

EHDS

EU4H-2021-PJ2
EHDS2 Pilot ("HealthData@EU pilot")
101079839

D6.2 - Annex 1

Functional Analysis for the integration of HealthDCAT-AP into the HealthData@EU Central Platform

Pascal Derycke
30 Septembre 2024



Co-funded by
the European Union

Disclaimer: Co-funded by the European Union. Views and opinions expressed are however 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.



Disclaimer

The content of this deliverable represents the views of the author(s) only and is his/her/their sole responsibility; it cannot be considered to reflect the views of the European Commission and/or the Consumers, Health, Agriculture and Food Executive Agency or any other body of the European Union. The European Commission and the Agency do not accept any responsibility for use of its contents.

Copyright Notice

Copyright © 2024 EHDS2 Consortium Partners. All rights reserved. For more information on the project, please see <https://www.ehds2pilot.eu>



Contents

- 1. HealthDCAT-AP 5
- 2. Functional requirements expressed by Use cases 5
 - 2.1 Use cases enhancing the Dataset webpage by incorporating additional information. 6
 - Use case 1 – Displaying label and value of “Acronym”..... 6
 - Use case 2 – Displaying label and value of “Legal basis”..... 8
 - Use case 3 – Displaying label and value of “Retention Period” 10
 - Use case 4 – Displaying label and value of “Number of records” 12
 - Use case 5 – Displaying label and value of “Number of unique individuals” 14
 - Use case 6 – Displaying the label and value of “Minimum age” 16
 - Use case 7 – Displaying label and value of “Maximum Age” 18
 - Use case 8 – Displaying label and value of “Applicable Legislation” 20
 - Use case 9 – Displaying label and value of “Health Data Access Body” 22
 - Use case 10 – Displaying label and value of “Publisher Type” 24
 - Use case 11 – Displaying label and value of “Publisher Notes” 26
 - Use case 12 – Displaying label and value of “Conforms To” 28
 - Use case 13 – Displaying label and value of “Coding systems” 30
 - Use case 14 – Displaying label and value of ‘Generated by’ 32
 - Use case 14 – Displaying label and value of “Purpose” 34
 - Use case 15 – Displaying label and value of “Provenance” 36
 - Use case 16 – Displaying label and value of “Population Coverage” 38
 - Use case 17 – Displaying label and value of “Dataset analytics” 40
 - Use case 18 – Displaying label “Sample distributions” and sub-labels “Link to sample”, “Format”, “Updated”, “Actions” with their relevant values..... 42
 - Use case 19 – Displaying label and value of “Categories” 45
 - Use case 20 – Displaying label and value of “Themes” 47
 - Use case 21 – Displaying label and value of “Health Codes” 49
 - Use case 22 – Displaying label and value of “Personal Data” 52
 - Use case 23 – Displaying label and value of “Trusted data holder” 54
 - 2.2 Use cases improving the EU datasets catalogue webpage with advanced filtering capabilities. 56
 - Use case 24 - Filter EU Datasets Catalogue by metadata values in 56
“Health Data Access Bodies” list..... 56
 - Use case 25 - Filter EU Datasets Catalogue by metadata values in Publisher list 59
 - Use case 26 – Filter EU Datasets Catalogue by metadata values in Categories list 62



Use case 27 – Filter EU Datasets Catalogue by metadata values in “Health themes” list	65
Use case 28 – Filter EU Datasets Catalogue by metadata values in “Coding systems” list	68
Use case 29 – Filter EU Datasets Catalogue by metadata values in “Data models” list	71
Use case 30 – Filter EU Datasets Catalogue by metadata values in “Age range” list	74
Use case 31 Filter EU Datasets Catalogue by metadata values in “Population Size” list	77
Use case 32 –Filter EU Datasets Catalogue by metadata values in “Retention Period” list	80
Use case 33 – Filter EU Datasets Catalogue by metadata values in “Update Frequency” list.	84
Use case 34 – Filter EU Datasets Catalogue by metadata values in “dataset type” list	87
Use case 35 – Filter EU Datasets Catalogue by metadata values in “Personal data” list	90
Use case 36 – Filter EU Datasets Catalogue by metadata value “Trusted Data Holder”	93
Use case 37 – SPARQL endpoint	96
3. Annexes:	99
Annex 1 Controlled vocabularies for the specific use of the HealthDCAT-AP	99
Annex 2: RDF HealthDCAT-AP template	102
Annex 3: Filtering capabilities of a HealthDCAT-AP catalogue	103



1. HealthDCAT-AP

The Proof of Concept (PoC) for Health DCAT-AP aims to augment the standard Data Catalogue Vocabulary Application Profile for data portals in Europe (DCAT-AP) by incorporating specialised metadata elements specific to the health sector. This enhancement aligns with the European Health Data Space (EHDS) Regulation proposal, ensuring relevance and compliance.

The cardinalities specified in the draft version of [HealthDCAT-AP](#) are currently tentative and will be further discussed in the frame of the Second Joint Action Towards the European Health Data Space – [TEHDAS2](#) ensuring that the final model is both robust and suited to the unique needs of the health sector.

The current approach to cardinalities in HealthDCAT-AP involves defining three sets of metadata elements - mandatory, recommended, and optional - based on the access rights to the dataset. These sets are tailored according to the data type: open data, protected data, and sensitive data (such as personal data).

2. Functional requirements expressed by Use cases

This section outlines a series of user stories related to the implementation of the HealthDCAT-AP for the HealthData@EU Central Platform. The user stories are organised to reflect the functional requirements of the platform through the following types of use cases:

- Data Display: Enhancing the Dataset webpage by incorporating additional information.
- Data Searchability: Improving the EU Datasets Catalogue webpage with advanced filtering capabilities.

[EHDS Regulation Whereas 60](#)

The EU datasets catalogue could be aligned with the data.europa.eu initiative and without prejudice to the requirements set out in the Regulation (EU) 2022/868. Interoperability should be ensured between the EU datasets catalogue, the national datasets catalogues and the dataset catalogues from European infrastructures and other relevant data sharing infrastructures.

The functional requirements are defined within the context of transitioning and implementing HealthDCAT-AP from the existing "[EU Data Portal](#)" system to an enhanced system with new capabilities, referred to as the "EU Health Data Portal" (i.e., EU Datasets Catalogue).



2.1 Use cases enhancing the Dataset webpage by incorporating additional information.

Use case 1 – Displaying label and value of “Acronym”

User story name	As an anonymous user of the HealthData@EU Central Platform, I want to see the acronym of each dataset displayed on its corresponding Dataset webpage, This will allow me to quickly and easily identify and reference the dataset by its metadata..
Role(s) concerned	All users
Necessary right(s)	N/A
Precondition scenario	The HealthData@EU Central Platform hosts the EU Datasets Catalogue which includes metadata records with the metadata property [dct:alternative rdfs:Literal (Open data 0..n) (Protected data 0..n) (Sensitive data 0..n)] The user can access the HealthData@EU Central Platform from his computer or mobile (an Internet connection is required).
Scenario definition	1) The user visits the Dataset webpage of the HealthData@EU Central Platform. 2) If metadata are available, then: The label “Acronym” and the value of acronym of the dataset is displayed
Alternative scenarios	2) If metadata are not available, then: The label “Acronym” and its value are not displayed



Acceptance criteria	<p>Given the metadata record of a dataset potentially includes value about the dataset's acronym, When a user visits the Dataset webpage, Then the information about the acronym is prominently displayed on the Dataset webpage for the user to easily identify and reference the dataset.</p>
Data generated	N/A
Frequency use	High
Simultaneous conn.	High
RDF snippet	<code>dct:alternative "LINK-VACC"@en ;</code>



Use case 2 – Displaying label and value of “Legal basis”

User story name	As an anonymous user of the HealthData@EU Central Platform, I want to view the information about legal basis for the dataset, so that I can swiftly and effortlessly determine the conditions and limitations for the reuse of the dataset as the legal compliance and relevance of the dataset for my project's scope, ensuring it aligns with legal and ethical standards.
Role(s) concerned	All users
Necessary right(s)	N/A
Precondition scenario	<p>The HealthData@EU Central Platform contains the EU Datasets Catalogue composed of metadata records with the metadata property [dpv:hasLegalBasis dpv:LegalBasis (Open data 0..n) (Protected data 0..n) (Sensitive data 1..n)]</p> <p>The user has access to the HealthData@EU Central Platform from his computer or mobile (an Internet connection is required).</p>
Scenario definition	<p>1) The user visits the Dataset webpage of the HealthData@EU Central Platform.</p> <p>2) The label “Legal basis” and the value of legal basis of the dataset is displayed</p>
Alternative scenarios	N/A
Acceptance criteria	Given the metadata record of a dataset potentially includes information about the dataset's legal basis, When a user visits the Dataset webpage, Then the legal basis is prominently displayed on the web page for the user to easily appreciate the conditions and limitations for the reuse of the dataset.



Data generated	NA
Frequency use	High
Simultaneous conn.	High
RDF snippet	<pre> dpv:hasLegalBasis [a dpv:LegalBasis ; dct:description "CSI Deliberation no. 21/028 of february 18, 2021 "@en; dct:source <https://www.ehealth.fgov.be/ehealthplatform/file/view/AXkNfdPml9vUU fvGGfJr?filename=21-028-f212-AFMPS-vaccinnet-modifi%C3%A9e%20le %2018%20juin%202021.pdf >] ; </pre>



Use case 3 – Displaying label and value of “Retention Period”

User story name	As an anonymous user of the HealthData@EU Central Platform, I want to easily find the value of the retention period for the dataset, so that I can quickly evaluate whether the data's availability aligns with the timeline and demands of my project.
Role(s) concerned	All users
Necessary right(s)	N/A
Precondition scenario	<p>The HealthData@EU Central Platform contains the EU Datasets Catalogue composed of metadata records with the metadata property [healthdcatap:retentionPeriod dct:PeriodOfTime (Open data 0..n) (Protected data 0..n) (Sensitive data 0..n)]</p> <p>The user has access to the HealthData@EU Central Platform, from his computer or mobile (an Internet connection is required).</p>
Scenario definition	<ol style="list-style-type: none"> 1) The user visits a Dataset webpage of the HealthData@EU Central Platform, 2) If metadata are available then: the label “Retention period” and value of retention period is displayed.
Alternative scenarios	<ol style="list-style-type: none"> 3) If metadata are not available, then: The label “Retention period” and its value are not displayed



Acceptance criteria	<p>Given the metadata record of a dataset potentially includes information about the dataset’s retention period,</p> <p>When a user visits the Dataset webpage,</p> <p>Then the dataset's retention period is prominently displayed on the webpage for the user to easily appreciate the conditions and limitations for the reuse of the dataset.</p>
Data generated	N/A
Frequency use	High
Simultaneous conn.	High
RDF snippet	<pre>healthdcatap:retentionPeriod [a dct:PeriodOfTime; dcat:endDate "2034-12-31"^^<http://www.w3.org/2001/XMLSchema#date>; dcat:startDate "2020-03-01"^^<http://www.w3.org/2001/XMLSchema#date>; rdfs:comment: "Provide complementary information"@en];</pre>



Use case 4 – Displaying label and value of “Number of records”

User story name	<p>As an anonymous user of the HealthData@EU Central Platform, I want to easily find the information about (approximate) Dataset Number of Records, so that I can quickly assess its adequacy and applicability to the scope and requirements of my project.</p>
Role(s) concerned	All users
Necessary right(s)	N/A
Precondition scenario	<p>The HealthData@EU Central Platform contains the EU Datasets Catalogue composed of metadata records with the metadata property [healthdcatap:numberofRecords rdfs:Integer (Open data 0..n) (Protected data 0..n) (Sensitive data 1..1)]</p> <p>The user has access to the HealthData@EU Central Platform from his computer or mobile (an Internet connection is required).</p>
Scenario definition	<ol style="list-style-type: none"> 1) The user visits a Dataset webpage of the HealthData@EU Central Platform 2) The label “Number of Records” and value of Number of Records is displayed.
Alternative scenarios	N/A
Acceptance criteria	<p>Given the metadata record of a dataset potentially includes information about the Dataset Number of Records, When a user visits the Dataset webpage, Then the information about Dataset Number of Records is prominently displayed on the webpage for the user to easily appreciate the data scope and usability or relevance of the dataset.</p>



