses if the dataset is publicly accessible and under which conditions	Multiclass evaluation	Multiple choice question	How would you classify the level of data accessibility of this dataset?	0: Data accessibility is private – not available to parties outside the organisation owning the data. 1: Data accessibility is restricted – The dataset is accessible only to specific, authorized organizations or individuals,

Quality dimension	Quality criterion	Criterion description	Metrics	Metric description	Mathematical/theoretical formulation	Possible values
						potentially after an application or review process. 2: Data accessibility is public, with access conditions limiting the use of data for service development and research.
	API availability	Assesses if the operator of the recharging or refuelling point provides an API	Multiclass evaluation	Multiple choice question	Is there an application programme interface (API) available to retrieve data from the recharging or refuelling point for AF?	0: No; 1: Yes, limited/selection (of) data types available in the dataset 2: Yes, all data types accessible through API
	NAP accessibility	Assesses if the dataset is published at the National Access Point	Binary classification	Yes/No question	Is the dataset registered through the respective National Access Point?	0: No; 1: Yes
Currency	Update frequency (static data)	Assesses the frequency on which static data on Alternative Fuels are updated after a change in the real-world	Binary classification	Yes/No question	What is the update frequency for static data?	0: More than 24 hours after the real-world change occurred 1: Less than 24 hours after the real-world change occurred
	Update frequency (dynamic data)	Assesses the frequency on which dynamic data on electric recharging points are updated after a change in the real-world	Binary classification	Yes/No question	What is the update frequency for dynamic data from electric charging points?	0: More than one minute after the real-world changes occurred 1: Less than 1 minute after the real-world change occurred
Consistency	Compliance with recommended data standards	Assesses whether the Alternative Fuels dataset is compliant with the recommended data standards	Multiclass evaluation	Multiple choice question	Does the dataset comply with any standardized data standard (model)?	0: No, it uses a proprietary (non-standardized) format 1: Yes, it uses a commonly used standard other than DATEXII (e.g., OCPI) 2: Yes, it uses DATEX II (Alternative Fuels data model)
Relevance	Data duplication	Assesses the occurrence of data duplication within the dataset	Binary classification	Yes/No question	Have any data duplication been detected throughout the dataset?	0: No; 1: Yes



4. Proposal for Quality Assessment Methods

4.1. Generic approach

In a real-life Quality Assessment for Alternative Fuels data, we assume that a dedicated person or organisation (maybe a quality manager at the data-issuing organisation) is conducting a pre-defined procedure to calculate and compile the quality criteria and metrics, as introduced in Section 3, and to document the calculation results.

Such assessment may be done by assignment (i.e., in a limited manner) or on a regular basis (e.g., every 6 months). The assessment object is supposed to be a specific Alternative Fuels dataset, which consists of one or several data records for a specific recharging or refuelling infrastructure. The records may contain static and/or dynamic data.

Some quality criteria require a reference basis, often referred to as ground truth, for accurate calculation. In the context of AF data, one approach could be to conduct site visits to verify the alignment between the dataset and the real-world conditions it represents.

Depending on the above-mentioned prerequisites, the choice and concrete usage of the quality criteria and metrics might be different in each quality assessment. For reasons of understandability and transparency, we propose that each documentation of a Quality Assessment includes the following meta information:

- Author of Quality Assessment
- Date of Quality Assessment
- What AF dataset has been assessed? (Which currency? Which spatial coverage?)
- What reference basis has been used, if any?
- What quality criteria and metrics from the AF QF have been applied?
- Any remarks or additional notes about the calculation of the quality criteria and metrics?

The documentation of the Quality Assessment, e.g., tables with calculation results, including the above meta-information, should be available to potential data consumers. A preferred way to do so is to publish the documentation as metadata on the NAP. The NAPCORE Metadata specification² [6] has introduced the dedicated metadata property “Quality Information” to facilitate such documentation.

