

Annex 4b: The HealthDCAT-AP mapping tables to the national health data catalogues and pilot descriptions

Denmark

HealthDCAT-AP Dataset property name	Property description in implementing act Art. 77 draft	HealthDCAT-AP Dataset property definition (R7)	HealthDCAT-AP Dataset usage note (R7)	NHdC property name	NHdC property definition	Additional information	Example description cancer registry	Example description on biobank data
Title	Provides the official name or title of the dataset.	A name given to the Dataset.	Provide a title(s) for your Dataset, which can be repeated in multiple languages. Example: Healthy Brain Study - Physiological Data RDF example: dct:title	Titel	Datasættets læsbare navn. Obs: der er mulighed for at angive titel i flere sprog.	Property name and definition are only provided in Danish as the Danish metadata catalogue has not yet been implemented/translated to English. An English version of the catalogue will not be prioritized before 2028.	Cancerregisteret	Danish National Genome Database
Access rights	Describes the level of access applicable to the dataset (e.g. public, restricted or non public).	Information on that indicates whether the Dataset is publicly accessible, has public access or is restricted.	This property allows you to indicate whether the dataset is publicly accessible, has restrictions, or is not publicly available. This is a mandatory property in HealthDCAT-AP. Use one of the following values :public, :restricted,	Adgangsbegretninger	Datasættets følsomhed. Der angives enten non-public (personlige sundhedsdata), restricted (ikke personlige sundhedsdata, yet been men ikke offentligt tilgængeligt), public (data er offentligt tilgængeligt).	Property name and definition are only provided in Danish as the Danish metadata catalogue has not yet been implemented/translated to English. An English version of the catalogue will not be prioritized before 2028.	NON_PUBLIC	NON_PUBLIC

not public. :non-public, which means:
 NON_PUBLIC — The dataset is not publicly available and access is subject to restrictions.
 PUBLIC — The dataset is openly accessible to everyone.
 RESTRICTED — The dataset has access restrictions but may be available under specific conditions.
 RDF example: dct:accessRights

Application	Identifies U	The	The ELI of the EHDS	Lovgivning	Referer til gældende lovgivning, der påvirker oprettelse og deling af datasættet. I dansk kontekst bør der refereres til Sundhedsloven (og medfølgende bekendtgørelse), Databeskyttelsesloven, EHDS og/eller GDPR.	Property name and definition are only provided in Danish as the Danish metadata catalogue has not yet been implemented/translated to English. An English version of the catalogue will not be prioritized before 2028.	https://www.retsinformation.dk/eli/Ita/2018/1049	http://data.europa.eu/eli/reg/2025/327/oj , https://www.retsinformation.dk/eli/Ita/2018/728
	nion legal acts or national legislation that specifically govern the establishment of the dataset, where such legal acts or legislation exists.	n that mandate s the creation or manage ment, maintenance or availability of the dataset,	Regulation, published in March 2025, can now be used as the applicable legislation where relevant: http://data.europa.eu/eli/reg/2025/327/oj . As multiple legislations may apply to the resource the maximum cardinality is not limited. RDF example: dcatap:applicableLegislation					

Code values	Identifies classification systems or controlled vocabularies reflected in the content of the dataset (e.g. diseases, e thesaurus, classification, coding procedure classification), where applicable.	A free-text description of any coding systems used that apply to the dataset (e.g. U07.1, Y59.0). No controlled terminology is required at this stage.	Use this field to describe, in free text, the code values of the coding systems used for the dataset (e.g. healthdcats, p:hasCodeValues).
--------------------	---	--	---

Code values is not a mandatory property in HealthDCAT-AP release 6 (or release 7) and is not included in the Danish metadata catalogue. If the property becomes mandatory in a later release of HealthDCAT-AP it will be implemented in the Danish data model. Even though not all metadata properties will be implemented and published in the Danish metadata catalogue, Danish data holders will at all times be able to send every property of the HealthDCAT-AP to the Danish metadata catalogue database and from there the metadata will be sent to the HealthData@EU central platform and be published in the EU dataset catalogue.

Coding system	Identifies standardised coding schemes used to encode data elements within the dataset (e.g. ICD 10 CM, SNOMED CT, DRGs), where applicable.	Coding systems in use (ex: ICD-10-CM, DGRs, SNOME D=CT).	This property allows you to specify the coding systems used within your dataset. For example, if a dataset uses ICD-10 for disease classification, this property allows data users to search for datasets with the same coding system. As a machine-actionable property, it also facilitates automated processes, making dataset discovery more efficient. Specify the standardised coding systems used in the dataset (e.g. ICD-10-CM, SNOMED-CT). The coding systems controlled vocabulary (https://hdeu-dcat.acceptance.data.health.europa.eu/resource/authority/coding-system/) MUST be used. RDF example: health dcatap:hasCodingSystem	—	Coding system' is not a mandatory property in HealthDCAT-AP release 6 (or release 7) and is not included in the Danish metadata catalogue. If the property becomes mandatory in a later release of HealthDCAT-AP it will be implemented in the Danish data model. Even though not all metadata properties will be implemented and published in the Danish metadata catalogue, Danish data holders will at all times be able to send every property of the HealthDCAT-AP to the Danish metadata catalogue database and from there the metadata will be sent to the HealthData@EU central platform and be published in the EU dataset catalogue.	—	
----------------------	---	--	--	---	--	---	--

Contact point	Provides contact details for enquiries related to the dataset.	Contact information that can be used for sending comments about the Dataset.	This property informs about a contact point (Individual, Organization, Location, Group) that can answer questions about the dataset. Details on how to describe these are provided under class vcard:Kind. RDF example: dcat:contactPoint	URL Email	Webside til kontakt - fx side hvor man kan skrive spørgsmål eller lign. Email til kontakt.	Contact point is not a property like most others. Instead it contains two underlying properties. In the Danish metadata catalogue the two underlying and mandatory properties 'URL' and 'Email' has been implemented. Property names and definition are only provided in Danish as the Danish metadata catalogue has not yet been implemented/translated to English. An English version of the catalogue will not be prioritized before 2028.	https://sundhedsdatastyrelsen.dk	https://www.ngc.dk/
Conforms to	Identifies a standard or technical specification that defines the structural or syntactic organisation of the dataset as a whole.	An established standard to which the resource conforms.	If your data conforms to an established standard or specification, use this property to indicate which one. The Health Standards controlled vocabulary (https://hdeu-f.conformsdcat.acceptance.data.health.europa.eu/resource/authority/standard/) MUST be used. RDF example: dct:conformsTo	—	—	Conforms to' is not a mandatory property in HealthDCAT-AP release 6 (or release 7) and is not included in the Danish metadata catalogue. If the property becomes mandatory in a later release of HealthDCAT-AP it will be implemented in the Danish data model. Even though not all metadata properties will be implemented and published in the Danish metadata catalogue, Danish data holders will at all times be able to send every property of	—	—

				the HealthDCAT-AP to the Danish metadata catalogue database and from there the metadata will be send to the HealthData@EU central platform and be published in the EU dataset catalogue.		
Custodian	Identifies the Party responsible for the dataset. The entity responsible for the dataset is the one that holds the data and maintains the dataset. In the EHDS context, this field represents the legally defined health data holder, (Article 2(6) EHDS Regulation), i.e. the entity responsible for care and holding and maintaining the dataset. This entity may differ from the resource publisher.	Use this property to identify the entity that technically holds and maintains the dataset. In the EHDS context, this field represents the legally defined health data holder, (Article 2(6) EHDS Regulation), i.e. the entity responsible for care and holding and maintaining the dataset. This entity may differ from the resource publisher.	—	Custodian' is not a mandatory property in HealthDCAT-AP release 6 (or release 7) and is not included in the Danish metadata catalogue. If the property becomes mandatory in a later release of HealthDCAT-AP it will be implemented in the Danish data model. Even though not all metadata properties will be implemented and published in the Danish metadata catalogue, Danish data holders will at all times be able to send every property of the HealthDCAT-AP to the Danish metadata catalogue database and from there the metadata will be send to the HealthData@EU central platform and be published in the EU dataset catalogue.	—	—

Description	Provides a free text summary describing the content and purpose of the dataset.	A free-text information about the Dataset.	Brief description of the dataset. This property can be repeated for including several language versions of the description. RDF example: dct:description	Description	Beskrivelse af datasættet, der er forståelig og informativ for forskere og andre databruere. Herunder bør særlige opmærksomhedspunkter for datasættet beskrives fx databrud eller andre forhold, påvirker pålidelighed, usikkerhed, kompletthed og sammenlignelighed, der kan være vigtige for en bruger.	Property name and definition are only provided in Danish as the Danish metadata catalogue has not yet been implemented/translated to English. An English version of the catalogue will not be prioritized before 2028.	Cancerregisteret indeholder oplysninger om danske cancerpatienter. \nRegisteret er et incidensregister, dvs. at det alene indeholder oplysninger om nye tilfælde af cancer. Cancerregisteret rummer bl.a. oplysninger om: diagnosetidspunktet for sygdommen, den anatomiske lokalisering, hvilken type kræft det er (histologi), udbredelse af sygdommen og hvordan diagnosen er stillet. \nIndberetning til Cancerregisteret skete på frivillig basis frem til marts 1987, hvorefter det blev obligatorisk. Fra 2004 overgik Cancerregisteret fra indberetning på papirblanket til elektroniske indberetning via Landspatientregisteret (LPR) for offentlige hospitaler og via Sundhedsdatastyrelsens Elektroniske Indberetningssystem (SEI) for de praktiserende læger og private hospitaler. Samtidig overgik man i 2004 fra manuel kodning til det automatiske cancerlogik kodningssystem. Registeret har siden 2004 været kendt som det moderniserede Cancerregister i takt med implementeringen af digitale løsninger	Database of 38,000+ patients treated in the Danish healthcare system from 2019-(ongoing). Data includes raw sequencing data (FASTQ for all patients; CRAM and VCF files for most)
--------------------	---	--	---	--------------------	---	--	---	---

	wide), where applicable.		s MAY be used. RDF example: dct:sp atial			
Personal data / Has personal data	Indicates whether the dataset contains personal data	Key	This property is used to indicate the types of personal data contained in the dataset itself (i.e. the actual data, not the meaning of dataset, I metadata record). Article 4, indicates the Data Privacy Regulation point (1), of association with Personal Data at (EU) 2016/679 of the European Parliament and of the Council ² .	—	Has personal data' is not a mandatory property in HealthDCAT-AP release 6 (or release 7) and is not included in the Danish metadata catalogue. If the property becomes mandatory in a later release of HealthDCAT-AP it will be implemented in the Danish data model. Even though not all metadata properties will be implemented and published in the Danish metadata catalogue, Danish data holders will at all times be able to send every property of the HealthDCAT-AP to the Danish metadata catalogue database and from there the metadata will be send to the HealthData@EU central platform and be published in the EU dataset catalogue.	—
Health data access body	The competent health data access body responsible for facilitating access to the	Health Data Access Body supportinelectronic g access to data in the	Member state-designated authority that facilitates the secondary use of HDABs assess the information provided by the health data	HDAB	In Denmark we only have one HDAB. The name of the HDAB is not published in the Danish metadata catalogue but part of the Danish data model and the HDAB name is automatically set on	Sundhedsdatastyrelsen Sundhedsdatastyrelsen

(EU) ion 101-E0035) of the
2025/327. Regulati EHDS Regulation.
on on theUse the controlled
Europea vocabulary
n Health http://13.81.34.152:11
Data 01/resource/authority/
Space healthcategories/ to
laying ensure standardised
down a classification across
list of all datasets
categorieRDF
s of example: healthdcat
electroni p:healthCategory
c data for
secondar
y use,
Art.51.

Health theme	Indicates a high level health topic or theme associated with the dataset (e.g. cardiovascular disease, cancer).	A category of the Dataset or tag describing the Dataset. (http://13.81.34.152:101/resource/authority/health-theme) MUST be used. RDF example: healthdcatap:healthTheme	Sundhedstemaer	Liste af sundhedstemaer på URI form. Skal vælges fra wikidat a: http://13.81.34.152:1101/resource/authority/health-theme, fx "ENVIRONMENTAL_HEALTH"	Property name and definition are only provided in Danish as the Danish metadata catalogue has not yet been implemented/translated to English. An English version of the catalogue will not be prioritized before 2028.	CANCER_DISEASE	CANCER_DISEASE, HEALTH_PRODUCTS
Identifier	Provides a unique and persistent identifier for the dataset (e.g. uniform resource	The main For Health DCAT Application Profile the identifier is mandatory and its value must be unique from the origin database of the	Datasæt ID	Et unik id, der identificerer datasæt t. Skal navngives/nummereres af sundhedsdataindehaver (t abeltitlen kan evt. anvendes). Det er vigtigt,	Property name and definition are only provided in Danish as the Danish metadata catalogue has not yet been implemented/translated to English. An English version	CAR	danish-national-genome-database-research-0001

identifier (URI), digital object identifier (DOI) or catalogue identifier).
 unique identifier in the context of the Catalogue e.
 Dataset. In the process of dataset transmission to the next database, the 'identifier' property is overwritten by the new database identifier, and the original identifier becomes the value of the 'other identifier' property.
 RDF example: dct:identifier

at navn/nummer (kan også være en kombination af tekst og tal) er unikt for tabellen på tværs af alle dataindehavers tabeller. Der arbejdes på sigt, fra Visionens side, på en central løsning af ID'er for både dataindehaver, datasæt, tabeller og variable.

of the catalogue will not be prioritized before 2028.

Structured data / Is structured
 Indicates whether the dataset is structured, meaning that it is organised according to a predefined schema or model that defines its variables, attributes or fields.
 Indicates whether the Dataset contains structured data. Set it to true when a machine-readable description for which of the data variables can be published, and machine-to-machine readable structured variables do not apply. When the value is true, at least one healthdcatap:hasVariables should be provided.

Structured data' is not a property in HealthDCAT-AP release 6 which is the release the Danish metadata catalogue has implemented at this point. The property will be implemented in the Danish data model in the near future when the HealthDCAT-AP data standard are finalized.

Keyword
 Provides keywords or tags describing
 A keyword or tag describing
 Add keywords to increase dataset discoverability. You can include keywords

Nøgleord
 En liste af nøgleord der beskriver datasættet. Det anbefales, at der benyttes kendte beskrivende
 Property name and definition are only provided in Danish as the Danish metadata catalogue has not

"Kræft", "Cancer", "Neoplasi", "Neoplasmer", "Onkologi", "Tumor", "Malignitet", "Histologi", "TNM"

genome, genomics, whole genome sequencing

the dataset g the and Dataset. supporting search and discovery.	in different languages, submitting each keyword as a separate entry. Example: Physiological measures, Heart Rate, Stress Measures. RDF example: dcat:keyword	velser for eksempel fra Wikidata.	yet been implemented/translated to English. An English version of the catalogue will not be prioritized before 2028.
--	--	--	---

Language	Indicates the language or languages in which the dataset content is expressed.	A language Dataset.	For this property the values from the EU Vocabularies Languages Named Authority List must be used. If your Dataset contains multiple languages, this property can be repeated. RDF example: dct:language	Language' is not a mandatory property in HealthDCAT-AP release 6 (or release 7) and is not included in the Danish metadata catalogue. If the property becomes mandatory in a later release of HealthDCAT-AP it will be implemented in the Danish data model. Even though not all metadata properties will be implemented and published in the Danish metadata catalogue, Danish data holders will at all times be able to send every property of the HealthDCAT-AP to the Danish metadata catalogue database and from there the metadata will be send to the HealthData@EU central
------------------------	---	----------------------------	---	---

					platform and be published in the EU dataset catalogue.		
Number of records	Indicates the total number of records contained inof the dataset, number where of applicable.	Size of the dataset in terms of records.	Use this property to indicate the total number of records in the dataset, giving users an idea of the dataset's size and whether it suits their analysis or research needs. You can provide the exact count or an approximate average if the exact figure is unavailable. It reflects the total data entries (e.g., consultations, prescriptions, lab results, hospital admissions) according to the dataset's content. RDF example: health dcatap:numberOfRecords	Antal registreringer	Antal registreringer angives som et estimat (eksempelvis hele millioner, tusinder, hundrede eller lign.), der giver brugere n en forståelse af datasættet s størrelse. Jo mere præcis jo bedre, men i levende registre, hvor dette tal ændrer sig hele tiden, kan dette tal ikke være præcist.	Property name and definition are only provided in Danish as the Danish metadata catalogue has not yet been implemented/translated to English. An English version of the catalogue will not be prioritized before 2028.	(no metadata provided by data holder) 60000

Number of unique individuals	Provides an estimated number of unique natural persons represented in the dataset, where the dataset contains personal data and where applicable.	Number of records for unique individuals represented in the dataset. It helps users understand the dataset's scale and assess whether it meets their needs for secondary use. Provide the total or an approximate count of distinct individuals whose data is included. RDF example: health dcatap:numberOfUniqueIndividuals	Antal unikke individer Antal individer angives som et estimat (eksempelvis hele millioner, tusinder, hundrede eller lign.), der giver brugeren en forståelse af datasættets størrelse. Jo mere præcis jo bedre, men i levende registre, hvor dette tal ændrer sig hele tiden, kan dette tal ikke være præcist.	Property name and definition are only provided in Danish as the Danish metadata catalogue has not yet been implemented/translated to English. An English version of the catalogue will not be prioritized before 2028.	(no metadata provided by data holder)	38000
-------------------------------------	---	---	--	--	---------------------------------------	-------

Provenance	Describes the origin, source or creation process of the dataset.	A statement about the lineage of a Dataset.	Information about how the data was collected, including methodologies, tools, and protocols used, this property can be repeated for different versions of the provenance (e.g. the provenance in different languages). RDF example: dct:provenance	Original Hvordan er datasættet dannet?	Property name and definition are only provided in Danish as the Danish metadata catalogue has not yet been implemented/translated to English. An English version of the catalogue will not be prioritized before 2028.	Indsamling af oplysninger til Cancerregisteret sker i henhold til "Bekendtgørelse om lægers anmeldelse til Cancerregisteret af kræftsygdomme v.". Indberetninger er til og med 2003 foretaget på papirblanket. Fra 2004 er indberetningen fra sygehuse sket gennem Landspatientregisteret og for primærsektoren gennem Sundhedsdatastyrelsens elektroniske indberetningssystem (SEI) fra 2005. Til og med 2003 blev alle informationer kodet manuelt af koderne. Fra 2004 sker den primære databehandling af alle oplysninger med relevans eller mulig relevans for Cancerregisteret i den "automatiserede cancerlogik". Denne logik sammenholder nye oplysninger i Landspatientregisteret og Landsregisteret for Patologi med eksisterende oplysninger i Cancerregisteret og foretager derefter den initiale vurdering af, om de nye oplysninger enten skal føre til en ny sygdom i Cancerregisteret, om oplysninger skal supplere	The National Genome Database of the Danish National Genome Center contains data from genomic analyses (whole genome sequencing and Total RNA sequencing) performed on patients treated in the Danish national healthcare system. Collection started on July 1st 2019, and is still ongoing. Genomic analyses were performed primarily on Illumina's NGS platform."
-------------------	--	---	---	--	--	---	--

n
eksisterende anmeldt kræftsygdom, eller
om oplysningerne ikke tilfører noget nyt.
Efter den automatiserede databehandling foretages en manuel vurdering og databehandling i 10 - 20 pct. af tilfældene. Dette kan i nogle tilfælde medføre, at koderne fra Cancerregisteret er nødt til at gå tilbage til indberetteren for at få verificeret sygdommen.
I Cancerregisteret udføres en række valideringsrutiner med henblik på at opdage fejlklassificeringer og lignende tvivlsomme tilfælde. Dette gælder ikke mindst verifikationen af metastaser kontra primærtumorer samt afklaring af, hvorvidt 2 eller flere tumorer inden for samme organ(system) skal tolkes som en eller flere sygdomme.
Ved afslutningen af et år, og inden offentliggørelse af de endelige tal fra Cancerregisteret, foretages desuden en samkørsel med Dødsårsagsregisteret. Herved kan der identificeres personer med anmeldelsespligtige cancersygdomme, som ikke er

i Cancerregisteret. Genfindes personerne med anmeldelsespligtig sygdom i Landspatientregisteret eller i Landsregisteret for Patologi, registreres personen på baggrund af disse data i Cancerregisteret. Resterende tilfælde, der ikke kan bekræftes ved kobling til Landspatientregisteret eller Landsregisteret for Patologi, indtastes i Cancerregisteret med cancerdiagnosen fra Dødsårsagsregisteret og i makroskopisk datagrundlag registreres at Dødsårsagsregisteret er eneste datakilde.

Ved ændringer og fejlelser i de indberettede oplysninger vil cancerlogikken blive kørt om for de konkrete tilfælde, således at Cancerregisteret baseres på de korrekte oplysninger.

I det moderniserede Cancerregister fra 2004 inddrages yderligere en kilde i forhold til tidligere – patologisk-anatomiske diagnoser (SNOMED-patologi) fra Landsregisteret for Patologi.

Publisher	Indicates the entity making the dataset available, which may differ from the creator or custodian.	An entity (organisation) responsible for making the Dataset available.	Use this property to identify the agent that publishes the dataset description in the EHDS context, this may be the data holder or another organisation such as an intermediation entity. RDF example: dct:publisher	Udgiver Fx Sundhedsdatastyrelsen.	Property name and definition are only provided in Danish as the Danish metadata catalogue has not yet been implemented/translated to English. An English version of the catalogue will not be prioritized before 2028.	Sundhedsdatastyrelsen	Danish National Genome Center
Temporal coverage	Indicates the time period covered by the dataset (e.g. years of data collection), where temporal information exists.	A temporal period that the Dataset covers.	The start and end date of the period that the dataset covers. This property makes use of a Period of Time class This property makes use of the class dct:PeriodOfTime (more details on PeriodOfTime class tab) The start and end of the interval SHOULD be given by using properties dcat:startDate or time:hasBeginning, and dcat:endDate or time:hasEnd, respectively. RDF example: dct:temporal	Indsættelsesperioden Datasættets start- og slutdato (hvis det er afsluttet).	Property name and definition are only provided in Danish as the Danish metadata catalogue has not yet been implemented/translated to English. An English version of the catalogue will not be prioritized before 2028.	1.1.1943 0.00	2019-07-01 - 2026-02-01

Type	Indicates the general category or nature of the dataset (e.g. registry, administrative dataset, research dataset, synthetic dataset).	A type of Dataset. The EU Dataset-type controlled vocabulary (http://publications.europa.eu/resource/authority/dataset-type) SHOULD be used. RDF example: dct:type	Datasættype Vælges fra en predefineret liste.	Property name and definition are only provided in Danish as the Danish metadata catalogue has not yet been implemented/translated to English. An English version of the catalogue will not be prioritized before 2028.	HVD	HVD
Variables	A structured description of the variables contained in the dataset, including where applicable the variable name, definition, data type, coding system or value domain.	Links the Dataset to a CSVW TableGroup that provides fields in a machine-readable way. This machine-property becomes mandatory when healthdata:hasStructuredData is true. example: healthdata:p:hasVariables	Tabletitel (no description as the it will be too extensive due to the number of properties listed) (tabletitel) Tekniske navne (variabletitel) Beskrivelse (variabletitel) Datatype (variabletitel)	'Variables' is not a property like most others. Instead it consists of several underlying entities; 'TableGroup', 'Table' and 'Column' and related properties. In the Danish metadata catalogue the underlying and mandatory properties 'Title' (table), 'Name' (variable), 'Title' (variable), 'Description' (variable), 'Datatype' (variable) as well as the additional property 'Keyword' (table) have been implemented. Property names are only provided in Danish as the Danish metadata catalogue has not yet been implemented/translated to English. An English version	Tabletitel: Cancergruppering_I CD10 TableKeyword: - Variable titel: Cancer_diag_group_ICD10 Variable description: ICD10 - diagnosegruppering - undergrupper Variable datatype: char (3)	Tabletitel: Person TableKeyword: Individual, patient, person Variable technical name: id Variable titel: Person ID Variable description: Identifikation af en enkelt person, der kan kobles til CPR."@da, "Identification of a single person, that can be mapped to a Danish CPR Variable datatype: string

				of the catalogue will not be prioritized before 2028.		
Was generated by	Identifies the activity or process through which the dataset was generated (e.g. registry, clinical study, administrative process).	An activity associated with generation of a dataset. The Health activity provides controlled vocabulary (https://hdeu-business.dcat.acceptance.data.health.europa.eu/resource/authority/health-activity) MUST be used. RDF example: prov:wasGeneratedBy	—	Was generated by' is not a mandatory property in HealthDCAT-AP release 6 (or release 7) and is not included in the Danish metadata catalogue. If the property becomes mandatory in a later release of HealthDCAT-AP it will be implemented in the Danish data model. Even though not all metadata properties will be implemented and published in the Danish metadata catalogue, Danish data holders will at all times be able to send every property of the HealthDCAT-AP to the Danish metadata catalogue database and from there the metadata will be sent to the HealthData@EU central platform and be published in the EU dataset catalogue.	—	—
Dataset distribution	An available Distribution for the Dataset.	For Health DCAT Application Profile a distribution is mandatory independently of the access rights. A distribution must include information on the Health Data	URL adgang Lovgivning	URL til distribution. Refererer til gældende lovgivning, der påvirker oprettelse og deling af distributionen. I dansk kontekst bør der referes til Sundhedsloven (og medfølgende bekendtgørelse), Databeskyttelsesloven, EHDS og/eller GDPR.	Dataset distribution' is not a property like most others. Instead it contains of several underlying properties. In the Danish metadata catalogue the two underlying and mandatory properties 'Access URL' and 'Applicable legislation' has	https://www.retsinformation.dk/eli/lta/2018/1049 <http://data.europa.eu/eli/reg/2025/327/oj>, <https://www.retsinformation.dk/eli/lta/2018/728>

Access Body
supporting data
access.
RDF example: dcat:di
stribution

been implemented.
Property names
and definition are only
provided in Danish as the
Danish metadata catalogue
has not yet been
implemented/translated to
English. An English version
of the catalogue will not be
prioritized before 2028.

Theme	A	A Dataset may be category associated with of the multiple themes. Dataset. The EU data-theme controlled vocabulary (http://publications.europa.eu/resource/authority/data-theme) MUST be used. RDF example: dcat:theme	Datasættema	Vælges fra en liste, og der skal som minimum vælges temaet "HEAL".	Property name and definition are only provided in Danish as the Danish metadata catalogue has not yet been implemented/translated to English. An English version of the catalogue will not be prioritized before 2028.	HEAL	HEAL, TECH
--------------	---	--	--------------------	--	--	------	------------

Finland

HealthDCAT-AP Dataset	Property description in implementi	HealthDCAT-AP Dataset property	HealthDCAT-AP Dataset property usage note (R7)	NHDC property name	NHDC property definition	Additional information	Example description National cancer registry	Comments	Example description on biobank data	Comments2
-----------------------	------------------------------------	--------------------------------	--	--------------------	--------------------------	------------------------	--	----------	-------------------------------------	-----------

property name	Act 77 draft	Art. definition (R7)							
Title	Provides the official name or title of the dataset.	A name given to the Dataset.	Provide a title(s) for your Dataset, which can be repeated in multiple languages. Example: Healthy Brain Study - Physiological Data RDF example: dct:title	Title	Name of the dataset	Finnish Cancer Registry	The description is based on publicly available information. Data resource experts did not participate in preparing the description.	Auria Biobank's ovarian cancer serum and plasma samples	The description is based on publicly available information. Data resource experts did not participate in preparing the description.
Access rights	Describes the level of access applicable to the dataset (e.g. public, restricted or non public).	Information that indicates whether the Dataset is publicly accessible, has access restrictions or is not public.	This property allows you to indicate whether the dataset is publicly accessible, has restrictions, or is not publicly available. This is a mandatory property in HealthDCAT-AP. Use one of the following values :public, :restricted, :non-public, which means: NON_PUBLIC — The dataset is not publicly available and access is subject to restrictions.		We don't have exactly matching property. We have Terms of use , which indicates terms of use of the dataset with a vocabulary: Open data/data permit/Biobank permit/Contract/Data will not be transferred	Non public	In our current catalogue: Data permit	Non public	

PUBLIC — The dataset is openly accessible to everyone.
 RESTRICTED — The dataset has access restrictions but may be available under specific conditions.
 RDF example:
 dct:accessRights

Applicable legislation	Identifies Union legal acts or national legislation that specifically govern the establishment, maintenance or availability of the dataset, where such legal acts or legislation exists.	The legislation that mandates the creation or management of the Dataset.	The ELI of the EHDS Regulation, published in March 2025, can now be used as the applicable legislation where relevant: http://data.europa.eu/eli/reg/2025/327/oj . As multiple legislations may apply to the resource the maximum cardinality is not limited. RDF example: dcatap:applicable Legislation	(Is data resource under Secondary Use Act? (Editor side))	Indicates whether data resource belongs under Secondary Use Act: Yes/No	It is not exactly matching property.	http://data.europa.eu/eli/reg/2025/327/oj	In our current catalogue: Under secondary use act	http://data.europa.eu/eli/reg/2025/327/oj	
Code values	Identifies classification systems or controlled vocabularies reflected in the content of the dataset	A free-text description of any health classification, terminology, ontology,	Use this field to describe, in free text, the code values of the coding systems used that apply to the dataset (e.g. U07.1, Y59.0). No			We don't have this property in our national catalogue. We have code lists, and sometimes	IDC-10: Malignant tumours (C00-C96), Other tumours (D00-D48 and D76), Others (N85, N87, N89-N90).	It is unclear how this HealthDC AT-AP property will actually		Not known

(e.g. disease classifications, procedure classifications), where applicable.

thesaurus, or coding system used in the dataset.

controlled vocabulary is required at this stage.
RDF example:
healthdcatap:hasCodeValues

also codes and values in the variable level. The codes and values are described when the code list has not been published in an internal source, so usually these codes and values are for internal code lists used in the organisation, not from ICD-10 or SNOMED etc.

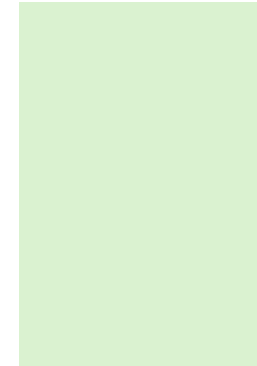
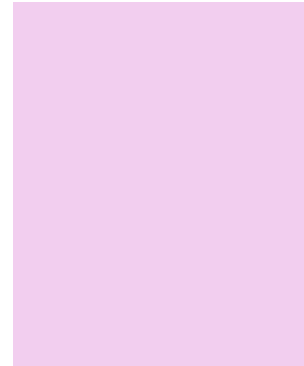
be utilized.

Coding system	Identifies standardised coding schemes used to encode data elements within the dataset (e.g. ICD 10 CM, SNOMED CT, DRGs), where applicable.	Coding systems in use (ex: ICD-10-CM, DGRs, SNOMED =CT).	This property allows you to specify the coding systems used within your dataset. For example, if a dataset uses ICD-10 for disease classification, this property allows data users to search for datasets with the same coding system. As a machine-actionable property, it also	We don't have exactly matching property. We have code lists described in variable level: Code list used in a specific variable. Possibility to either link to a code list published in external source, or	Some of the main code lists could be hand-picked manually and the information imported in the description, but it is hardly worth importing information automatically.	International Classification of Diseases, 10th Revision (ICD-10) International Classification of Diseases for Oncology, 3rd Edition (IDC-O-3)			Not known
----------------------	---	--	--	---	--	--	--	--	-----------

			facilitates automated processes, making dataset discovery more efficient. Specify the standardised coding systems used in the dataset (e.g. ICD-10-CM, SNOMED-CT). The coding systems controlled vocabulary (https://hdeu-dcat.acceptance.data.health.europa.eu/resource/authority/coding-system/) MUST be used. RDF example: healthdcatap:hasCodingSystem		provide the code list, codes and values information inside the tool.				
Contact point	Provides contact details for enquiries related to the dataset.	Contact information that can be used for sending comments about the Dataset.	This property informs about a contact point (Individual, Organization, Location, Group) that can answer questions about the dataset. Details on how to describe these are provided under class vcard:Kind. RDF example: dcat:contactPoint	(Contact information)	Informs the contact person (minimum requirement) Not exactly in the same format.	Contact Page: https://cancerregistry.fi/information/contacts/ Email: kirjaamo@cancer.fi	Currently in the Finnish catalogue: Elli Hirvonen, Contact person, phone and email. We have currently guided to use (at) instead of @ in the	Contact Page: https://www.auria.fi/biopankki/en/#ota_yhteytta Email: info@auriabiopankki.fi	

							email addresses in the Finnish catalogue, but the Dataset Description Assistant doesn't allow this.		
Conforms to	Identifies a standard or technical specification that defines the structural or syntactic organisation of the dataset as a whole.	An established standard to which the described resource conforms.	If your data conforms to an established standard or specification, use this property to indicate which one. The Health Standards controlled vocabulary (https://hdeu-dcat.acceptance.data.health.europa.eu/resource/authority/standard/) MUST be used. RDF example: <code>dct:conformsTo</code>						
Custodian	Identifies the organisation responsible for holding, managing and maintaining the dataset over time (health data holder).	Party that accepts accountability and responsibility for the data and ensures appropriate care and maintenance of the	Use this property to identify the entity that technically holds and maintains the dataset. In the EHDS context, this field represents the legally defined health data holder, (Article	Organisation	Indicates data holder / data controller / registrar/ registration authority	Not exactly in the same format.	Name: Finnish Institute for Health and Welfare (THL) Type: Public health institute Contact Page: https://thl.fi/en/front-page Email: kirjaamo@thl.fi		Name: Auria Biobank, The wellbeing services county of Southwest Finland Type: Biobank Contact Page: https://www.auria.fi/biopankki/ Email: auria-info@varha.fi

resource
[ISO-
19115].
2(6) EHDS
Regulation), i.e.
the entity
responsible for
holding and
maintaining the
dataset. This
entity may differ
from the
publisher.
RDF example:
geodcatap:custo
dian



Description	Provides a free text summary describing the content and purpose of the dataset.	A free-text information about the Dataset.	Brief description of the dataset. This property can be repeated for including several language versions of the description. RDF example: dct:description	Description	Description of the dataset	The Cancer Registry stores information necessary for the Finnish Institute for Health and Welfare (THL) to perform its statutory tasks in terms of the incidence of cancer, prevention, early detection, treatment and medical rehabilitation, and the use of services. Data are collected for statistical and scientific research purposes. Registration of cancers began nationwide in 1953. Since 1961, the submission of cancer reports has been mandatory. The Mass Screening Registry was established in 1968 in connection with the main register and has collected cervical cancer screening data since 1963, breast cancer screening data since 1987 and nationwide colorectal cancer screening data since 2022. Persons with cancer living in Finland are registered in the Cancer Registry. The data to be stored includes the patient's personal information, cancer information and death certificate information. As part of the Cancer Registry, cervical,		Serum and plasma samples and related clinical data of patients with ovarian cancer from MUPET-HERCULES-DECIDER study have been transferred to Auria Biobank. The patients are treated in Turku University Hospital since 2009. Longitudinal series have been collected from the samples, e.g. before treatment - during treatment - after treatment. The samples are stored at a temperature of -80 °C. The collection is ongoing. Medical data sources include: Medical imaging, laboratory results, medical texts, medication data, diagnoses, medical procedures, visits and wards, radiotherapy treatments.	
--------------------	---	--	--	--------------------	----------------------------	--	--	---	--

breast cancer and colorectal cancer screening registries are maintained. Data on the target population of national mass screenings are registered in these mass screening registers. Data is collected from screening invitations and examinations throughout the country.

Dataset distribution	Provides a technical representation or format in which the dataset is made available for secondary use.	An available Distribution for the Dataset.	For Health DCAT Application Profile a distribution is mandatory independently of the access rights. A distribution must include information on the Health Data Access Body supporting data access. RDF example: dcat:distribution				Access URL: https://findata.fi/en/ Applicable Legislation: http://data.europa.eu/eli/reg/2025/327/oj	It is unclear how this HealthDC AT-AP property will actually be utilized.	Access URL: https://findata.fi/en/ Applicable Legislation: http://data.europa.eu/eli/reg/2025/327/oj	
Geographical coverage	Indicates the geographic area to which the dataset relates (e.g. Member State, region or EU wide), where applicable.	A geographic region that is covered by the Dataset.	The EU Vocabularies Name Authority Lists MAY be used for continents, countries and places that are in those lists; if a particular location is not in one of the mentioned Named Authority Lists, Geonames URIs MAY be used. RDF example: dct:spatial	(Regional coverage)	Geographic region covered by the dataset	Not exactly the same. Currently will be in HealthDCAT-AP at the country level, and in our national catalogue often presented at the municipal level in free text.	Finland	Currently in the Finnish catalogue: National	Finland	The actual geographical coverage is most probably Southwest Finland, but currently not possible to specify in this level in the Dataset Description Assistant
Personal data / Has personal data	Indicates whether the dataset contains personal data within the meaning of Article 4, point (1), of	Key elements that represent an individual in the dataset, Indicates	This property is used to indicate the types of personal data contained in the dataset itself (i.e. the actual data, not the metadata record).			If this property will be utilized as the implementing act description suggests, we could	Gender, Identifier, Birth Date, Age, Medical Health	It is unclear how this HealthDC AT-AP property will actually	Gender, Medical Health	We don't know the data well enough to be able to describe in detail.

Regulation (EU) 2016/679 of the European Parliament and of the Council².

association with Personal Data.

The Data Privacy Vocabulary for Personal Data at <https://w3id.org/dpv/dpv-pd> MUST be used for this property.

Examples of allowed values include:
Gender → <https://w3id.org/dpv/dpv-pd#Gender>
Age → <https://w3id.org/dpv/dpv-pd#Age>
RDF example:
dpv:hasPersonalData

generate this information from: Does the data resource contain personal data? (Editor side only, not in catalogue). Currently this question is used to trigger the information field need for Data privacy notices on editor side only.

be utilized.

Health data access body	The competent health data access body responsible for facilitating access to the dataset under the European Health Data Space, where the dataset is subject to restricted or non public access.	Health Data Access Body supporting access to data in the Member State.	Member state-designated authority that facilitates the secondary use of electronic health data. HDABs assess the information provided by the health data applicant and decide on health data requests and access applications, authorise and issue data permits, obtain data from data holders and make data	Health Data Access Body Name: Findata - Finnish Social and Health Data Permit Authority Health Data Access Body Type: Government and Public Sector Organizations Health Data Access Body Contact Page: https://findata.fi/en/ Health Data Access Body Email: info@findata.fi		Health Data Access Body Name: Findata - Finnish Social and Health Data Permit Authority Health Data Access Body Type: Government and Public Sector Organizations Health Data Access Body Contact Page: https://findata.fi/en/ Health Data Access Body Email: info@findata.fi	
--------------------------------	---	--	--	---	--	---	--

			available in Secure Processing Environments. HDABs systematically track the data request and data access applications received and the data permits issued. As per Article 58 of the EHDS, HDABs are required to publicly list information on the data permits issued. (Article 55; Recital 52 of EHDS Regulation). RDF example: healthdcatap:hda b			
Health category / Health data category	Identifies the categories of electronic health data contained in the dataset and listed in Article 51(1) of Regulation (EU) 2025/327.	The health category to which this dataset belongs as described in the Commission Regulation on the European Health Data Space	This property indicates the EHDS data category the dataset belongs to, based on Article 51 (https://eur-lex.europa.eu/legislation/content/EN/TXT/HTML/?uri=OJ:L_202500327#ntr35-L_202500327EN.000101-E0035) of the EHDS	Data from population-based health data registries such as public health registries	Health data from biobanks and associated databases	

laying down a list of categories of electronic data for secondary use, Art.51.

Regulation. Use the controlled vocabulary <http://13.81.34.152:1101/resource/authority/healthcategories/> to ensure standardised classification across all datasets
RDF example: `healthdcatap:healthCategory`

Health theme	Indicates a high level health topic or theme associated with the dataset (e.g. cardiovascular disease, cancer).	A category of the Dataset or tag describing the Dataset.	A dataset may be associated with multiple health themes. The Health Theme controlled vocabulary (http://13.81.34.152:1101/resource/authority/health-theme/) MUST be used. RDF example: <code>healthdcatap:healthTheme</code>	URN	Unique and persistent identifier	Cancer	Cancer
Identifier	Provides a unique and persistent identifier for the dataset (e.g. uniform resource identifier (URI), digital object identifier	The main identifier for the Dataset, e.g. the URI or other unique identifier in the context of the	For Health DCAT Application Profile the identifier is mandatory and its value must be unique from the origin database of the Dataset. In the process of dataset transmission to the next database,			https://urn.fi/URN:NBN:fi:thl:ak:21085403-7be8-4f93-bf05-231518c642a0	bbmri-eric:ID:FI_001:collection:Auria_Hercules

	(DOI) or catalogue identifier).	Catalogue .	the 'identifier' property is overwritten by the new database identifier, and the original identifier becomes the value of the 'other identifier' property. RDF example: dct:identifier				
Structured data / Is structured	Indicates whether the dataset is structured, meaning that it is organised according to a predefined schema or model that defines its variables, attributes or fields.	Indicates whether the Dataset contains structured data for which a machine-readable description of the data variables can be provided.	Use this boolean property to state whether the dataset contains structured data. Set it to true when a machine-readable description of the data variables can be published, and to false when such structured variables do not apply. When the value is true, at least one healthdcatap:hasVariables should be provided.		Yes		Yes
Keyword	Provides keywords or tags describing the dataset and supporting search and discovery.	A keyword or tag describing the Dataset.	Add keywords to increase dataset discoverability. You can include keywords in different languages, submitting each keyword as a separate entry. Example:	Keywords Keywords describing the dataset. Ontologies FinMesh, YSO and TERO are being used.	cancer, cancerous diseases, registers, screening, mass screening, breast cancer, cervical cancer, colorectal cancer, prevention, national cancer registry		ovarian cancer, serum samples, plasma samples, clinical data

Physiological measures, Heart Rate, Stress Measures. RDF example: dcat:keyword

Language	Indicates the language or languages in which the dataset content is expressed.	A language of the Dataset.	For this property the values from the EU Vocabularies Languages Named Authority List must be used. If your Dataset contains multiple languages, this property can be repeated. RDF example: dct:language	No particular field for the language of the dataset but descriptions are available in FIN/ENG/SWE	FIN, SWE	1 700 000	Information was found that there would be roughly 1,7 million registered cancer cases. Unclear if it should be interpreted as number of records. The Dataset Description	500	It says 500 samples, but it is not the same as number of records. https://directory.bbmri-eric.eu/ERIC/directory/#/collection/bbmri-eric:ID:FI_001:collection:Auria_Hercules
Number of records	Indicates the total number of records contained in the dataset, where applicable.	Size of the dataset in terms of the number of records.	Use this property to indicate the total number of records in the dataset, giving users an idea of the dataset's size and whether it suits their analysis or research needs. You can provide the exact count or an approximate average if the exact figure is unavailable. It reflects the total data entries (e.g.,	We don't have this exact property in our catalogue. Instead, we have Number of observation units.					

consultations,
prescriptions, lab
results, hospital
admissions)
according to the
dataset's content.
RDF example:
healthdcatap:nu
mberOfRecords

n
Assistant
only
accepts
numbers.

Number
of
unique
individu
als

Provides an estimated number of unique natural persons represented in the dataset, where the dataset contains personal data and where applicable.

Number of
records
for unique
individuals

Use this property to indicate the number of unique individuals represented in the dataset. It helps users understand the dataset's scale and assess whether it meets their needs for secondary use. Provide the total or an approximate count of distinct individuals whose data is included. RDF example: `healthcatap:numbOfUniqueIndivduals`

3600000

Information found by Googling, AI overview: The Finnish Cancer Registry and Mass Screening Registers do not disclose a single, static aggregate of all historical unique individuals, as this number is constantly expanding. However, over 3.6 million unique individuals have been registered since

100

Number of sample donors:
100.

					1953, with over 334,000 living Finns previously diagnosed with cancer still on the active roster.		
Provenance	Describes the origin, source or creation process of the dataset.	A statement about the lineage of a Dataset.	Information about how the data was collected, including methodologies, tools, and protocols used, this property can be repeated for different versions of the provenance (e.g. the provenance in different languages). RDF example: dct:provenance		Regular sources of information in the register: -Healthcare organizations. - Municipal data on cancer patients and data on emigration and immigration are obtained from the Population Information System of the Digital and Population Data Services Agency. - The monitoring data of the Mass Screening Register and the data of the target population for mass screenings are also obtained from the Population Information System of the Digital and Population Data Services Agency. - Data on socioeconomic status, occupation and education, as well as data on causes of death, are available from Statistics Finland.		

Publisher	Indicates the entity making the dataset available, which may differ from the creator or custodian.	An entity (organisation) responsible for making the Dataset available.	Use this property to identify the agent that publishes the dataset description in the catalogue. In the EHDS context, this may be the data holder or another organisation such as an intermediation entity. RDF example: dct:publisher		We only have one organisation information in our catalogue, and it is the data holder.	Publisher Name: The Finnish Cancer Society Publisher Type: Public health registry Publisher Contact Page: https://cancerregistry.fi/ Publisher Email: kirjaamo@cancer.fi	THL is the registrar of the Cancer Registry. The Finnish Cancer Society is the technical administrator of the Cancer Registry and the processor of personal data, and has also made the metadata description.	Name: Auria Biobank, The wellbeing services county of Southwest Finland Type: Biobank Contact Page: https://www.auria.fi/biopankki/ Email: auria-info@varha.fi	
Temporal coverage	Indicates the time period covered by the dataset (e.g. years of data collection), where temporal information exists.	A temporal period that the Dataset covers.	The start and end date of the period that the dataset covers. This property makes use of a Period of Time class This property makes use of the class <code>dct:PeriodOfTime</code> (more details on <code>PeriodOfTime</code> class tab) The start and end of the interval SHOULD be given by using properties	Reference date	Temporal period that the Dataset covers	1.1.1953-		1.1.2009-	

				dcat:startDate or time:hasBeginning, and dcat:endDate or time:hasEnd, respectively. RDF example: dct:temporal					
Type	Indicates the general category or nature of the dataset (e.g. registry, administrative dataset, research dataset, synthetic dataset).	A type of the Dataset.	A Dataset may be associated with multiple types. The EU Dataset-type controlled vocabulary (http://publications.europa.eu/resource/authority/dataset-type) SHOULD be used. RDF example: dct:type	(Data resource type)	Type of the dataset	Not the same vocabulary.	Registry (but currently in the Dataset Description Assistant: Statistical data)	The current vocabulary in the Dataset description assistant is not usable and is not in line with the description in the implementing act draft.	None of the options currently in the Dataset Description Assistant work. The ones mentioned in the implementing act: maybe research dataset.
Variables	A structured description of the variables contained in the dataset, including, where applicable, the variable name, definition, data type, coding system or value domain.	Links the Dataset to a CSVW TableGroup that provides a machine-readable description of the data variables.	Use this property to link a structured dataset to a CSVW TableGroup that documents its data variables, columns, or fields in a machine-readable way. This property becomes mandatory when healthdcatap:hasStructuredData is true.	Variables	In the Finnish model no Table descriptions currently. We have Variable group which might have the same information content in some cases, but it's not used thoroughly.		For example: Table URL: https://aineistokatalogi.fi/catalog/studies/21085403-7be8-4f93-bf05-231518c642a0/datasets/ca8a40b0-1e54-447f-9bd2-70cfe82296f4 Table Title: Basic variables Table Keyword: cancer registry, basic variables Column Name: dg_date Column Title: Date of diagnosis Column Description: Date of cancer diagnosis.		

RDF example:
healthdcatap:has
Variables

Column Datatype: date
Column Property URL: -

Was generated by	Identifies the activity or process through which the dataset was generated (e.g. registry, clinical study, administrative process).	An activity that generated, or provides the business context for, the creation of the dataset.	The activity associated with generation of a dataset. The Health activity controlled vocabulary (https://hdeu-dcat.acceptance.data.health.europa.eu/resource/authority/health-activity) MUST be used. RDF example: prov:wasGeneratedBy	Registry (but currently in the Dataset Description Assistant: National Health Registries)	The current vocabulary in the Dataset description assistant seems not to be in line with the description in the implementing act draft.	Biobank/sample collection Patient admission, care and discharge
Theme		A category of the Dataset.	A Dataset may be associated with multiple themes. The EU data-theme controlled vocabulary (http://publications.europa.eu/resource/authority/data-theme) MUST be used. RDF example: dcat:theme	Health		Health

Iceland

HealthDCAT-AP Dataset property name	Property description in implementing act Art. 77 draft	HealthDCAT-AP Dataset property definition (R7)	HealthDCAT-AP Dataset property usage note (R7)	NHD C property name	NHD C property definition	Additional information	Example description on National cancer registry	Comments
Title	Provides the official name or title of the dataset.	A name given to the Dataset.	Provide a title(s) for your Dataset, which can be repeated in multiple languages. Example: Healthy Brain Study - Physiological Data RDF example: dct:title				The Icelandic Cancer Register	
Access rights	Describes the level of access applicable to the dataset (e.g. public, restricted or non public).	Information that indicates whether the Dataset is publicly accessible, has access restrictions or is not public.	This property allows you to indicate whether the dataset is publicly accessible, has restrictions, or is not publicly available. This is a mandatory property in HealthDCAT-AP. Use one of the following values :public, :restricted, :non-public, which means:				Non-public	The dataset has access restrictions; access may be granted under specific conditions for scientific research purposes.

NON_PUBLIC — The dataset is not publicly available and access is subject to restrictions.

PUBLIC — The dataset is openly accessible to everyone.

RESTRICTED — The dataset has access restrictions but may be available under specific conditions.

RDF example: dct:accessRights

Applicable legislation	Identifies Union legal acts or national legislation that specifically govern the establishment, maintenance or availability of the dataset, where such legal acts or legislation exists.	The legislation that mandates the creation or management of the Dataset.	The ELI of the EHDS Regulation, published in March 2025, can now be used as the applicable legislation where relevant: http://data.europa.eu/eli/reg/2025/327/oj. As multiple legislations may apply to the resource the maximum cardinality is not limited. RDF example: dcatap:applicableLegislation	The Icelandic Cancer Register <i>Krabbameinsskrá Íslands fékk lagagrunn árið 2007 (lög nr 41/2007 og reglugerð nr 548/2008) og er hún því samkvæmt lögum ein af heilbrigðisskrám Landlæknis. Samkvæmt lögum um landlækni og lýðheilsu er tilgangur þess að halda heilbrigðisskrár m.a. sá að nýta gögnin til vísindarannsókna.</i>	Most probably the link to EHDS Regulation will be given here.
Code values	Identifies classification systems or controlled vocabularies reflected in the content of the dataset (e.g. disease classifications, procedure classifications), where applicable.	A free-text description of any health classification, terminology, ontology, thesaurus, or coding system used in the dataset.	Use this field to describe, in free text, the code values of the coding systems used that apply to the dataset (e.g. U07.1, Y59.0). No controlled vocabulary is required at this stage. RDF example: healthdcatap:hasCodeValues	<i>ICD-10 topography codes: C00–C97 (malignant neoplasms), D00–D09 (carcinoma in situ), D22, D32–D35, D42–D44, D45–D47 (benign and uncertain behaviour neoplasms of CNS, endocrine and haematopoietic tissue); ICD-O-3 topography codes (e.g. C50.9 breast, C61.9 prostate); ICD-O-3 morphology codes (e.g. 8500/3 infiltrating duct carcinoma, 9732/3 multiple myeloma); Basis of diagnosis codes: 0 (death certificate only), 1 (clinical), 2 (investigation without histology), 4 (biochemical), 5</i>	Krabbameinsskránin er í samræmi við leiðbeiningar Evrópusamtaka krabbameinsskráa (European Network of Cancer Registries, ENCR). Skráð er samkvæmt alþjóðlegri sjúkdómaflokkun (International Classification of

				(cytology), 6 (histology of metastasis), 7 (histology of primary tumour); Laterality codes: 0 (not paired organ), 1 (right), 2 (left), 4 (bilateral), 9 (unknown); Sex codes: 1 (male/boy), 2 (female/girl), 7 (non-binary/child); