



ASSESSMENT SUMMARY v1.0.0

Data Quality Vocabulary (DQV)¹

World Wide Web Consortium (W3C)²

¹ DQV specification: [Data on the Web Best Practices: Data Quality Vocabulary \(w3.org\)](https://www.w3.org/Data/BestPractices/DataQualityVocabulary/)

² W3C: <https://www.w3.org/>

Change Control

Modification		Details	
Version 1.0.0			
Initial version			

TABLE OF CONTENT

1. INTRODUCTION..... 4

2. ASSESSMENT SUMMARY 4

2.1. EIF Interoperability Principles.....4

2.2. EIF Interoperability Layers7

3. ASSESSMENT RESULTS 9

1. INTRODUCTION

The present document is a summary of the assessment of DQV carried out by the CAMSS Team using the CAMSS EIF assessment scenario³. The purpose of this scenario is assessing the compliance of a standard or specification with the European Interoperability Framework (EIF).⁴

2. ASSESSMENT SUMMARY

Data Quality Vocabulary (DQV) provides a framework in which the quality of a dataset can be described, whether by the dataset publisher or by a broader community of users. It does not provide a formal, complete definition of quality, rather, it sets out a consistent means by which information can be provided such that a potential user of a dataset can make his/her own judgment about its fitness for purpose. Moreover, DQV can be a good asset for semantic interoperability as it provides a standardized metadata framework, as well as facilitates semantic data integration. Given that DQV is intended to describe the quality of datasets, organizations make it easier for others to discover and assess the suitability of their data for various purposes, such as data discover and reuse and the automated processing of data.

2.1. EIF Interoperability Principles

Interoperability principles are fundamental behavioural aspects that drive interoperability actions. They are relevant to the process of establishing interoperable European public services. They describe the context in which European public services are designed and implemented.

The specification fully supports the principles setting context for EU actions on interoperability:

- **Subsidiarity and proportionality**

Data Quality Vocabulary (DQV) is not included in any national catalogue of recommended specifications whose Member State NIF has a high performance on interoperability according to the National Interoperability Framework Observatory (NIFO)⁵ factsheets.

The specification fully supports the principles setting context for EU actions on interoperability:

- **Openness**

DQV is a vocabulary for expressing data quality and according to the W3C Data on the Web Best Practices Working Group, it is foreseen as an extension to the DCAT vocabulary [vocab-dcat] to cover the quality of the data, how frequently is it updated, whether it accepts user corrections and persistence commitments. The W3C has a defined and publicly available Process for the

³ CAMSS Assessment EIF scenario v.6.0.0: <https://ec.europa.eu/eusurvey/runner/CAMSSAssessmentEIFScenario6>

⁴ ISA2 Programme: https://ec.europa.eu/isa2/eif_en

⁵ NIFO Factsheets: <https://joinup.ec.europa.eu/collection/nifo-national-interoperability-framework-observatory/digital-publicadministration-factsheets-2023>

Development and approval process of the specification as a recommended standard, including a public review. W3C has also Royalty-Free IPR licenses granted under the W3C Patent Policy apply to all W3C specifications.

Since 2015, three major Data Quality Vocabulary releases have been published and several minor changes have been recommended, as well as W3C has a page that collects implementations of the Data Quality Vocabulary, for instance, DQV has a wide market acceptance. Furthermore, the Web Best Practices Community Group continues the work started by the Data on the Web Best Practices Working Group in investigating topics such as data versioning, subsetting, data access and metadata.

- **Transparency**

DQV plays a crucial role in promoting transparency, interoperability, and quality assessment by providing an standardised metadata about data quality, which can indirectly contribute to their visibility and usability in the context of data publishing and sharing. In addition, it can be integrated with service descriptions or metadata schemas to provide additional context about the quality of services.

On the other hand, DQV can enhance transparency and accountability by making the quality of data used in service interfaces more visible to stakeholders, including service users, regulators, and auditors. This transparency fosters trust in public administration services and encourages responsible data management practices.

- **Reusability**

DQV focuses on describing fundamental data quality concepts such as accuracy, completeness, consistency, and timeliness. These concepts are applicable across different domains, including business, government, healthcare, education, and more. By providing a common vocabulary for expressing these concepts, DQV facilitates interoperability and consistency in data quality assessment across various domains.

- **Technological neutrality and data portability**

DQV does not prescribe a specific data representation or serialisation format, making it technology agnostic in this aspect. DQV can also be implemented and utilized across different platforms and environments without being tied to any particular data format or technology stack, which makes it platform agnostic.

The W3C Data Quality Vocabulary (DQV) specification is designed to be modular and flexible, allowing for partial implementations based on specific use cases and requirements. Furthermore, users can define their own quality dimensions, metrics, and indicators, or extend existing ones to accommodate domain-specific requirements, as well as create domain-specific extensions or profiles of DQV that focus on particular domains or industries.

On the other hand, DQV addresses and enables data portability between systems/applications supporting the implementation or evolution of European public services by providing them with the means to describe and manage data quality, which can support data portability initiatives within the European public services ecosystem.

The Technical Specification partially supports the principles related to generic user needs and expectations:

- **User-centricity**

DQV provides a standardised vocabulary for describing data quality information, including dimensions, metrics, indicators, and constraints. By adopting DQV, organisations can ensure consistency and interoperability in how data quality is represented and communicated. This standardised approach enables the reuse of quality-related information across different datasets, systems, and applications.

- **Inclusion and accessibility**

Accessibility is a quality dimension in DQV as the degree to which data can be accessed in a specific context of use, particularly by people who need supporting technology or special configuration because of some disability. Besides, accessibility is mentioned across the specification for different examples of the vocabulary.

