

**M5.1 Draft guideline on data description – public consultation questions**

**Guideline for health data holders on their duties regarding data description**

TEHDAS2 – Second Joint Action Towards the European Health Data Space

19 December 2024

**This project has been co-funded by the 4th EU Health Programme (2021–2027) under Grant Agreement no 101176773.**

**Disclaimer**

Views and opinions expressed in this deliverable represent those of the author(s) only and do not necessarily reflect those of the European Union or HaDEA. Neither the European Union nor the granting authority can be held responsible for them.

**Copyright Notice**

Copyright © 2024 TEHDAS2 Consortium Partners. All rights reserved. For more information on the project, please see [www.tehdas.eu](http://www.tehdas.eu).

**Contents**

[1 Part A questions for generic feedback 2](#_Toc188283017)

[1.1 Demography 2](#_Toc188283018)

[1.2 Quality 2](#_Toc188283019)

[1.3 Generic feedback 3](#_Toc188283020)

[2 Part B questions for specific feedback 3](#_Toc188283021)

[2.1 Part 1: Role and experience 3](#_Toc188283022)

[2.2 Part 2: Properties of the proposed solution HealthDCAT-AP 4](#_Toc188283023)

[2.3 Part 3: Usability of the proposed solution HealthDCAT-AP 7](#_Toc188283024)

# Part A questions for generic feedback

These questions will be asked in each public consultation to provide an understanding of the recipients’ demographics, the quality of the document and to gather generic feedback. Questions marked with an asterisk are mandatory.

## Demography

**Country**\* [-List of countries-, European Union, Outside European Union]

**Type of the responder**\* [Public organisation, Private organisation, Non-governmental organisation (NGO), Academic or research institution, Industry association, Individual expert or professional, Patient representative, Citizen, Other (please specify)]

**Sector**\* [Health care provider, Government/public administration, Research and development, Medical devices, Pharmaceutical industry, Education and academia, Information technology, Data Management/processing, Patient advocacy, Legal and compliance, Other (please specify)]

**Organisation size**\* [Micro (1–9 employees), Small to medium enterprise (10–249 employees), Large enterprise (250+ employees), Not applicable (for individual respondents)]

**Professional role/function** [open text field]

## Quality

**From your perspective, how ready is the document to meet the expected needs?**\* [Early draft, Advanced draft, Well-developed, Needs finalisation only]

**What is level of quality of the document?**\* [Low, Medium, High]

**Is the document clear and easy to understand?**\* [Rate 1 (not clear nor easy to understand) – 4 (very clear and easy to understand)]

**How well does the document address the key issues and challenges related to its subject matter?**\* [Rate 1 (not well) – 4 (very well)]

**How feasible and implementable do you find the recommendations or technical specifications presented in the document?**\* [Rate 1 (not feasible and implementable at all) – 4 (very feasible and implementable)]

## Generic feedback

**Do you have any suggestions for improving the document? Are there any additional topics or areas that should be covered? [Please provide feedback and ideas for enhancing the document]** [max. 750 characters]

# Part B questions for specific feedback

These questions are specific for TEHDAS2 WP5 Task 5.1, for Milestone 5.1 on Data Description Duties for Health Data Holders.

## Part 1: Role and experience

**1.1 Are you a health data holder, according to the definition of the EHDS?\*** [Yes, No]

*Natural or legal person, public authority, agency or other body in the healthcare or the care sectors, including reimbursement services where necessary, as well as any natural or legal person developing products or services intended for the health, healthcare or care sectors, developing or manufacturing wellness applications, performing research in relation to the healthcare or care sectors or acting as a mortality registry, as well as any Union institution, body, office or agency, that has either:*

1. *the right or obligation, in accordance with applicable Union or national law and in its capacity as a controller or joint controller, to process personal electronic health data for the provision of healthcare or care or for the purposes of public health, reimbursement, research, innovation, policy making, official statistics or patient safety or for regulatory purposes; or*
2. *the ability to make available non-personal electronic health data through the control of the technical design of a product and related services, including by registering, providing, restricting access to or exchanging such data.*

*Please note that the following categories of health data holders shall be exempt from the obligations on health data holders: (a) natural persons, including individual researchers; (b) legal persons that qualify as microenterprises.*

* 1. **If yes, do you already create metadata for the health data you hold?** [Yes, No]
	2. **If yes, which metadata standard are you using?** [text field]
	3. **If no, why not?** [text field]
	4. **Do you already share or expose your metadata?** [Yes, No]
	5. **If yes, how?** [text field]

## Part 2: Properties of the proposed solution HealthDCAT-AP

**2.1 In your opinion, is the HealthDCAT-AP sufficient for health data holders to fully and accurately describe their datasets?\*** [Rate 1 (No not sufficient) – 4 (Yes very sufficient)]

**2.2 In what ways could the HealthDCAT-AP be tailored to better support the unique characteristics of (your) health datasets?\* [text field]?\*** [text field]

**2.3 What aspects of (your) health datasets are most challenging to describe using HealthDCAT-AP?\*** [text field]

**2.4 How would you compare the HealthDCAT-AP to other metadata standards you have used/are using? What aspects stand out as strengths or weaknesses?** [text field]

**2.5 Feedback on the current cardinalities to describe personal electronic health datasets through HealthDCAT-AP** [matrix, scale selection] *Note.* M = mandatory | R = Recommended | O = Optional