4.2. Aggregated view on quality criteria and metrics

A special task during the Quality Assessment might be an aggregation of the calculation results for individual criteria and metrics. The benefit of such aggregation is a better view on the entire quality of an AF dataset, and better comparability when assessing multiple AF datasets.

As a first step, we propose an approach that allows us to obtain aggregated quality indicators for each quality dimension. Given the input values for each metric – which are assigned by the evaluator – we use the equation (1) to calculate an aggregated quality indicator on the level of one quality dimension. It is a scaled weighted mean of the relative contributions of each individual metric to the quality indicator (these contributions are indicated by the fraction on the right side of the numerator).

$$Q_i = \frac{\sum_j \left(W_{i,j} \times A_{i,j} \times \frac{\text{Input value}_{i,j} - \text{Worst}_{i,j}}{\text{Best}_{i,j} - \text{Worst}_{i,j}} \right)}{\sum_j (W_{i,j} \times A_{i,j})} \quad (1)$$

Where:

- Q_i – Quality indicator for the dimension i

² <https://napcore.eu/metadata/>

- "*i*" – quality dimension identification
- "*j*" – the identification of the metric in each quality dimension
- $W_{i,j}$ – parameter weight for the "*j*"-th metric across the "*i*"-th dimension
- $A_{i,j}$ – applicability for the assessed dataset – binary value ... "0" for non-relevant (N/A), "1" for relevant
- $Worst_{i,j}$ – worst possible value considered for the computation of the respective metric
- $Best_{i,j}$ – best possible value for the computation of the respective metric
- $Input\ value_{i,j}$ – the value assigned by the evaluator for the respective metric. It is worth mentioning that if the assigned value is outside of the interval $\langle Worst_{i,j}; Best_{i,j} \rangle$ (which may happen for quantitative metrics), the nearest value from that interval is used instead

It is important to note that the worst and best values considered for the computation might not be the same as the possible values outlined in Table 3 (last column). For some quantitative variables, the worst or best value would tend to infinity, making it impossible to calculate the quality indicator. Evaluators are free to edit/adjust the weights of each metric according to their preference. Once again, we reinforce the importance of allowing for greater flexibility during the assessment process.

As a second step, after obtaining a quality indicator for each quality dimension, equation (2) details the process to obtain the general quality indicator for the dataset.

$$Q_{AF} = \frac{\sum_i (W_i \times Q_i)}{\sum_i (W_i)} \quad (2)$$

Where:

- Q_{AF} – overall quality indicator of the AF dataset
- "*i*" – quality dimension identification
- W_i – parameter weight for the "*i*"-th dimension
- Q_i – quality indicator of the "*i*"-th dimension

The weight of each dimension is calculated as the average of the weights of each applicable metric belonging to that dimension.

5. Conclusions

This report details the development of a QF for Alternative Fuels data. We formulated applicable criteria based on a comprehensive literature review and the outcomes of legislation efforts, as well as insights from other European projects and initiatives. Specific metrics for each criterion were designed to enable evaluators to obtain measurable values. We also proposed a method to obtain these metrics for a particular dataset and a procedure to aggregate values for each quality dimension.

This approach emphasises flexibility, acknowledging that overall quality dimensions are not universally applicable and must be assessed for each use case. Different applications and domains may prioritize certain dimensions over others based on specific requirements and operational environments. Consequently, some quality dimensions were not considered in the final set as they were deemed out of scope or difficult to measure. However, the QF comprises six quality dimensions, each including one or more criteria and their respective metrics. These dimensions are:

- Completeness
- Correctness
- Accessibility
- Currency
- Consistency
- Relevance

A key challenge in applying the proposed data quality framework lies in the difficulty of establishing a reliable ground truth. Generating such a reference, for instance, through on-site visits, can be highly resource-intensive, especially considering that the intended criteria are to be assessed at the level of individual charging or fuelling points. Given the scale and granularity required, conducting site visits for each location would be challenging. This limitation may restrict the practical applicability of the framework and should be acknowledged. However, the flexibility of the method—allowing the selection and weighting of criteria based on data availability and user relevance—offers a pragmatic approach to mitigate this issue. While not all ground truth data may be attainable, the framework is designed to adapt accordingly, thus balancing theoretical rigour with practical constraints.