Data generated	N/A
Frequency use	High
Simultaneous conn.	High
RDF snippet	<code>healthdcatap:numberOfrecords: "124866488";</code>



Use case 5 – Displaying label and value of “Number of unique individuals”

User story name	As an anonymous user of the HealthData@EU Central Platform, I want to easily find the information about approximate number of unique individuals in the dataset, so that I can quickly assess its adequacy and applicability to the scope and requirements of my project.
Role(s) concerned	All users
Necessary right(s)	N/A
Precondition scenario	<p>The HealthData@EU Central Platform contains the EU Datasets Catalogue composed of metadata records with the metadata property [healthdcatap:numberofUniqueIndividuals rdfs:Integer (open data 0..1)(protected data 0..1)(sensitive data 1..1)]</p> <p>The user has access to the HealthData@EU Central Platform from his computer or mobile (an Internet connection is required).</p>
Scenario definition	<ol style="list-style-type: none"> 1) The user visits a Dataset webpage of the HealthData@EU Central Platform. 2) The label “Number of unique individuals” and value of Number of unique individuals of the dataset is displayed.
Alternative scenarios	N/A



Acceptance criteria	<p>Given the metadata record of a dataset potentially includes information about the Number of unique individuals, When a user visits the Dataset webpage, Then the information about Number of unique individuals is prominently displayed on the webpage for the user to easily appreciate the conditions and limitations (data scope, usability, relevance) for the reuse of the dataset.</p>
Data generated	N/A
Frequency use	High
Simultaneous conn.	High
RDF snippet	healthdcatap:numberOfUniqueIndividuals: "8914722";



Use case 6 – Displaying the label and value of “Minimum age”

User story name	<p>As an anonymous user of the HealthData@EU Central Platform, I want to know what is the minimum age of the population represented in dataset, so that I can quickly and accurately assess whether the dataset is relevant and adequate for the specific age demographic I am targeting in my project.</p>
Role(s) concerned	All users
Necessary right(s)	N/A
Precondition scenario	<p>The HealthData@EU Central Platform contains the EU Datasets Catalogue composed of metadata records with the metadata property [healthdcatap:minTypicalAge rdfs:Integer (Open data 0..n) (Protected data 0..n) (Sensitive data 1..1)]</p> <p>The user has access to the HealthData@EU Central Platform from his computer or mobile (an Internet connection is required).</p>
Scenario definition	<ol style="list-style-type: none"> 1) The user visits a Dataset webpage of the HealthData@EU Central Platform 2) The label “Minimum age” and value of Minimum age of the population represented in the dataset is displayed.
Alternative scenarios	N/A
Acceptance criteria	<p>Given the metadata record of a dataset potentially includes information about the minimum age of the population represented in the dataset, When a user visits the Dataset webpage, Then the minimum age of the population represented in the dataset is prominently displayed on the webpage for the user to easily appreciate the conditions and limitations for the reuse of the dataset.</p>



Data generated	N/A
Frequency use	High
Simultaneous conn.	High
RDF snippet	healthdcat:minTypicalAge: "18";



Use case 7 – Displaying label and value of “Maximum Age”

User story name	<p>As an anonymous user of the HealthData@EU Central Platform, I want to know what is the maximum age of the population represented in each dataset, so that I can quickly and accurately assess whether the dataset is relevant and adequate for the specific age demographic I am targeting in my project.</p>
Role(s) concerned	All users
Necessary right(s)	N/A
Precondition scenario	<p>The HealthData@EU Central Platform contains the EU Datasets Catalogue composed of metadata records with the metadata property [healthdcatap:maxTypicalAge rdfs:Integer (Open data 0..n) (Protected data 0..n) (Sensitive data 1..1)]</p> <p>The user has access to the HealthData@EU Central Platform from his computer or mobile (an Internet connection is required).</p>
Scenario definition	<ol style="list-style-type: none"> 1) The user visits a Dataset webpage of the HealthData@EU Central Platform 2) The label “Maximum age” and value of maximum age of the population represented in the dataset is displayed.
Alternative scenarios	N/A
Acceptance criteria	<p>Given the metadata record of a dataset potentially includes information about the maximum age of the population represented in the dataset, When a user visits the Dataset webpage, Then the maximum age of the population represented in the dataset is prominently displayed on the webpage for the user to easily appreciate the conditions and limitations for the reuse of the dataset.</p>



Data generated	N/A
Frequency use	High
Simultaneous conn.	High
RDF snippet	healthdcat:maxTypicalAge: "64";



Use case 8 – Displaying label and value of “Applicable Legislation”

User story name	<p>As an anonymous user of the HealthData@EU Central Platform, I want to easily find information about which Applicable Legislations are associated with the dataset, so that I can quickly and accurately determine the legal compliance and relevance of the dataset for my project's scope, ensuring it aligns with legal and ethical standards.</p>
Role(s) concerned	All users
Necessary right(s)	N/A
Precondition scenario	<p>The HealthData@EU Central Platform contains the EU Datasets Catalogue composed of metadata records with the metadata property [dcatap:applicableLegislation rdfs:Resource (Open data 1..n) (Protected data 1..n) (Sensitive data 1..n)]</p> <p>The user has access to the HealthData@EU Central Platform from his computer or mobile (an Internet connection is required).</p>
Scenario definition	<ol style="list-style-type: none"> 1) The user visits a single Datasets webpage of the HealthData@EU Central Platform 2) The label “Applicable Legislation” and value of Applicable Legislation for the dataset is displayed.
Alternative scenarios	N/A
Acceptance criteria	<p>Given the metadata record of a dataset potentially includes information about the Applicable Legislation for the dataset, When a user visits the Dataset webpage, Then the Applicable Legislation for the dataset is prominently displayed on the webpage for the user to easily appreciate the legal compliance and the conditions and limitations for the reuse of the dataset.</p>



Data generated	N/A
Frequency use	High
Simultaneous conn.	High
RDF snippet	<code>dcatap:applicableLegislation</code> <code><http://data.europa.eu/eli/reg/2022/868/oj>;</code>



Use case 9 – Displaying label and value of “Health Data Access Body”

User story name	<p>As an anonymous user of the HealthData@EU Central Platform, I want to identify the specific Health Data Access Body responsible for cataloguing health datasets, so that I can quickly understand which organisation to contact for accessing the dataset and ensure that I'm adhering to the correct protocols and procedures for dataset access requests.</p>
Role(s) concerned	All users
Necessary right(s)	N/A
Precondition scenario	<p>The HealthData@EU Central Platform contains the EU Datasets Catalogue composed of metadata records with the metadata property [healthdcatap:hdab foaf:Agent (Open data 1..1) (Protected data 1..1) (Sensitive data 1..1)]</p> <p>The user has access to the HealthData@EU Central Platform from his computer or mobile (an Internet connection is required).</p>
Scenario definition	<ol style="list-style-type: none"> 1) The user visits a single Dataset webpage of the HealthData@EU Central Platform 2) The label “Health Data Access Body” and value of Health Data Access Body for the dataset is displayed.
Alternative scenarios	N/A



Acceptance criteria	<p>Given the metadata record of a dataset potentially includes contact information about the Health Data Access Body for the dataset, When a user visits the Dataset webpage, Then the Health Data Access Body for the dataset is prominently displayed on the webpage for the user to quickly understand which organisation to contact for dataset access requests.</p>
Data generated	N/A
Frequency use	High
Simultaneous conn.	High
RDF snippet	<pre> healthdcatap:hdab [a foaf:Organization; locn:address [a locn:Address; locn:adminUnitL1 "http://publications.europa.eu/resource/authority/country/BEL"; locn:fullAddress "Galileelaan 5, Bus 2"; locn:postCode "1210"; locn:postName "Saint-Josse-ten-Noode"]; foaf:homepage <https://www.hda.belgium.be>; foaf:mbox <mailto:info@hda.fgov.be>; foaf:name "Belgian Health Data Agency"]; </pre>



Use case 10 – Displaying label and value of “Publisher Type”

User story name	<p>As an anonymous user of the HealthData@EU Central Platform, I want to easily and accurately identify the Publisher Type of the organisation responsible for the dataset, so that I can quickly assess the dataset's credibility and relevance to my project.</p>
Role(s) concerned	All users
Necessary right(s)	N/A
Precondition scenario	<p>The HealthData@EU Central Platform contains the EU Datasets Catalogue composed of metadata records with the metadata property [healthdcatap:publisherType (Open data 0..n) (Protected data 0..n) (Sensitive data 1..n)]</p> <p>The user has access to the HealthData@EU Central Platform from his computer or mobile (an Internet connection is required).</p>
Scenario definition	<p>1) The user visits a Dataset webpage of the HealthData@EU Central Platform</p> <p>2) If metadata are available, then:</p> <p>The label "Publisher Type" and value of Publisher Type of the publisher is displayed.</p>
Alternative scenarios	<p>2) If metadata are not available, then:</p> <p>The label "Publisher Type" and its value are not displayed.</p>



Acceptance criteria	<p>Given the metadata record of a dataset potentially includes information about the Publisher Type, When a user visits the Dataset webpage, Then the Publisher Type is prominently displayed on the webpage for the user to quickly understand which type of organisation is responsible for the publication of the dataset.</p>
Data generated	N/A
Frequency use	High
Simultaneous conn.	High
RDF snippet	<pre>healthcatap:publisherType <http://healthdataportal.eu/resource/authority/publisher- type/NationalPublicHealthInstitut>;</pre>



Use case 11 – Displaying label and value of “Publisher Notes”

User story name	As an anonymous user of the HealthData@EU Central Platform, I want to read a detailed description of the publisher responsible for each dataset, including their background and expertise, so that I can more effectively evaluate the dataset's credibility and relevance to my project needs.
Role(s) concerned	All users
Necessary right(s)	N/A
Precondition scenario	<p>The HealthData@EU Central Platform contains the EU Datasets Catalogue composed of metadata records with the metadata property [healthdcatap:publisherNote rdfs:Literal (Open data 0..n) (Protected data 0..n) (Sensitive data 1..n)]</p> <p>The user has access to the HealthData@EU Central Platform from his computer or mobile (an Internet connection is required).</p>
Scenario definition	<ol style="list-style-type: none"> 1) The user visits a Dataset webpage of the HealthData@EU Central Platform 2) If metadata are available, then: The label “Publisher notes” and value of Publisher notes is displayed.
Alternative scenarios	<ol style="list-style-type: none"> 2) If metadata are not available, then: The label “Publisher notes” and its value are not displayed.



Acceptance criteria	<p>Given the metadata record of a dataset potentially includes a note about the publisher, When a user visits the Dataset webpage, Then a note about the publisher is prominently displayed on the webpage for the user to effectively evaluate the dataset's credibility and relevance to his project needs.</p>
Data generated	N/A
Frequency use	High
Simultaneous conn.	High
RDF snippet	<p>healthdcatap:publisherNote "Sciensano is a research institute and the national public health institute of Belgium. It is a so-called federal scientific institution that operates under the authority of the federal minister of Public Health and the federal minister of Agriculture of Belgium."@en;</p>



Use case 12 – Displaying label and value of “Conforms To”

User story name	<p>As an anonymous user of the HealthData@EU Central Platform, I want to know the data model (Standard) of each dataset, including its structure, format, schema, and any interrelationships among data elements, so that I can efficiently determine the dataset's compatibility and readiness for integration into my specific project.</p>
Role(s) concerned	All users
Necessary right(s)	N/A
Precondition scenario	<p>The HealthData@EU Central Platform contains the EU Datasets Catalogue composed of metadata records with the metadata property [dct:conformsTo dct:Standard (Open data 0..n) (Protected data 0..n) (Sensitive data 0..n)]</p> <p>The user has access to the HealthData@EU Central Platform from his computer or mobile (an Internet connection is required).</p>
Scenario definition	<ol style="list-style-type: none"> 1) The user visits a Dataset webpage of the HealthData@EU Central Platform 2) If metadata are available, then: The label “Conforms to” and value of Dataset conformance to the data model of the dataset is displayed. 3) The user can access additional information about the data model by clicking on the name of data model.
Alternative scenarios	<ol style="list-style-type: none"> 2) If metadata are not available, then: The label “Conforms to” and is value are not displayed.



