



Brussels, XXX  
[...] (2026) XXX draft

ANNEX

**ANNEX**

**to the**

**COMMISSION IMPLEMENTING REGULATION (EU) .../...**

**laying down the minimum metadata elements and their characteristics to be provided by health data holders for dataset descriptions for the secondary use of electronic health data under Article 77(1) of Regulation (EU) 2025/327 establishing the European Health Data Space**

## ANNEX

### *Part A - Technical reference*

The metadata framework referred to in Article 1 is based on the HealthDCAT-AP, which is maintained and made publicly available by the Commission<sup>1</sup>.

### *Part C - Metadata elements*

The elements listed hereafter shall be expressed using the technical specifications of the HealthDCAT-AP as referenced in Part D of this Annex. Where applicable, controlled vocabularies defined in HealthDCAT-AP must be used.

| <b>Property</b>        | <b>Description</b>   |
|------------------------|--|
| Access rights          | Describes the level of access applicable to the dataset (e.g. public, restricted or non-public).   |
| Applicable legislation | Identifies Union legal acts or national legislation that specifically govern the establishment, maintenance or availability of the dataset, where such legal acts or legislation exists. |
| Code values            | Identifies classification systems or controlled vocabularies reflected in the content of the dataset (e.g. disease classifications, procedure classifications), where applicable.        |
| Coding system          | Identifies standardised coding schemes used to encode data elements within the dataset (e.g. ICD-10-CM, SNOMED-CT, DRGs), where applicable.  |
| Contact point          | Provides contact details for enquiries related to the dataset.   |
| Conforms to            | Identifies a standard or technical specification that defines the structural or syntactic organisation of the dataset as a whole.  |
| Custodian              | Identifies the organisation responsible for holding, managing and maintaining the dataset over time (health data holder).  |
| Description            | Provides a free-text summary describing the content and purpose of the dataset.  |
| Distribution           | Provides a technical representation or format in which the dataset is made available for secondary use.  |
| Geographical coverage  | Indicates the geographic area to which the dataset relates (e.g. Member State, region or EU-wide), where applicable.   |
| Has personal data      | Indicates whether the dataset contains personal data within the meaning of Article 4, point (1), of Regulation (EU) 2016/679 of the European Parliament                                  |

<sup>1</sup> [HealthDCAT-AP Release 6](#)

|                              |   |
|------------------------------|---|
|                              | and of the Council <sup>2</sup> .   |
| Health data access body      | The competent health data access body responsible for facilitating access to the dataset under the European Health Data Space, where the dataset is subject to restricted or non-public access. |
| Health category              | Identifies the categories of electronic health data contained in the dataset and listed in Article 51(1) of Regulation (EU) 2025/327.   |
| Health theme                 | Indicates a high-level health topic or theme associated with the dataset (e.g. cardiovascular disease, cancer).   |
| Identifier                   | Provides a unique and persistent identifier for the dataset (e.g. uniform resource identifier (URI), digital object identifier (DOI) or catalogue identifier).                                  |
| Is structured                | Indicates whether the dataset is structured, meaning that it is organised according to a predefined schema or model that defines its variables, attributes or fields.                           |
| Keyword                      | Provides keywords or tags describing the dataset and supporting search and discovery.   |
| Language                     | Indicates the language or languages in which the dataset content is expressed.  |
| Number of records            | Indicates the total number of records contained in the dataset, where applicable.   |
| Number of unique individuals | Provides an estimated number of unique natural persons represented in the dataset, where the dataset contains personal data and where applicable.   |
| <b>Provenance</b>            | Describes the origin, source or creation process of the dataset.  |
| <b>Publisher</b>             | Indicates the entity making the dataset available, which may differ from the creator or custodian.  |
| <b>Temporal coverage</b>     | Indicates the time period covered by the dataset (e.g. years of data collection), where temporal information exists.  |
| <b>Title</b>                 | Provides the official name or title of the dataset.   |
| <b>Type</b>                  | Indicates the general category or nature of the dataset (e.g. registry, administrative dataset, research dataset, synthetic dataset).   |
| <b>Variables</b>             | A structured description of the variables contained in the dataset, including, where applicable, the variable name, definition, data type, coding system or value domain.                       |

<sup>2</sup> Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (OJ L 119, 4.5.2016, p. 1, ELI: <http://data.europa.eu/eli/reg/2016/679/oj>).

|                         |   |
|-------------------------|---|
| <b>Was generated by</b> | Identifies the activity or process through which the dataset was generated (e.g. registry, clinical study, administrative process). |
|-------------------------|---|

*Part C - Controlled vocabularies*

When providing dataset descriptions, health data holders shall use the controlled vocabularies listed below:

| <b>Metadata property</b> | <b>Controlled vocabulary</b>  | <b>Source</b>                             |
|--------------------------|---|---|
| Health data category     | The categories listed in Article 51(1), points (a) to (q) of Regulation (EU) 2025/327 | Article 51(1) of Regulation (EU) 2025/327 |

Where controlled vocabularies are required for the minimum metadata elements laid down in this Regulation, those defined in the HealthDCAT-AP shall be used.