It is worth mentioning that the initial goal was to develop a case study using a real-world Alternative Fuels dataset, either from a recharging or refuelling point. However, this was not achieved during the timeframe of the NAPCORE project due to two main reasons:

- Time/resource constraints: the legislative developments of the EC (namely, the IR), crucial for the completion of the quality criteria, were only acknowledged in late 2024. As a result, the consolidated criteria were only finalised in Q1 2025, when the project was already approaching its end.
- Challenges in accessing real-world datasets: Similar to the difficulties faced during the development of the Alternative Fuels demonstrator [7], obtaining access to a real-world Alternative Fuels dataset – especially through the NAPs – proved challenging. Although we managed to collect a few datasets, they used non-standardised formats, which hindered proper interpretation and reuse for quality assessment purposes. This scenario is expected to be improved since the AFIR mandates operators and owners of AF infrastructure to ensure the availability of data in the prescribed data format (DATEX II).

It is important to highlight that legislative developments have significantly improved since the beginning of this work. With the publishing of AFIR – especially article 20 – and the inclusion of additional data types and recommendations from the AFIR IR on data, there is now a stronger foundation for setting up a proper Alternative Fuels dataset and ensuring compliance with the criteria outlined in this QF.



While this QF is still in its early stages and has not yet been tested in a real-world environment, it represents a concrete set of criteria serving as a baseline for establishing common definitions of what constitutes good data quality in Alternative Fuels contexts. It also provides a defined methodology for measuring and evaluating the quality of a dataset, as well as comparing the quality of different datasets. As the landscape of Alternative Fuels data provision and availability is expected to evolve significantly in the coming years, further enhancements will be necessary to update both the criteria and the associated metrics.



References

- [1] European Union, “REGULATION (EU) 2023/1804 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 13 September 2023 on the deployment of alternative fuels infrastructure, and repealing Directive 2014/94/EU,” Sep. 2023. Accessed: Jul. 26, 2024. [Online]. Available: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32023R1804>
- [2] European Commission, “COMMISSION DELEGATED REGULATION (EU) 2022/670 of 2 February 2022 supplementing Directive 2010/40/EU of the European Parliament and of the Council with regard to the provision of EU-wide real-time traffic information services,” Apr. 2022. Accessed: Jul. 26, 2024. [Online]. Available: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32022R0670>
- [3] European Commission, “COMMISSION IMPLEMENTING REGULATION (EU) 2025/655 of 2 April 2025 laying down rules for the application of Regulation (EU) 2023/1804 of the European Parliament and of the Council as regards specifications and procedures relating to the availability and accessibility of data on alternative fuels infrastructure,” May 2025. Accessed: June 03, 2025. [Online]. Available: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=OJ:L_202500655
- [4] European Commission, “STF Sub-group on Data,” Sub-group on a common data approach for electromobility and other alternative fuels (RegEx E03321/7). Accessed: Jan. 03, 2025. [Online]. Available: https://transport.ec.europa.eu/transport-themes/clean-transport/sustainable-transport-forum-stf/active-sub-groups/sub-group-common-data-approach-electromobility-and-other-alternative-fuels-regex-e033217_en
- [5] A. M. Carvalho, S. Soares, J. Montenegro, and L. Conceição, “Data Quality: revisiting dimensions towards new framework development,” in: *Procedia Computer Science* Volume 253, 2025, Pages 247-256. Available: <https://doi.org/10.1016/j.procs.2025.01.088>
- [6] NAPCORE, “Metadata,” mobilityDCAT-AP. Accessed: Jul. 11, 2024. [Online]. Available: <https://napcore.eu/metadata/>
- [7] NAPCORE, “Annex C – Alternative Fuels Demonstrator: Interoperability demonstrators with real-world use of data and different Levels of Service – 2nd round (Milestone 2.10, Version 1.3, Final Draft),” 2025.