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Property** | **Definition** | **Current proposal** | **Should be Mandatory** | **Should be Recommended** | **Should be Optional** |
| dcat:keyword | Keywords or tags describing the dataset | M |  |  |  |
| dct:provenance | Information about how the data was collected, including methodologies, tools, and protocols used | M |  |  |  |
| dpv:hasPurpose | A free text statement of the purpose of the processing of data or personal data. | M |  |  |  |
| dct:issued | The date of formal issuance (e.g.: publication) of the dataset | O |  |  |  |
| dct:modified | The most recent date on which the dataset was changed or modified. | O |  |  |  |
| healthdcatap:populationcoverage | A definition of the population within the dataset | R |  |  |  |
| dct:spatial | The geographic area covered by the dataset. | M |  |  |  |
| dct:temporal | A temporal period that the Dataset Series covers. | R |  |  |  |
| dct:language | The language(s) of the dataset content. | R |  |  |  |
| healthdcatap:retentionPeriod | A temporal period which the dataset is available for secondary use. | O |  |  |  |
| dct:accrualPeriodicity | Frequency of updates to the dataset (e.g., daily, monthly, annually). | R |  |  |  |
| dcat:spatialResolutionInMeters | The minimum spatial separation resolvable in a dataset, measured in meters. | O |  |  |  |
| dcat:temporalResolution  | The minimum time period resolvable in the dataset. | R |  |  |  |
| dct:publisher | The organisation or entity responsible for publishing the dataset. | M |  |  |  |
| healthdcatap:publishertype | A type of organisation that makes the Dataset available. | M |  |  |  |
| healthdcatap:publishernote | A description of the publisher activities | O |  |  |  |
| healthdcatap:hdab | Health Data Access Body supporting access to data in the Member State. | M |  |  |  |
| dct:creator | An entity responsible for producing the dataset. | O |  |  |  |
| dcat:contactPoint | Contact information for inquiries about the dataset. | M |  |  |  |
| foaf:page | A page or document about this Dataset. | R |  |  |  |
| dcat:landingPage | A web page that provides access to the Dataset, its Distributions and/or additional information. | R |  |  |  |
| dcat:qualifiedRelation | A description of a relationship with another resource. | O  |  |  |  |
| prov:qualifiedAttribution  | An Agent having some form of responsibility for the resource. | O |  |  |  |
| prov:wasGeneratedBy | An activity that generated, or provides the business context for, the creation of the dataset. | O |  |  |  |
| dqv:hasQualityAnnotation  | Dataset, including rating, quality certificate, feedback that can be associated to the dataset. | R |  |  |  |
| dcat:hasVersion | A related Dataset that is a version, edition, or adaptation of the described Dataset. | O |  |  |  |
| dcat:version | The version indicator (name or identifier) of a resource. | O |  |  |  |
| adms:versionnotes | A description of the differences between this version and a previous version of the Dataset. | O |  |  |  |
| dcat:isVersionOf | inverse of dcat:hasVersion | O |  |  |  |
| dct:relation | A related resource | R |  |  |  |
| dct:isReferencedBy  | A related resource, such as a publication, that references, cites, or otherwise points to the dataset. | R |  |  |  |
| dct:source | A related Dataset from which the described Dataset is derived. | R |  |  |  |
| dpv:hasLegalBasis | The legal basis used to justify processing of personal data | R |  |  |  |
| adms:identifier | A secondary identifier of the Dataset, such as MAST/ADS17, DataCite18, DOI19, EZID20 or W3ID21 | O |  |  |  |
| dcat:theme | The theme or category of the dataset: Health (M by default, with the option to add other themes) | M |  |  |  |
| dcatap:applicableLegislation | The legislation that mandates the creation or management of the Dataset. | M |  |  |  |
| healthdcatap:healthCategory | The health category to which this dataset belongs as described in the EHDS Article 51 | M |  |  |  |
| healthdcatap:healthTheme | A category of the Dataset or tag describing the Dataset. | M |  |  |  |
| dct:type | A type of the Dataset: for health datasets containing personal level information, the type of the dataset MUST take the value "personal data". | M |  |  |  |
| dpv:hasPersonalData | Key elements that represent an individual in the dataset. Example: Age Exact, Blood type, Current Employment, etc. | R |  |  |  |
| dct:conformsTo | An implementing rule or other specification. | R |  |  |  |
| healthdcatap:hasCodingSystem | Coding systems in use (ex: ICD-10-CM, DGRs, SNOMED=CT, ...) | R |  |  |  |
| healthdcatap:minTypicalAgehealthdcatap:maxTypicalAge | Minimum/maximum typical age of the population within the dataset | R |  |  |  |
| healthdcatap:numberOfRecords | Size of the dataset in terms of the number of records. | R |  |  |  |
| healthdcatap:numberOfUniqueIndividuals  | Number of records for unique individuals. | R |  |  |  |
| dcat:distribution | An available distribution for the dataset. For personal electronic health data, a distribution must include the landing page of the Health Data Access Body supporting data access. | M |  |  |  |
| adms:sample | A sample distribution of the dataset. | M |  |  |  |
| healthdcatap:analytics | An analytics distribution of the dataset. | R  |  |  |  |
| dct:identifier | The main identifier for the Dataset, e.g. the URI or other unique identifier in the context of the Catalogue. | M |  |  |  |

## Part 3: Usability of the proposed solution HealthDCAT-AP

**3.1 How adequate do you find the HealthDCAT-AP standard for describing (your) health datasets?\*** [Rate 1 (Not adequate) – 4 (Very adequate)]

**3.2 Are there specific challenges you will face in the implementation of HealthDCAT-AP?** [text field]

**3.3. How do you foresee using the HealthDCAT-AP standard to improve data discoverability and interoperability?\*** [text field]

**3.4 On a scale of 1 to 4, how clear did you find the guideline on data description?\*** [Rate 1 (not clear) – 4 (very clear)]

**3.5 What improvements would this guideline need in order to be clear and useful?\*** [text field]

**3.5 What kind of additional support, such as training, documentation, or tailored tools, would help you better adopt the standard?** [text field]