- **Privacy**

By implementing DQV, public administrations can assess the quality of the data they manage, including personal data. DQV enables public administrations to identify and address data quality issues that could compromise the security or privacy of personal data. Confidentiality is also mentioned as a security matter in DQV, and the specification introduces certain aspects that can contribute to enabling it.

Moreover, DQV has not been found included in any national or European initiative covering privacy aspects.

- **Security**

The specification addresses many security aspects, but without specific provisions that allow it, such as confidentiality or data integrity, as well as, it is possible to explore the quality of data in security aspects like data accuracy and data integrity. By using DQV to describe and communicate the integrity measures applied to a dataset, organizations can help ensure that the data remains secure, as well as can help to define and document specific scopes of security. On the other hand, there are aspects that we cannot explore about their quality of data, such as the secure exchange and processing of data.

- **Multilingualism**

DQV enables the representation of information about the quality of a set using the “Multilingual importing” quality metric. This metric applies to linksets between datasets that include SKOS concepts. It considers the concepts that have been matched and quantifies the information gain the links provide if one adds the preferred labels or the alternative labels of the concepts from the object dataset to the descriptions of the concepts from the subject dataset.

The Technical Specification partially supports the foundation principles for cooperation among public administrations:

- **Administrative Simplification**

By providing information about quality data sets, DQV can help avoid administrative burden. Although it is not a unique quality version, it can offer a first impression of it, and it can be useful to obtaining information about the quality of datasets. Moreover, the possibility of publish and exchange quality metadata, encourages the implementation of digital services, supporting the principle of digital-first.

- **Preservation of information**

The purpose of DQV is not related to long term preservation of electronic records. Therefore this criterion is considered not applicable to this specification.

- **Assessment of effectiveness and efficiency**

There are already existing studies and documents assessing and documenting DQV features and providing possible improvements of its performance among other aspects. In "DQAF: Towards DQV-Based Dataset Quality Annotation Using the Web Annotation Data Model"⁶ it is proposed to use the DQV specification for a Dataset Quality Annotation Framework (DQAF). Furthermore, there are documentation such as "Introducing the Data Quality Vocabulary (DQV)"⁷, which makes and assessment of DQV and explains its components as well as how it works.

2.2. EIF Interoperability Layers

The interoperability model which is applicable to all digital public services includes: - Four layers of interoperability: legal, organisational, semantic and technical; - A cross-cutting component of the four layers, 'integrated public service governance'; - A background layer, 'interoperability governance'.

The Technical Specification partially supports the implementation of digital public services complying with the EIF interoperability model:

- **Interoperability governance**

DQV is already associated with EIRA ABBs in the European Library Of Specifications (ELIS). More specifically, DQV can define the interoperability aspects of the "Linked Data", "Linked Open Data", "Open Data", "Controlled Vocabulary", "Data", "Data Format", "Data Model", "Hash Code", "Metadata", and "Ontology" ABBs of the EIRA Technical Infrastructure View. Moreover, the specification is not recommended by any European Member State, but is included in open repository of standards at national⁸ and European level⁹. However, DQV has been used in the ALIGNED project¹⁰.

⁶ DQAF: Towards DQV-Based Dataset Quality Annotation Using the Web Annotation Data Model: <https://ieeexplore.ieee.org/document/7799649>

⁷ Introducing the Data Quality Vocabulary (DQV): <https://www.semantic-web-journal.net/system/files/swj2079.pdf>

⁸ Norwegian Digitalisation Agency: <https://eurogeographics.org/app/uploads/2019/06/6-SDQ2020-Data-quality-in-an-e-Government-perspective.pdf>

⁹ Metadata Quality Assessment Methodology: <https://data.europa.eu/mqa/methodology?locale=en>

¹⁰ ALIGNED Project: <https://cordis.europa.eu/project/id/644055/reporting>

- **Legal interoperability**

The specification is not considered a European Standard

- **Organisational interoperability**

DQV is useful for describing the quality of the different categories of a dataset. It can help identify where a dataset has more or less quality. One of the principles which is considered by the W3C working group is that DQV should be easy to map to as well as to re-use and extend.

- **Semantic interoperability**

Joinup offers several services that aim to help e-Government professionals share their experience with each other. Joinup supports them to find, choose, re-use, develop and implement interoperability solutions. DQV is mentioned in some Joinup entries as a discussion topic, although is not distributed in the platform.

3. ASSESSMENT RESULTS

This section presents an overview of the results of the CAMSS assessments for DQV. The CAMSS “Strength” indicator measures the reliability of the assessment by calculating the number of answered (applicable) criteria. On the other hand, the number of favourable answers and the number of unfavourable ones are used to calculate the “Automated Score” per category and an “Overall Score”.

Category	Automated Score	Assessment Strength	Compliance Level
EIF Principle setting the context for EU actions on interoperability	20/100 (20%)	100%	Ad-hoc
Core interoperability principles	1640/1700 (96%)	100%	Seamless
Principles related to generic user needs and expectations	1080/1200 (85%)	67%	Seamless
Foundation principles for cooperation among public administrations	500/500 (100%)	80%	Seamless
Interoperability layers*	820/1000 (82%)	100%	Seamless
Overall Score	3560/4000 (89%)	89%	

**The technical interoperability layer is covered by the criteria corresponding to the core interoperability principle "Openness".*

With 89% of assessment strength, this assessment can be considered representative of the high specification compliance with the EIF principles and recommendations. The Overall Automated Score of 89% (3560/4000) demonstrates that DQV highly supports the European Interoperability Framework in the domains where it applies.