Acceptance criteria	<p>Given the metadata record of a dataset potentially includes information about the dataset conformance to data model, When a user visits the Dataset webpage, Then the information about Dataset conformance to data model of the dataset is prominently displayed on the webpage.</p>
Data generated	N/A
Frequency use	High
Simultaneous conn.	High
RDF snippet	<code>dct:conformsTo <https://www.wikidata.org/wiki/Q19597236>;</code>
Comments:	dct:conformsTo is a property of DCAT-AP 3.0.0 and might be already supported by the data.europa.eu portal.



Use case 13 – Displaying label and value of “Coding systems”

User story name	As an anonymous user of the HealthData@EU Central Platform, I want to clearly understand the specific medical coding systems (such as ICD-10, SNOMED CT) used in each dataset, so that I can accurately evaluate the dataset's relevance and applicability to my project.
Role(s) concerned	All users
Necessary right(s)	N/A
Precondition scenario	<p>The HealthData@EU Central Platform contains the EU Datasets Catalogue composed of metadata records with the metadata property [healthdcatap:hasCodingSystem dct:Standard (Open data 0..n) (Protected data 0..n) (Sensitive data 0..n)]</p> <p>The user has access to the HealthData@EU Central Platform from his computer or mobile (an Internet connection is required).</p>
Scenario definition	<ol style="list-style-type: none"> 1) The user visits a Dataset webpage of the HealthData@EU Central Platform 2) If metadata are available, then: The label “Coding systems” and value of coding systems used in the dataset are displayed. 3) The user can access additional information about the coding systems by clicking on the name of the coding system.
Alternative scenarios	<ol style="list-style-type: none"> 2) If metadata are not available, then: The label “Coding Systems” and its value are not displayed.
Acceptance criteria	<p>Given the metadata record of a dataset potentially includes information about the coding systems used by the dataset, When a user visits the Dataset webpage, Then the coding systems are prominently displayed on the webpage.</p>



Data generated	N/A
Frequency use	High
Simultaneous conn.	High
RDF snippet	<pre>healthdcatap:hasCodingSystem <https://www.wikidata.org/wiki/Property:P1690>, <https://www.wikidata.org/wiki/Property:P4229>;</pre>



Use case 14 – Displaying label and value of ‘Generated by’

User story name	<p>As an anonymous user of the HealthData@EU Central Platform, I want to see information about how the dataset was generated, e.g., research, survey, clinical trial, or public health monitoring, so that I can better assess the dataset's relevance, methodology, and potential application in my own field of study.</p>
Role(s) concerned	All users
Necessary right(s)	N/A
Precondition scenario	<p>The HealthData@EU Central Platform contains the EU Datasets Catalogue composed of metadata records with the metadata property [prov:wasGeneratedBy prov:Activity (Open data 0..n) (Protected data 0..n) (Sensitive data 0..n)]</p> <p>The user has access to the HealthData@EU Central Platform from his computer or mobile (an Internet connection is required).</p>
Scenario definition	<ol style="list-style-type: none"> 1) The user visits a Dataset webpage of the HealthData@EU Central Platform 2) If metadata are available, then: <p>The label “Generated by” and value of activity involved in generating the dataset is displayed.</p>
Alternative scenarios	<ol style="list-style-type: none"> 2) If metadata are not available, then: <p>The label “Generated by” and its value are not displayed.</p>



Acceptance criteria	<p>Given the metadata record of a dataset potentially includes information about the type of activity involved in generating the dataset, When a user visits the Dataset webpage, Then the information about the type of activity is prominently displayed on the webpage.</p>
Data generated	N/A
Frequency use	High
Simultaneous conn.	High
RDF snippet	<pre> prov:wasGeneratedBy [a prov:Activity; dct:type <http://dbpedia.org/resource/Record_linkage> ; foaf:page <https://www.sciensano.be/sites/default/ files/information_letter_linkvacc_20221010_en.pdf>; rdfs:label "Pseudonymisation and Anonymisation"@en ; rdfs:seeAlso <https://www.ehealth.fgov.be/ehealthplatform/fr/ service- codage-anonymisation-et-ttp> prov:startedAtTime "2021-01- 01"^^xsd:date ; prov:wasAssociatedWith [a prov:Agent; foaf:givenName "Dr. Joris van Loenhout"^^xsd:string; foaf:mbox <mailto:Joris.VanLoenhout@sciensano.be>; foaf:homePage <https://www.sciensano.be/fr/people/joris-van- loenhout>; prov:actedOnBehalfOf [a prov:Agent, prov:Organization; foaf:name "Sciensano"];]; </pre>



Use case 14 – Displaying label and value of “Purpose”

User story name	<p>As an anonymous user of the HealthData@EU Central Platform, I want to understand the specific purposes for which datasets are processed, including the reasons for data collection, and intended uses, so that I can ensure my project's activities and objectives are in scope and comply with the original consent and purposes for which the data was collected.</p>
Role(s) concerned	All users
Necessary right(s)	N/A
Precondition scenario	<p>The HealthData@EU Central Platform contains the EU Datasets Catalogue composed of metadata records with the metadata property [dpv:hasPurpose dpv:Purpose (Open data 0..n) (Protected data 0..n) (Sensitive data 1..n)]</p> <p>The user has access to the HealthData@EU Central Platform from his computer or mobile (an Internet connection is required).</p>
Scenario definition	<ol style="list-style-type: none"> 1) The user visits a Dataset webpage of the HealthData@EU Central Platform 2) If metadata are available, then: <p>The label “Purpose” and value of the specific purpose for which the data is processed is displayed.</p>
Alternative scenarios	N/A



Acceptance criteria	<p>Given the metadata record of a dataset potentially includes information about the specific purpose for which personal data is processed, When a user visits the Dataset webpage, Then the information about the specific purpose for which personal data is processed is prominently displayed on the webpage.</p>
Data generated	N/A
Frequency use	High
Simultaneous conn.	High
RDF snippet	<pre> dpv:hasPurpose [a dpv:Purpose ; dct:description "The primary objective of Sciensano's LINK-VACC project is to monitor COVID-19 vaccines post-authorization and evaluate the public health value of prioritizing vaccination for people with comorbidities. This involves assessing the vaccines' effectiveness and safety in the broader population context, beyond the limited scope of clinical trials, and determining future vaccination policies in public health emergencies such as epidemics or pandemics"@en;]; </pre>



Use case 15 – Displaying label and value of “Provenance”

User story name	As an anonymous user of the HealthData@EU Central Platform, I want to know the information about dataset provenance, such as lineage of each dataset, including its source, history of changes, and any processing it has undergone, so that I can assess its quality, reliability, and suitability for my research or analysis projects.
Role(s) concerned	All users
Necessary right(s)	N/A
Precondition scenario	<p>The HealthData@EU Central Platform contains the EU Datasets Catalogue composed of metadata records with the metadata property [dct:provenance dct:ProvenanceStatement (Open data 0..n) (Protected data 0..n) (Sensitive data 1..n)]</p> <p>The user has access to the HealthData@EU Central Platform from his computer or mobile (an Internet connection is required).</p>
Scenario definition	<ol style="list-style-type: none"> 1) The user visits a Dataset webpage of the HealthData@EU Central Platform 2) The label “Provenance” and value of Provenance are displayed.
Alternative scenarios	N/A
Acceptance criteria	<p>Given the metadata record of a dataset potentially includes information about the Provenance of dataset,</p> <p>When a user visits the Dataset webpage, Then the information about the provenance of the dataset is prominently displayed on the webpage.</p>



Data generated	N/A
Frequency use	High
Simultaneous conn.	High
RDF snippet	<pre>dct:provenance [a dct:ProvenanceStatement; rdfs:label "The data for the LINK-VACC project is sourced from several existing databases, including Vaccinnet+, HealthData COVID-19 database (Contact tracing and Clinic database), CoBRHA, STATBEL, and the AIM database. These databases collectively provide comprehensive demographic, clinical, and socio-economic data relevant to the project's objectives"@en];</pre>
Comments:	dct:provenance is a property of DCAT-AP 3.0



Use case 16 – Displaying label and value of “Population Coverage”

User story name	<p>As an anonymous user of the HealthData@EU Central Platform, I want to see information about the population coverage for dataset, including demographic details like age range, geographic location, and socio-economic status, so that I can accurately determine the dataset's relevance and applicability to my research or analysis projects, particularly in ensuring representativeness and diversity of the study population.</p>
Role(s) concerned	All users
Necessary right(s)	N/A
Precondition scenario	<p>The HealthData@EU Central Platform contains the EU Datasets Catalogue composed of metadata records with the metadata property [healthdcatap:populationCoverage rdfs:Literal (Open data 0..n) (Protected data 0..n) (Sensitive data 1..n)]</p> <p>The user has access to the HealthData@EU Central Platform from his computer or mobile (an Internet connection is required).</p>
Scenario definition	<ol style="list-style-type: none"> 1) The user visits a Dataset webpage of the HealthData@EU Central Platform 2) The label “Population coverage” and value of population coverage for the dataset are displayed.
Alternative scenarios	N/A
Acceptance criteria	<p>Given the metadata record of a dataset potentially includes information about the population coverage of the dataset, When a user visits the Dataset webpage, Then the information about the population coverage is prominently displayed on the webpage.</p>



Data generated	N/A
Frequency use	High
Simultaneous conn.	High
RDF snippet	<p>healthdcatap:populationCoverage "The population targeted by the LINK-VACC project comprises all individuals in Belgium who have received a COVID-19 vaccine, undergone testing for COVID-19, or have been hospitalized with a confirmed diagnosis of COVID-19. The project also considers healthcare professionals and the general Belgian population for understanding vaccination coverage and effectiveness, especially among those with comorbidities and varying socio-economic backgrounds"@en</p>



Use case 17 – Displaying label and value of “Dataset analytics”

User story name	<p>As an anonymous user of the HealthData@EU Central Platform, I want to access comprehensive dataset analytics for dataset, such as technical reports, utility metrics like usage statistics and citation counts, or data services like augmented analytics dashboards, so that I can assess the dataset's quality, relevance, or analytical insights, ensuring the data aligns with my specific research objectives or compliance requirements.</p>
Role(s) concerned	All users
Necessary right(s)	N/A
Precondition scenario	<p>The HealthData@EU Central Platform contains the EU Datasets Catalogue composed of metadata records with the metadata property [healthdcatap:analytics dcat:Distribution (Open data 0..n) (Protected data 0..n) (Sensitive data 0..n)]</p> <p>The user has access to the HealthData@EU Central Platform from his computer or mobile (an Internet connection is required).</p>
Scenario definition	<ol style="list-style-type: none"> 1) The user visits a Dataset webpage of the HealthData@EU Central Platform 2) If metadata are available, then: The label “Datasets analytics” and links to the Dataset's analytics are displayed.
Alternative scenarios	<ol style="list-style-type: none"> 2) If metadata are not available, then: The label “Datasets analytics” and links to the Dataset's analytics are not displayed.



Acceptance criteria	<p>Given the metadata record of a dataset potentially includes links providing access to dataset's analytics, When a user visits the dataset's webpage, Then the links providing access to dataset's analytics are prominently displayed on the webpage.</p>
Data generated	N/A
Frequency use	High
Simultaneous conn.	High
RDF snippet	<pre>healthdcatap:analytics [a dcat:Distribution; dcatap:applicableLegislation <http://data.europa.eu/eli/reg/2022/868/oj>; dct:title "Technical report number of unique study subjects available by environment for project HDBP0250"@en; dcat:accessURL <https://fair.healthdata.be/sites/ default/files/distribution/ d43a158e-7d13-4660-bbc3-9d3f8d5501e5/ Technical_report_number_of_unique_study_ subjects_available_by_environment_ for_project_HDBP0250.csv> ; dct:format <http://publications.europa.eu/resource/authority/file- type/CSV>; dcat:mediaType <http://www.iana.org/assignments/media-types/text/tab- separated-values>.];</pre>



Use case 18 – Displaying label “Sample distributions” and sub-labels “Link to sample”, “Format”, “Updated”, “Actions” with their relevant values

User story name	<p>As an anonymous user of the HealthData@EU Central Platform, I want to see representative samples from dataset, including a variety of data points and a sample of the data structure (e.g., data dictionary RDF-ized with the CSVW vocabulary) so that I can make an informed decision about the dataset's appropriateness for my research needs and determine if it aligns with my study's scope and objectives.</p>
Role(s) concerned	All users
Necessary right(s)	N/A
Precondition scenario	<p>The HealthData@EU Central Platform contains the EU Datasets Catalogue composed of metadata records with the metadata property [adms:sample dcat:Distribution (Open data 0..n) (Protected data 0..n) (Sensitive data 1..n)]</p> <p>The user has access to the HealthData@EU Central Platform from his computer or mobile (an Internet connection is required).</p>



<p>Scenario definition</p>	<p>1) The user visits a Dataset webpage of the HealthData@EU Central Platform</p> <p>2) If metadata are available, then:</p> <p>The label "Sample distributions" and sub-labels "Link to sample", "Format", "Updated", and "Actions" with their relevant values are displayed.</p> <p>For sub-label "Link to sample" the value name of sample distribution is displayed. An element of type accordion is also displayed and contains additional detail for the sample.</p> <p>For sub-label "Format" the value of format is displayed (rdf, cvs, etc)</p> <p>For sub-label "Updated" the value of creating sample distribution is displayed.</p> <p>For sub-label "Actions" the linked button "Preview" is displayed. User can preview the file by clicking on the Preview link.</p>
<p>Alternative scenarios</p>	<p>2) If metadata are not available, then:</p> <p>The label "Sample distributions" and links are not displayed.</p>
<p>Acceptance criteria</p>	<p>Given the metadata record of a dataset potentially includes links providing access to sample data, When a user visits the Dataset webpage, Then he can see the links providing access to sample data which are prominently displayed on the webpage.</p>
<p>Data generated</p>	<p>N/A</p>
<p>Frequency use</p>	<p>High</p>
<p>Simultaneous conn.</p>	<p>High</p>



RDF snippet	<pre>adms:sample [a dcat:Distribution ; dct:description "Synthetic data of the HealthPilot Use Case"@en; dcat:downloadURL <https://github.com/CAVDgit/ EHDS2_UC_Sciensano/blob/main/ use_case_1_synthetic_data_10K_individuals.csv>; dcat:mediaType <http://www.iana.org/assignments/media-types/text/tab-separated-values> ;];</pre>
Comments:	<p>adms:sample is a property of DCAT-AP 3.0 of range Class distribution which might not have been implemented in the data.europa.eu technical stack.</p>



Use case 19 – Displaying label and value of “Categories”

User story name	As an anonymous user of the HealthData@EU Central Platform, I want to quickly identify the categories associated with each dataset as defined by the EHDS2 Regulation, so that I can efficiently locate and utilise datasets that are most relevant to my project.
Role(s) concerned	All users
Necessary right(s)	N/A
Precondition scenario	<p>The HealthData@EU Central Platform contains the EU Datasets Catalogue composed of metadata records with the metadata property [healthdcatap:healthCategory, subproperty of dct:subject skos:Concept 1..n]</p> <p>The user has access to the HealthData@EU Central Platform from his computer or mobile (an Internet connection is required).</p>
Scenario definition	<ol style="list-style-type: none"> 1) The user visits a Dataset webpage of the HealthData@EU Central Platform 2) The label “Categories” and value of categories are displayed.
Alternative scenarios	N/A
Acceptance criteria	<p>Given the metadata record of a dataset includes a list of categories for the dataset,</p> <p>When a user visits the Dataset webpage, Then the categories are displayed in a dedicated "Categories" section ensuring immediate visibility.</p>



Data generated	N/A
Frequency use	High
Simultaneous conn.	High
RDF snippet	<code>healthdcatap:healthCategory <http://healthdata.ec.europa.eu/.../(o)>;</code>
Comments:	<ul style="list-style-type: none"> · The list of categories is defined by Art. 33 of the EHDS Regulation proposal. · A controlled vocabulary must be created and maintained for the list of categories.



Use case 20 – Displaying label and value of “Themes”

User story name	As an anonymous user of the HealthData@EU Central Platform, I want to quickly identify the themes associated with each dataset, so that I can efficiently locate and utilise datasets that are most relevant to my project.
Role(s) concerned	All users
Necessary right(s)	N/A
Precondition scenario	<p>The HealthData@EU Central Platform contains the EU Datasets Catalogue composed of metadata records with the metadata property [healthdcatap:healthTheme, subproperty of dct:subject skos:Concept (Open data 1..n) (Protected data 1..n) (Sensitive data 1..n)]</p> <p>The user has access to the HealthData@EU Central Platform from his computer or mobile (an Internet connection is required).</p>
Scenario definition	<ol style="list-style-type: none"> 1) The user visits a Dataset webpage of the HealthData@EU Central Platform 2) The label “Themes” and value of themes are displayed.
Alternative scenarios	<p>2) If the metadata are not available, then:</p> <p>The label “Themes” and value of themes are not displayed.</p>
Acceptance criteria	<p>Given the metadata record of a dataset includes a list of themes for the dataset,</p> <p>When a user visits the Dataset webpage, Then the themes are displayed in a dedicated "Themes" section ensuring immediate visibility.</p>
Data generated	N/A
Frequency use	High



Simultaneous	High
conn.	
RDF snippet	healthdcatap:healthTheme < https://www.wikidata.org/wiki/Q58624061 >, < https://www.wikidata.org/wiki/Q7907952 >;



Use case 21 – Displaying label and value of “Health Codes”

User story name	As an anonymous user of the HealthData@EU Central Platform, I want to quickly identify the health codes (for example, codes from a medical thesaurus) associated with each dataset, so that I can efficiently locate and utilise datasets that are most relevant to my project.
Role(s) concerned	All users
Necessary right(s)	N/A
Precondition scenario	<p>The HealthData@EU Central Platform contains the EU Datasets Catalogue composed of metadata records with the metadata property [healthdcatap:hasCodeValues, skos:Concept (Open data 0..n) (Protected data 0..n) (Sensitive data 0..n)]</p> <p>The user has access to the HealthData@EU Central Platform from his computer or mobile (an Internet connection is required).</p>
Scenario definition	<ol style="list-style-type: none"> 1) The user visits a Dataset webpage of the HealthData@EU Central Platform 2) If the metadata are available, then: The label “Health Codes” and value of the health codes are displayed.
Alternative scenarios	<ol style="list-style-type: none"> 2) If the metadata are not available, then: The label “Health Codes” are value of the health codes are not displayed.



Acceptance criteria	<p>Given the metadata record of a dataset includes a list of codes for the dataset, When a user visits the Dataset webpage, Then these codes are displayed in a dedicated "Health codes" section ensuring immediate visibility.</p>
Data generated	N/A
Frequency use	High
Simultaneous conn.	High



RDF snippet	<pre> healthdcatap:hasCodeValues [a skos:Concept; skos:inScheme [a skos:ConceptScheme; dct:identifier "https://www.wikidata.org/wiki/ Property:P494"^^ <http://www.w3.org/2001/XMLSchema#anyURI>; skos:prefLabel "International Classification of Diseases, 10th Revision (ICD- 10)"@en; skos:definition "ICD-10 is a medical classification list by the World Health Organization."@en ; skos:notation "ICD-10"; owl:versionInfo "Version:2019"]; dct:identifier "https://icd.who.int/browse10/ 2019/en#/Y59.0"^^ <http://www.w3.org/2001/XMLSchema#anyURI>; skos:notation "Y59.0"; skos:prefLabel "Viral vaccines"@en]; </pre>
-------------	--



Use case 22 – Displaying label and value of “Personal Data”

User story name	<p>As an anonymous user of the HealthData@EU Central Platform, I want to quickly and clearly identify the specific types of sensitive personal data (e.g., names, addresses, health information) contained in each dataset, so that I can efficiently determine the datasets that align with the scope and requirements of my project, ensuring a targeted and relevant selection.</p>
Role(s) concerned	All users
Necessary right(s)	N/A
Precondition scenario	<p>The HealthData@EU Central Platform contains the EU Datasets Catalogue composed of metadata records with the metadata property [dpv:hasPersonalData dpv:PersonalData (Open data 0..n) (Protected data 0..n) (Sensitive data 0..n)]</p> <p>The user has access to the HealthData@EU Central Platform from his computer or mobile (an Internet connection is required).</p>
Scenario definition	<ol style="list-style-type: none"> 1) The user visits a Dataset webpage of the HealthData@EU Central Platform 2) If the metadata are available, then: Label “Persona Data” and value of personal data are displayed
Alternative scenarios	<ol style="list-style-type: none"> 2) If the metadata are not available, then: Label “Persona Data” and value of personal data are not displayed



Acceptance criteria	<p>Given the metadata record of a dataset includes specific types of sensitive personal data for the dataset, When a user visits the dataset's webpage, Then the specific types of sensitive personal data are displayed in a dedicated "Personal data" section ensuring immediate visibility.</p>
Data generated	N/A
Frequency use	High
Simultaneous conn.	High
RDF snippet	<p>dpv:hasPersonalData dpv-pd:Gender, dpv-pd:Age, dpv-pd:Location, dpv-pd:Nationality, dpv-pd:Education, dpv-pd:HealthRecord;</p>



Use case 23 – Displaying label and value of “Trusted data holder”

User story name	<p>As an anonymous user of the HealthData@EU Central Platform, I want to quickly and clearly identify the datasets that are published by Trusted Data Holders, so that I can efficiently determine the datasets that a have simplified data permit procedures.</p>
Role(s) concerned	All users
Necessary right(s)	N/A
Precondition scenario	<p>The HealthData@EU Central Platform contains the EU Datasets Catalogue composed of metadata records with the metadata property [healthdcatap:trustedDataHolder (Open data 0..n) (Protected data 0..n) (Sensitive data 0..n)]</p> <p>The user has access to the HealthData@EU Central Platform from his computer or mobile (an Internet connection is required).</p>
Scenario definition	<ol style="list-style-type: none"> 1) The user visits a Dataset webpage of the HealthData@EU Central Platform 2) If the metadata are available, then: Label “Trusted Data Holder” is displayed
Alternative scenarios	<ol style="list-style-type: none"> 2) If the metadata are not available, then: Label “Trusted Data Holder” is not displayed



Acceptance criteria	<p>Given the metadata record of a dataset published by a Trusted Data Holder, When a user visits the dataset's webpage, Then the label Trusted Data Holder is displayed ensuring immediate identification.</p>
Data generated	N/A
Frequency use	High
Simultaneous conn.	High
RDF snippet	<pre>healthdcatap:trustedDataHolder true;</pre>



2.2 Use cases improving the EU datasets catalogue webpage with advanced filtering capabilities.

Use case 24 - Filter EU Datasets Catalogue by metadata values in “Health Data Access Bodies” list

User story name	As an anonymous user, I want to be able to filter EU Datasets Catalogue by Health Data Access Bodies in the HealthData@EU Central Platform. So that I can quickly find and access data relevant to a specific Health Data Access Body .
Role(s) concerned	All users
Necessary right(s)	N/A
Precondition scenario	The HealthData@EU Central Platform contains the EU Datasets Catalogue with metadata records with the relevant metadata property [healthdcatap:hdab foaf:Agent (Open data 1..1) (Protected data 1..1) (Sensitive data 1..1)] The user has access to the HealthData@EU Central Platform from his computer or mobile (an Internet connection is required).
Scenario definition	<ol style="list-style-type: none"> 1) The user visits the EU Datasets Catalogue webpage of the HealthData@EU Central Platform 2) The user selects from the metadata record filter one HDAB or several HDABs 3) The result page displays the list of datasets from EU Datasets Catalogue corresponding to the selected HDABs.
Alternative scenarios	N/A



<p>Acceptance criteria</p>	<p>Health Data Access Bodies (HDABs) selection feature: The HealthData@EU Central Platform must include an option to filter datasets by HDABs. This feature should be easily accessible and intuitive to use.</p> <p>Comprehensive HDABs List: The filter should include the list of onboarded HDABs.</p> <p>Multiple HDABs Selection: Users should have the ability to select multiple HDABs at once.</p> <p>Seamless integration with other filters: This HDABs filter should work seamlessly with other existing filters in the catalogue, such as categories, health themes, etc., allowing for a dynamic refined search.</p> <p>Instant update of dataset list: Upon selecting a HDAB or a combination of HDABs, the list of datasets should update instantly to display only those that are relevant to the chosen HDABs.</p> <p>Clear indication of selected HDABs: The interface should clearly indicate which HDAB or HDABs have been selected at any given time.</p> <p>Reset option: There should be an option to easily reset or clear the HDAB selection to start a new search.</p> <p>Performance: The filtering process should be quick and not cause any significant delay in the user experience.</p> <p>Accessibility and usability: The feature should be accessible to users with disabilities and follow best practices for usability (ref: Directive (EU) 2016/2102).</p> <p>User Help and Documentation: There should be clear documentation or help available explaining how to use the HDAB filter.</p>
<p>Data generated</p>	<p>N/A</p>
<p>Frequency use</p>	<p>High (200-1000/day) to Very High (> 1000/day)</p>
<p>Simultaneous conn.</p>	<p>High (100 -1000) to Very High (> 1000)</p>



RDF snippet	<pre>healthdcatap:hdab [a foaf:Organization; locn:address [a locn:Address; locn:adminUnitL1 "BEL"; locn:postCode "1210"; locn:postName "Saint-Josse-ten-Noode"; locn:thoroughfare "Galileelaan 5, Bus 2"]; foaf:mbox <mailto:info@hda.fgov.be>; foaf:homepage <https://www.hda.belgium.be>; foaf:name "Belgian Health Data Agency"];</pre>
-------------	---



Use case 25 - Filter EU Datasets Catalogue by metadata values in Publisher list

User story name	<p>As an anonymous user, I want to be able to filter EU Datasets Catalogue by Publisher in the HealthData@EU Central Platform. So that I can quickly find and access data relevant to a specific Publisher.</p>
Role(s) concerned	All users
Necessary right(s)	N/A
Precondition scenario	<p>The HealthData@EU Central Platform contains metadata records with the relevant metadata property [dct:publisher foaf:Agent (Open data 0..n) (Protected data 1..1) (Sensitive data 1..1)]</p> <p>The user has access to the HealthData@EU Central Platform from his computer or mobile (an Internet connection is required).</p>
Scenario definition	<ol style="list-style-type: none"> 1) The user visits the EU Datasets Catalogue of the HealthData@EU Central Platform 2) The user selects from the metadata record filter one Publisher or several Publishers 3) The result page displays the list of datasets from EU Datasets Catalogue corresponding to the selected Publishers
Alternative scenarios	<p>If there are no values in "Publisher" list in EU Datasets Catalogue, then:</p> <p>The Filter Publisher is not displayed</p>



<p>Acceptance criteria</p>	<p>Publisher selection feature: The HealthData@EU Central Platform must include an option to filter metadata sets by Publisher. This feature should be easily accessible and intuitive to use.</p> <p>Comprehensive Publisher List: The filter should include the comprehensive list of Publishers.</p> <p>Multiple Publisher Selection: Users should have the ability to select multiple Publishers at once.</p> <p>Seamless integration with other filters: This Publishers filter should work seamlessly with other existing filters in the catalogue, such as categories, health themes, etc., allowing for a dynamic refined search.</p> <p>Instant update of dataset list: Upon selecting a Publisher or a combination of Publishers, the list of metadata sets should update instantly to display only those that are relevant to the chosen Publishers.</p> <p>Clear indication of selected Publishers: The interface should clearly indicate which Publisher or Publishers have been selected at any given time.</p> <p>Reset option: There should be an option to easily reset or clear the Publisher selection to start a new search.</p> <p>Performance: The filtering process should be quick and not cause any significant delay in the user experience.</p> <p>Accessibility and usability: The feature should be accessible to users with disabilities and follow best practices for usability (ref: Directive (EU) 2016/2102).</p> <p>User Help and Documentation: There should be clear documentation or help available explaining how to use the Publisher filter.</p>
<p>Data generated</p>	<p>N/A</p>
<p>Frequency use</p>	<p>High</p>
<p>Simultaneous conn.</p>	<p>High</p>



RDF snippet	<pre>dct:publisher [a foaf:Organization; locn:address [a locn:Address; foaf:name "Sciensano"; foaf:mbox <mailto:info@sciensano.be>; foaf:homepage <https://sciensano.be>;];];</pre>
RDF snippet	<pre>dct:publisher <https://org.belgif.be/id/CbeRegisteredEntity/0693876830>;</pre>



Use case 26 – Filter EU Datasets Catalogue by metadata values in Categories list

User story name	As an anonymous user, I want to be able to filter EU Datasets Catalogue by Categories in the HealthData@EU Central Platform. So that I can quickly find and access metadata relevant to a specific Category for my project.
Role(s) concerned	All users
Necessary right(s)	N/A
Precondition scenario	<p>The HealthData@EU Central Platform contains metadata records with the relevant metadata property [heathdcatap:healthCategory, subproperty of dct:subject skos:Concept (Open data 0..n) (Protected data 0..n) (Sensitive data 1..n)]</p> <p>The user has access to the HealthData@EU Central Platform from his computer or mobile (an Internet connection is required).</p>
Scenario definition	<ol style="list-style-type: none"> 1) The user visits the EU Datasets Catalogue webpage of the HealthData@EU Central Platform 2) The user selects from the metadata record filter one Category or several Categories <p>The result page displays the list of datasets from EU Datasets Catalogue corresponding to the selected Categories</p>
Alternative scenarios	N/A



Acceptance criteria

Categories selection feature: The HealthData@EU Central Platform must include an option to filter metadata sets by **Categories**. This feature should be easily accessible and intuitive to use.

Comprehensive Categories List: The filter should include the comprehensive list of Categories. The list of categories refers to the datasets as described in Commission Regulation on the European Health Data Space laying down a list of minimum categories of electronic data for secondary use, Art. 33. A controlled vocabulary that must be used for the HealthDCAT-AP property (i.e.: heathdcatap:healthCategory) is maintained by the HealthData@EU infrastructure.

Multiple Categories Selection: A metadata set may be associated with multiple categories. Users should have thus the ability to select multiple Categories at once.

Seamless integration with other filters: This Category filter should work seamlessly with other existing filters in the catalogue, such as HDABs, health themes, etc., allowing for a dynamic refined search.

Instant update of metadata set list: Upon selecting a Category or a combination of Categories, the list of metadata sets should update instantly to display only those that are relevant to the chosen Categories.

Clear indication of selected Categories: The interface should clearly indicate which Category or Categories have been selected at any given time.

Reset option: There should be an option to easily reset or clear the Category selection to start a new search.

Performance: The filtering process should be quick and not cause any significant delay in the user experience.

Accessibility and usability: The feature should be accessible to users with disabilities and follow best practices for usability (ref: Directive (EU) 2016/2102).

User Help and Documentation: There should be clear documentation or help available explaining how to use the Category filter.



Data generated	N/A
Frequency use	High
Simultaneous conn.	High
RDF snippet	healthdcatap:healthCategory <http://healthdata.ec.europa.eu/.../(o)>;
Comments:	A HealthDCAT-AP Controlled Vocabulary must be used for this property



Use case 27 – Filter EU Datasets Catalogue by metadata values in “Health themes” list

User story name	<p>As an anonymous user, I want to be able to filter EU Datasets Catalogue by Health Themes in the HealthData@EU Central Platform. So that I can quickly find and access metadata relevant to a specific Health Theme.</p>
Role(s) concerned	All users
Necessary right(s)	N/A
Precondition scenario	<p>The HealthData@EU Central Platform contains metadata records with the relevant metadata property [healthdcatap:healthTheme, subproperty of dct:subject skos:Concept (Open data 0..n) (Protected data 0..n) (Sensitive data 0..n)]</p> <p>The user has access to the HealthData@EU Central Platform from his computer or mobile (an Internet connection is required).</p>
Scenario definition	<ol style="list-style-type: none"> 1) The user visits the EU Datasets Catalogue webpage of the HealthData@EU Central Platform 2) The user selects from the metadata record Health Theme filter one or several Health Themes 3) The result page displays the list of datasets from EU Datasets Catalogue corresponding to the selected Health Themes
Alternative scenarios	<p>If there are no values in “Health themes” list in EU Datasets Catalogue, then:</p> <p>The Filter Health Themes is not displayed.</p>



Acceptance criteria	<p>Health Themes selection feature: The HealthData@EU Central Platform must include an option to filter metadata sets by Health Themes. This feature should be easily accessible and intuitive to use.</p> <p>Comprehensive Health Themes List: The filter should include the comprehensive list of Health Themes.</p> <p>Multiple Health Themes Selection: A data set may be associated with multiple Health Themes. Users should have thus the ability to select multiple Health Themes at once.</p> <p>Seamless integration with other filters: This Health Themes filter should work seamlessly with other existing filters in the catalogue, such as HDABs, Categories etc., allowing for a dynamic refined search.</p> <p>Instant update of dataset list: Upon selecting a Health Theme or a combination of Themes, the list of datasets should update instantly to display only those that are relevant to the chosen Categories.</p> <p>Clear indication of selected Health Themes: The interface should clearly indicate which Health Theme or Themes have been selected at any given time.</p> <p>Reset option: There should be an option to easily reset or clear the Health Themes selection to start a new search.</p> <p>Performance: The filtering process should be quick and not cause any significant delay in the user experience.</p> <p>Accessibility and usability: The feature should be accessible to users with disabilities and follow best practices for usability (ref: Directive (EU) 2016/2102).</p> <p>User Help and Documentation: There should be clear documentation or help available explaining how to use the Health Themes filter.</p>
Data generated	N/A
Frequency use	High
Simultaneous conn.	High
RDF snippet	healthdcatap:healthTheme



	https://www.wikidata.org/wiki/Q58624061 > ,
	https://www.wikidata.org/wiki/Q7907952 > ;
Comments:	It is foreseen to use the Wikidata ontology for this property



Use case 28 – Filter EU Datasets Catalogue by metadata values in “Coding systems” list

User story name	<p>As an anonymous user, I want to be able to filter EU Datasets Catalogue by Coding Systems in the HealthData@EU Central Platform. So that I can quickly find and access metadata relevant to a specific Coding System for my project.</p>
Role(s) concerned	All users
Necessary right(s)	N/A
Precondition scenario	<p>The HealthData@EU Central Platform contains metadata records with the relevant metadata property [healthdcatap:hasCodingSystem dct:Standard (Open data 0..n) (Protected data 0..n) (Sensitive data 0..n)]</p> <p>The user has access to the HealthData@EU Central Platform from his computer or mobile (an Internet connection is required).</p>
Scenario definition	<ol style="list-style-type: none"> 1) The user visits the EU Datasets Catalogue webpage of the HealthData@EU Central Platform 2) The user selects from the metadata record Coding Systems filter one or several Coding Systems 3) The result page displays the list of datasets from EU Datasets Catalogue corresponding to the selected Coding Systems
Alternative scenarios	<p>If there are no values in “Coding Systems” list in EU Datasets Catalogue, then:</p> <p>The Filter Coding Systems is not displayed</p>



<p>Acceptance criteria</p>	<p>Coding Systems selection feature: The HealthData@EU Central Platform must include an option to filter metadata sets by Coding Systems. This feature should be easily accessible and intuitive to use.</p> <p>Comprehensive Coding Systems List: The filter should include the comprehensive list of Coding Systems.</p> <p>Multiple Coding Systems Selection: A metadata set may be associated with multiple Coding Systems. Users should have thus the ability to select multiple Coding Systems at once.</p> <p>Seamless integration with other filters: This Coding Systems filter should work seamlessly with other existing filters in the catalogue, such as HDABs, Categories etc., allowing for a dynamic refined search.</p> <p>Instant update of dataset list: Upon selecting a Coding System or a combination of Coding Systems, the list of metadata sets should update instantly to display only those that are relevant to the chosen Categories.</p> <p>Clear indication of selected Coding Systems: The interface should clearly indicate which Coding System or Systems have been selected at any given time.</p> <p>Reset option: There should be an option to easily reset or clear the Coding Systems selection to start a new search.</p> <p>Performance: The filtering process should be quick and not cause any significant delay in the user experience.</p> <p>Accessibility and usability: The feature should be accessible to users with disabilities and follow best practices for usability (ref: Directive (EU) 2016/2102).</p> <p>User Help and Documentation: There should be clear documentation or help available explaining how to use the Coding Systems filter.</p>
<p>Data generated</p>	<p>N/A</p>
<p>Frequency use</p>	<p>High</p>
<p>Simultaneous conn.</p>	<p>High</p>



RDF snippet	healthdcatap:hasCodingSystem < https://www.wikidata.org/wiki/Property:P1690 >, < https://www.wikidata.org/wiki/Property:P4229 >;
Comments:	It is foreseen to use the Wikidata HTTP URIs for this property.



Use case 29 – Filter EU Datasets Catalogue by metadata values in “Data models” list

User story name	<p>As an anonymous user, I want to be able to filter EU Datasets Catalogue by Data models in the HealthData@EU Central Platform. So that I can quickly find and access data relevant to a specific Data Model.</p>
Role(s) concerned	All users
Necessary right(s)	N/A
Precondition scenario	<p>The HealthData@EU Central Platform contains metadata records with the relevant metadata property [dct:conformsTo dct:Standard (Open data 0..n) (Protected data 0..n) (Sensitive data 0..n)]</p> <p>The user has access to the HealthData@EU Central Platform from his computer or mobile (an Internet connection is required).</p>
Scenario definition	<ol style="list-style-type: none"> 1) The user visits EU Datasets Catalogue webpage of the HealthData@EU Central Platform 2) The user selects from the metadata record Data Models filter one or several Data Models 3) The result page displays the list of datasets from EU Datasets Catalogue corresponding to the selected Data Models
Alternative scenarios	<p>If there are no values in “Data models” list in EU Datasets Catalogue, then:</p> <p>The Filter Data models is not displayed</p>



<p>Acceptance criteria</p>	<p>Data Models selection feature: The HealthData@EU Central Platform must include an option to filter metadata sets by Data Models. This feature should be easily accessible and intuitive to use.</p> <p>Comprehensive Data Models List: The filter should include the comprehensive list of Data Models To WP 6:</p> <p>Multiple Data Models Selection: A metadata set may be associated with multiple Data Models. Users should have thus the ability to select multiple Data Models at once.</p> <p>Seamless integration with other filters: This Data Models filter should work seamlessly with other existing filters in the catalogue, such as HDABs, Categories etc., allowing for a dynamic refined search.</p> <p>Instant update of dataset list: Upon selecting a Data Model or a combination of Models, the list of datasets should update instantly to display only those that are relevant to the chosen Data Models.</p> <p>Clear indication of selected Data Models: The interface should clearly indicate which Data Model or Models have been selected at any given time.</p> <p>Reset option: There should be an option to easily reset or clear the Data Models selection to start a new search.</p> <p>Performance: The filtering process should be quick and not cause any significant delay in the user experience.</p> <p>Accessibility and usability: The feature should be accessible to users with disabilities and follow best practices for usability (ref: Directive (EU) 2016/2102).</p> <p>User Help and Documentation: There should be clear documentation or help available explaining how to use the Data Models filter.</p>
<p>Data generated</p>	<p>N/A</p>
<p>Frequency use</p>	<p>High</p>
<p>Simultaneous conn.</p>	<p>High</p>
<p>RDF snippet</p>	<p>dct:conformsTo <https://www.wikidata.org/wiki/Q19597236>;</p>



Comments:	It is foreseen to use the Wikidata HTTP URIs for this property.
-----------	---



Use case 30 – Filter EU Datasets Catalogue by metadata values in “Age range” list

User story name	<p>As an anonymous user, I want to be able to filter EU Datasets Catalogue by Age Range in the HealthData@EU Central Platform. So that I can quickly find and access data relevant to a specific Age Range.</p>
Role(s) concerned	All users
Necessary right(s)	N/A
Precondition scenario	<p>The HealthData@EU Central Platform contains metadata records with the relevant metadata properties [healthdcatap:minTypicalAge rdfs:Integer (Open data 0..n) (Protected data 0..n) (Sensitive data 1..1)] and [healthdcatap:maxTypicalAge rdfs:Integer 1(Open data 0..n) (Protected data 0..n) (Sensitive data 1..1)]</p> <p>The user has access to the HealthData@EU Central Platform from his computer or mobile (an Internet connection is required).</p>
Scenario definition	<p>The user visits EU Datasets Catalogue webpage of the HealthData@EU Central The user specifies in metadata record Age Range filter a minimum and maximum age in years (as integers)</p> <p>The result page displays the list of datasets from EU Datasets Catalogue corresponding to the selected Age Range</p>
Alternative scenarios	<p>If there are no results matching the criteria specified by the user, then:</p> <p>The result page displays the message “Dataset not found”.</p>



Acceptance criteria	<p>Age Range selection feature: The HealthData@EU Central Platform must include an option to filter metadata sets by Age Range. This feature should be easily accessible and intuitive to use.</p> <p>Seamless integration with other filters: This Age Range filter should work seamlessly with other existing filters in the catalogue, such as HDABs, Categories etc., allowing for a dynamic refined search.</p> <p>Instant update of dataset list: Upon selecting an Age Range, the list of datasets should update instantly to display only those that are relevant to the chosen Age Range.</p> <p>Clear indication of selected Age Range: The interface should clearly indicate which Age Range has been selected at any given time.</p> <p>Reset option: There should be an option to easily reset or clear the Age Range selection to start a new search.</p> <p>Performance: The filtering process should be quick and not cause any significant delay in the user experience.</p> <p>Accessibility and usability: The feature should be accessible to users with disabilities and follow best practices for usability (ref: Directive (EU) 2016/2102).</p> <p>User Help and Documentation: There should be clear documentation or help available explaining how to use the Age Range filter.</p>
Data generated	N/A
Frequency use	High
Simultaneous conn.	High
RDF snippet	<pre>healthdcat:minTypicalAge: "18"; healthdcat:maxTypicalAge: "64";</pre>



Comments:	The cardinality of the properties will be validated after the implementation phase.
-----------	---



Use case 31 Filter EU Datasets Catalogue by metadata values in “Population Size” list

User story name	<p>As an anonymous user,</p> <p>I want to be able to filter EU Datasets Catalogue by Population Size in the HealthData@EU Central Platform.</p> <p>So that I can quickly find and access metadata relevant to a specific Population Size.</p>
Role(s) concerned	All users
Necessary right(s)	N/A
Precondition scenario	<p>The HealthData@EU Central Platform contains metadata records with the relevant metadata properties [healthdcatap:numberofUniqueIndividuals rdfs:Integer (Open data 0..n) (Protected data 0..n) (Sensitive data 1..1)]</p> <p>The user has access to the HealthData@EU Central Platform from his computer or mobile (an Internet connection is required).</p>
Scenario definition	<ol style="list-style-type: none"> 1) The user visits EU Datasets Catalogue webpage of the HealthData@EU Central Platform 2) The user selects in metadata record Population Size filter a numerical value (as an integer) and specifies whether the Population Size should be greater than (>) or less than (<) this value. 3) The result page displays the list of datasets from EU Datasets Catalogue corresponding to the selected Population Size



Alternative scenarios	<p>If there are no results matching the criteria specified by user, then:</p> <p>The result page displays the message "Dataset not found".</p>
Acceptance criteria	<p>Population Size selection feature: The HealthData@EU Central Platform must include an option to filter metadata sets by Population Size. This feature should be easily accessible and intuitive to use.</p> <p>Seamless integration with other filters: This Population Size filter should work seamlessly with other existing filters in the catalogue, such as HDABs, Categories etc., allowing for a dynamic refined search.</p> <p>Instant update of dataset list: Upon selecting a Population Size criteria, the list of metadata sets should update instantly to display only those that are relevant to the chosen Population Size criteria.</p> <p>Clear indication of selected Population Size criteria: The interface should clearly indicate which Population Size criteria has been selected at any given time.</p> <p>Reset option: There should be an option to easily reset or clear the Population Size criteria to start a new search.</p> <p>Performance: The filtering process should be quick and not cause any significant delay in the user experience.</p> <p>Accessibility and usability: The feature should be accessible to users with disabilities and follow best practices for usability (ref: Directive (EU) 2016/2102).</p> <p>User Help and Documentation: There should be clear documentation or help available explaining how to use the Population Size filter.</p>
Data generated	N/A
Frequency use	High
Simultaneous conn.	High



RDF snippet	healthdcatap:numberOfUniqueIndividuals: "8914722";
-------------	--



Use case 32 –Filter EU Datasets Catalogue by metadata values in "Retention Period" list

User story name	<p>As an anonymous user,</p> <p>I want to be able to filter EU Datasets Catalogue by Retention Period in the HealthData@EU Central Platform.</p> <p>So that I can quickly find and access metadata relevant to a specific Time Period.</p>
Role(s) concerned	All users
Necessary right(s)	N/A
Precondition scenario	<p>The HealthData@EU Central Platform contains metadata records with the relevant metadata properties [dct:temporal dct:PeriodOfTime (Open data 0..n) (Protected data 0..n) (Sensitive data 0..n)]</p> <p>The user has access to the HealthData@EU Central Platform from his computer or mobile (an Internet connection is required).</p>



<p>Scenario definition</p>	<ol style="list-style-type: none"> 1) The user visits EU Datasets Catalogue webpage of the HealthData@EU Central 2) The user selects in metadata record Time Period filter a start date and an end date, or ongoing for metadata sets with continuous data collection. <p>The result page displays the list of datasets from EU Datasets Catalogue corresponding to the selected Time Period (that either fall within the selected Time Period or encompass the start or end of this period).</p> <ol style="list-style-type: none"> 3) Conditional upon the use of the Time Period filter, the 'Sort by' feature provides new options, enabling users to refine their search based on specific date ranges tailored to their research criteria. the 'Sort by' feature activates, offering three options for further refinement: <ul style="list-style-type: none"> o Start Date Range: Focuses on metadata sets that begin within the selected time frame. o End Date Range: Highlights metadata sets that conclude within the specified period. o Complete Overlap: Filters for metadata sets that entirely span the selected start and end dates."
<p>Alternative scenarios</p>	<p>If there are no results matching the criteria specified by user, then:</p> <p>The result page displays the message "Dataset not found".</p>



<p>Acceptance criteria</p>	<p>Time Period selection feature: The HealthData@EU Central Platform must include an option to filter metadata sets by Time Period. This feature should be easily accessible and intuitive to use.</p> <p>Seamless integration with other filters: This Time Period filter should work seamlessly with other existing filters in the catalogue, such as HDABs, Categories etc., allowing for a dynamic refined search.</p> <p>Instant update of dataset list: Upon selecting a Time Period, the list of metadata sets should update instantly to display only those that are relevant to the chosen Time Period.</p> <p>Clear indication of selected Time Period: The interface should clearly indicate which Time Period has been selected at any given time.</p> <p>Reset option: There should be an option to easily reset or clear the Time Period selection to start a new search.</p> <p>Performance: The filtering process should be quick and not cause any significant delay in the user experience.</p> <p>Accessibility and usability: The feature should be accessible to users with disabilities and follow best practices for usability (ref: Directive (EU) 2016/2102).</p> <p>User Help and Documentation: There should be clear documentation or help available explaining how to use the Time Period filter.</p>
<p>Data generated</p>	<p>N/A</p>
<p>Frequency use</p>	<p>High</p>
<p>Simultaneous conn.</p>	<p>High</p>



RDF snippet	<pre>dct:temporal [a dct:PeriodOfTime; dcat:endDate "2024-12- 31"^^<http://www.w3.org/2001/XMLSchema#date>; dcat:startDate "2021-01- 01"^^<http://www.w3.org/2001/XMLSchema#date>];</pre>
-------------	---



Use case 33 – Filter EU Datasets Catalogue by metadata values in “Update Frequency” list

User story name	<p>As an anonymous user,</p> <p>I want to be able to filter EU Datasets Catalogue by Update Frequency in the HealthData@EU Central Platform.</p> <p>So that I can quickly find and access metadata relevant to a specific Update Frequency.</p>
Role(s) concerned	All users
Necessary right(s)	N/A
Precondition scenario	<p>The HealthData@EU Central Platform contains metadata records with the relevant metadata properties [dct:accrualPeriodicity foaf:Document (Open data 0..1) (Protected data 0..1) (Sensitive data 0..1)]</p> <p>The user has access to the HealthData@EU Central Platform from his computer or mobile (an Internet connection is required).</p>
Scenario definition	<ol style="list-style-type: none"> 1) The user visits EU Datasets Catalogue webpage of the HealthData@EU Central Platform 2) The user selects from the metadata Update Frequency filter one value 3) The result page displays the list of datasets from EU Datasets
	Catalogue corresponding to the selected Update Frequency
Alternative scenarios	<p>If there are no results matching the criteria specified by user, then:</p> <p>The result page displays the message “Dataset not found”.</p>



Acceptance criteria	<p>Update frequency selection feature: The HealthData@EU Central Platform must include an option to filter metadata sets by HDABs. This feature should be easily accessible and intuitive to use.</p> <p>Complete Update Frequency List: The filter will feature a thorough list of Update Frequency options. All available values correspond to those in the controlled vocabulary of the Official Publications of the European Union (OP) [http://publications.europa.eu/resource/authority/frequency/].</p> <p>Multiple Update Frequency Selection: Users should have the ability to select multiple Update Frequency options at once.</p> <p>Seamless integration with other filters: This Update Frequency filter should work seamlessly with other existing filters in the catalogue, such as categories, health themes, etc., allowing for a dynamic refined search.</p> <p>Instant update of dataset list: Upon selecting an Update Frequency or a combination of Update Frequencies, the list of metadata sets should update instantly to display only those that are relevant to the chosen Update Frequencies.</p> <p>Clear indication of selected Update Frequencies: The interface should clearly indicate which HDAB or HDABs have been selected at any given time.</p> <p>Reset option: There should be an option to easily reset or clear the Update Frequencies selection to start a new search.</p> <p>Performance: The filtering process should be quick and not cause any significant delay in the user experience.</p> <p>Accessibility and usability: The feature should be accessible to users with disabilities and follow best practices for usability (ref: Directive (EU) 2016/2102).</p> <p>User Help and Documentation: There should be clear documentation or help available explaining how to use the Update Frequency filter.</p>
Data generated	N/A
Frequency use	High
Simultaneous conn.	High



RDF snippet	<pre>dct:accrualPeriodicity <http://publications.europa.eu/resource/authority/frequency/DAILY>, <http://publications.europa.eu/resource/authority/frequency/UPDATE_CONT>;</pre>
-------------	---



Use case 34 – Filter EU Datasets Catalogue by metadata values in “dataset type” list

User story name	<p>As an anonymous user of HealthData@EU Central Platform, I want to be able to filter EU Datasets Catalogue by selecting a specific dataset type such as Personal Data in the HealthData@EU Central Platform. So that I can quickly find and access metadata with detailed, individual-level insights for my research project (e.g.: potentially linkable datasets).</p>
Role(s) concerned	All users
Necessary right(s)	N/A
Precondition scenario	<p>The HealthData@EU Central Platform contains metadata records with the relevant metadata property [dct:type skos:Concept (Open data 0..1) (Protected data 0..1) (Sensitive data 1..1)]</p> <p>The user has access to the HealthData@EU Central Platform from his computer or mobile (an Internet connection is required).</p>
Scenario definition	<p>The user visits EU Datasets Catalogue webpage of the HealthData@EU Central. The user selects from the metadata record Dataset Type’s list one value.</p> <p>The result page displays the list of datasets from EU Datasets Catalogue corresponding to the selected filter</p>
Alternative scenarios	



<p>Acceptance criteria</p>	<p>Dataset Type selection feature: The HealthData@EU Central Platform must include an option to filter metadata sets by Dataset Type. This feature should be easily accessible and intuitive to use.</p> <p>This Dataset Type will utilise the 'dct:type' property associated with the OP controlled vocabulary "data-type".</p> <p>Seamless integration with other filters: This Dataset Type filter should work seamlessly with other existing filters in the catalogue, such as HDABs, health themes, etc., allowing for a dynamic refined search.</p> <p>Instant update of dataset list: Upon selecting the filter Dataset Type, the list of metadata sets should update instantly to display only those that are relevant to the activated filter.</p> <p>Clear indication of the filter Dataset Type: The interface should clearly indicate that the filter Dataset Type has been selected at any given time.</p> <p>Reset option: There should be an option to easily reset or clear the Dataset Type selection to start a new search.</p> <p>Performance: The filtering process should be quick and not cause any significant delay in the user experience.</p> <p>Accessibility and usability: The feature should be accessible to users with disabilities and follow best practices for usability (ref: Directive (EU) 2016/2102).</p> <p>User Help and Documentation: There should be clear documentation or help available explaining how to use the Dataset Type filter.</p>
<p>Data generated</p>	<p>N/A</p>
<p>Frequency use</p>	<p>High</p>
<p>Simultaneous conn.</p>	<p>High</p>



RDF snippet	<pre>dct:type <http://publications.europa.eu/resource/authority/dataset-type/PERSONAL_DATA>;</pre>
Comments:	<p>A new category, "Personal Data," must be introduced (see: http://publications.europa.eu/resource/authority/dataset-type/PERSONAL_DATA) allowing the selection of datasets containing Personal Data Only.</p> <p>Health catalogue implementers can choose to implement a Dataset Type filter based on the value Personal_Data: Yes or No (see Annex 3).</p>



Use case 35 – Filter EU Datasets Catalogue by metadata values in “Personal data” list

User story name	As an anonymous user of HealthData@EU Central Platform, I want to be able to filter EU Datasets Catalogue by those containing Personal Data in the HealthData@EU Central Platform. So that I can quickly find and access metadata with detailed, individual-level insights for my research project (e.g.: potentially linkable datasets).
Role(s) concerned	All users
Necessary right(s)	N/A
Precondition scenario	The HealthData@EU Central Platform contains metadata records with the relevant metadata property [dpv:hasPersonalData dpv:PersonalData (Open data 0..n) (Protected data 0..n) (Sensitive data 0..n)] The user has access to the HealthData@EU Central Platform from his computer or mobile (an Internet connection is required).
Scenario definition	The user visits EU Datasets Catalogue webpage of the HealthData@EU Central The user selects from the metadata record Personal data filter one or several values of the Personal Data list. The result page displays the list of datasets from EU Datasets Catalogue corresponding to the selected values for the filter
Alternative scenarios	If there are no values in “Personal Data” list in EU Datasets Catalogue, then: The Filter Personal data is not displayed



<p>Acceptance criteria</p>	<p>Personal Data selection feature: The HealthData@EU Central Platform must include an option to filter metadata sets containing Personal Data. This feature should be easily accessible and intuitive to use.</p> <p>Seamless integration with other filters: This Personal Data filter should work seamlessly with other existing filters in the catalogue, such as HDABs, health themes, etc., allowing for a dynamic refined search.</p> <p>Instant update of dataset list: Upon selecting the filter Personal Data Only, the list of metadata sets should update instantly to display only those that are relevant to the activated filter.</p> <p>Clear indication of the filter Personal Data Only: The interface should clearly indicate that the filter Personal Data Only has been selected at any given time.</p> <p>Reset option: There should be an option to easily reset or clear the Personal Filter Only selection to start a new search.</p> <p>Performance: The filtering process should be quick and not cause any significant delay in the user experience.</p> <p>Accessibility and usability: The feature should be accessible to users with disabilities and follow best practices for usability (ref: Directive (EU) 2016/2102).</p> <p>User Help and Documentation: There should be clear documentation or help available explaining how to use the Personal Data Only filter.</p>
<p>Data generated</p>	<p>N/A</p>
<p>Frequency use</p>	<p>High</p>
<p>Simultaneous conn.</p>	<p>High</p>



RDF snippet	<pre>dpv:hasPersonalData dpv-pd:Gender, dpv-pd:Age, dpv-pd:Location, dpv-pd:Nationality, dpv-pd:Education, dpv-pd:HealthRecord; >;</pre>
Comments:	



Use case 36 – Filter EU Datasets Catalogue by metadata value “Trusted Data Holder”

User story name	<p>As an anonymous user of HealthData@EU Central Platform, I want to be able to filter EU Datasets Catalogue by Trusted Data Holders, So that I can quickly find datasets that have simplified data permit procedures.</p>
Role(s) concerned	All users
Necessary right(s)	N/A
Precondition scenario	<p>The HealthData@EU Central Platform contains metadata records with the relevant metadata property [healthdcatap:trustedDataHolder xsd:boolean (Open data 0..1) (Protected data 0..1) (Sensitive data 0..1)]</p> <p>The user has access to the HealthData@EU Central Platform from his computer or mobile (an Internet connection is required).</p>
Scenario definition	<p>The user visits EU Datasets Catalogue webpage of the HealthData@EU Central. The user selects from the metadata record Trusted Data Holder filter “Yes” or “No</p> <p>The result page displays the list of datasets from EU Datasets Catalogue corresponding to the selected filter</p>
Alternative scenarios	<p>2) If there are “Trusted Data Holders” in EU Datasets Catalogue, then:</p> <p>The Filter Trusted Data Holder is not displayed</p>



<p>Acceptance criteria</p>	<p>Trusted Data Holder Only selection feature: The HealthData@EU Central Platform must include an option to filter metadata sets containing Publishers with the status “Trusted Data Holder”. This feature should be easily accessible and intuitive to use. This Trusted Data Holder Only filter will utilise the ‘healthdcatap:trustedDataholder’ property (True or False).</p> <p>Seamless integration with other filters: This Trusted Data Holder filter should work seamlessly with other existing filters in the catalogue, such as HDABs, health themes, etc., allowing for a dynamic refined search.</p> <p>Instant update of dataset list: Upon selecting the filter Trusted Data Holder Only, the list of metadata sets should update instantly to display only those that are relevant to the activated filter.</p> <p>Clear indication of the filter Trusted Data Holder Only: The interface should clearly indicate that the filter Trusted Data Holder Only has been selected at any given time.</p> <p>Reset option: There should be an option to easily reset or clear the Trusted Data Holder Only selection to start a new search.</p> <p>Performance: The filtering process should be quick and not cause any significant delay in the user experience.</p> <p>Accessibility and usability: The feature should be accessible to users with disabilities and follow best practices for usability (ref: Directive (EU) 2016/2102).</p> <p>User Help and Documentation: There should be clear documentation or help available explaining how to use the Trusted Data Holder Only filter.</p>
<p>Data generated</p>	<p>N/A</p>
<p>Frequency use</p>	<p>High</p>
<p>Simultaneous conn.</p>	<p>High</p>



RDF snippet	<code>healthdcatap:trustedDataHolder true;</code>
Comments:	



Use case 37 – SPARQL endpoint

User story name	<p>As an anonymous and advanced user of the HealthData@EU Central Platform, I want to access a SPARQL endpoint for the HealthData@EU Central Platform so that I can run complex queries over the dataset metadata records and extract specific, structured information according to my project’s needs.</p>
Role(s) concerned	All users
Necessary right(s)	N/A
Precondition scenario	<p>Healthdata.europa.eu infrastructure: The HealthData@EU Central Platform is populated with datasets that are formatted in RDF (Resource Description Framework), making them suitable for SPARQL queries.</p> <p>Metadata record accessibility: The RDF metadata records have been indexed in such a way that they can be queried via a SPARQL endpoint.</p> <p>User permissions: The anonymous user has the necessary permissions or access rights to query the HealthData@EU Central Platform.</p> <p>SPARQL compliance: The HealthData@EU Central Platform supports SPARQL protocol and can handle HTTP requests for SPARQL queries.</p> <p>System performance: The system hosting the HealthData@EU Central Platform has the capacity to process SPARQL queries without significant performance degradation.</p> <p>Query environment: A user interface or a direct network access point is available for users to submit their SPARQL queries to the endpoint.</p> <p>Documentation and support: Adequate documentation is provided for users on how to construct SPARQL queries for the HealthData@EU Central Platform, including examples and guidance on accessing the SPARQL endpoint.</p>



Scenario definition	<ol style="list-style-type: none"> 1) The user accesses the SPARQL interface on the HealthData@EU Central Platform to prepare a query. 2) The SPARQL endpoint processes the input query against the HealthData@EU Central platform database 3) the HealthData@EU Central platform database returns the results of the query to the interface. 4) The results of the query are presented to the user in a human-readable format.
Alternative scenarios	<ol style="list-style-type: none"> 1) The user visits the SPARQL interface of the HealthData@EU Central Platform 2) The SPARQL endpoint processes the input query against the HealthData@EU Central platform database 3) the HealthData@EU Central platform database is unable to find a match for the query and returns to the interface an error.
	4) A "No data found" message is displayed to the user.
Acceptance criteria	<p>Given thar the SPARQL endpoint is correctly set up to interface with the HealthData@EU Central Platform, When a user inputs a valid SPARQL query into the interface, Then the SPARQL endpoint should process the query against the HealthData@EU Central Platform and should return the results of the query to the user in a human-readable format.</p>
Data generated	NA
Frequency use	Low
Simultaneous conn.	Low
RDF snippet	N/A
Comments:	



Questions:	
------------	--



3. Annexes:

Annex 1 Controlled vocabularies for the specific use of the HealthDCAT-AP

Controlled vocabularies are predefined, authorised lists of terms that provide a consistent way to describe data, ensuring that everyone is using the same term to mean the same thing. This is especially important in data catalogues to facilitate search, discovery, and interoperability.

The ability to accurately categorise, tag, and describe datasets using a standardised vocabulary directly impacts the core functionalities of the catalogue.

Consider the 'Publisher Type' controlled vocabulary utilised in the HealthDCAT-Application Profile:

As a data holder generating a HealthDCAT-AP metadata record, **I want to** identify and select the Publisher Type that best represents my organisation. **So that** it facilitates the accurate and efficient annotation of my dataset.

Examples of publisher types include tags such as 'national public health institute,' 'cancer registry,' and others.

Controlled vocabularies used in DCAT-AP and Health DCAT-AP, maintained by the Publications Office:

DCAT-AP property	Name space of the OP controlled vocabulary in use
dct:spatial	http://publications.europa.eu/resource/authority/country/
dcat:theme	http://publications.europa.eu/resource/authority/data-theme/
dct:accessRights	http://publications.europa.eu/resource/authority/access-right/
dct:accrualPeriodicity	http://publications.europa.eu/resource/authority/frequency/
dct:language	http://publications.europa.eu/resource/authority/language/



dct:type (Dataset)	http://publications.europa.eu/resource/authority/dataset-type/(PERSONAL_DATA)*
dct:type (Distribution)	http://publications.europa.eu/resource/authority/file-type/

* The "dataset-type" controlled vocabulary requires a new "personal_data" entry.

Controlled vocabularies to develop as SKOS thesauri specifically for use with the Health DCAT-AP:

HealthDCAT-AP property	
healthdcatap:healthCategory [TWG session n°2b]	Based on EHDS Regulation art.33 (See example of High-value dataset categories)
healthdcatap:publisherType [TWG session n°2]	Registry, Biobank, Public Health Institute, Hospital, etc.
prov:qualifiedAttribution >dcat:hadRole [TWG session n°3 & 4]	Ex: http://inspire.ec.europa.eu/metadata-codelist/ResponsiblePartyRole/ (processor)
dcat:qualifiedRelation >dcat:hadRole [TWG session n°3 & 4]	Ex: http://www.iana.org/assignments/relation/ (related)
prov:wasGeneratedBy >dct:type [TWG session n°9]	Ex: http://dbpedia.org/resource/Record_linkage

Health DCAT-AP properties relying on the Wikidata knowledge base that is readily SPARQL-quarriable:

HealthDCAT-AP property	
healthdcatap:healthTheme [TWG session n°3 & 4]	Wikidata (" Wikidata: A large-scale collaborative ontological medical database ") (See example of MobilityDCAT-AP)



healthdcatap:hasCodeValues [TWG session n°4b]	Create a SKOS-based structure to represent relevant parts of the thesauri: LOINC, SNOMED-CT, ICD-10, FASTA, etc.
dct:conformsTo [TWG session n°3 & 4]	Wikidata (See: https://medical-data-models.org)



Annex 2: RDF HealthDCAT-AP template

The draft version of the HealthDCAT-AP has been published on a GitHub repository which can be accessed at the following url address: <https://healthdcat-ap.github.io>. The HealthDCAT-AP specification adheres to the W3C ReSpec documentation style, a framework designed for writing technical documents (<https://respec.org/docs/>). It serves as the single source of trust for health datasets catalogue implementers.

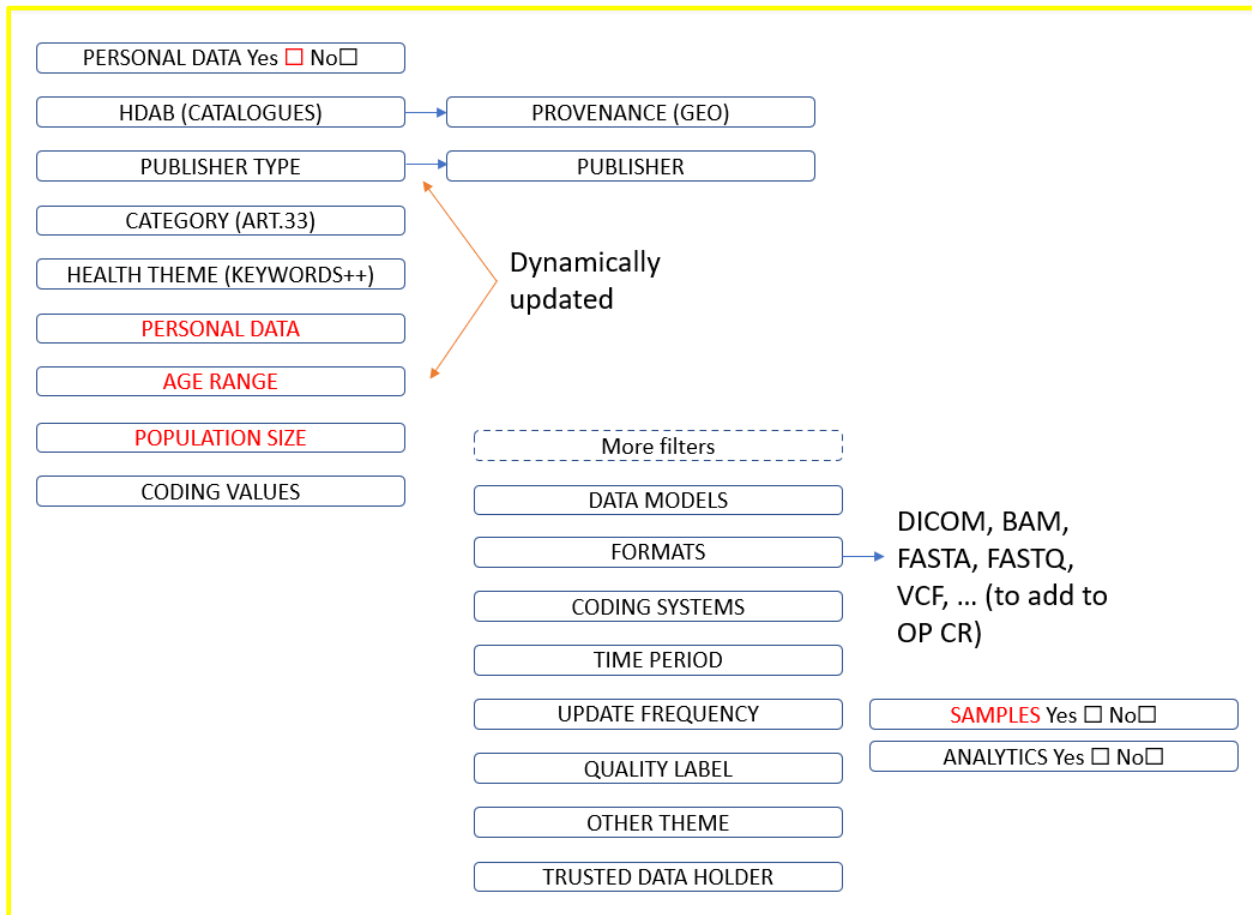
Examples of RDF snippet can be found from the healthDCAT-AP specification: <https://healthdcat-ap.github.io/#rdf-examples>



Annex 3: Filtering capabilities of a HealthDCAT-AP catalogue

HealthDCAT-AP, with its extensive set of properties, accommodates various data types. Some properties may not be applicable depending on the specific data type. Consequently, a data catalogue dynamically displays only the relevant information in a dataset's metadata record. This dynamic capability can also be applied to the catalogue's filters. Given that a health data catalogue (Art. 33) includes information about both personal and non-personal electronic health data, the available filters can dynamically adjust based on the user's search selections.

Refer to the highlighted red section in the diagram below: the dataset-type property (dct:type) is used as a filter. Selecting the [Yes] option will include all datasets marked as containing 'personal data.' If [Yes] is chosen, additional filters can be displayed to further refine datasets with personal-level information. It is up to the catalog implementers to determine how best to enhance the user experience with dynamic filter options.



Optimised Filters for HealthDCAT-AP Dataset Catalogues: Dynamic Display Based on Personal Data Inclusion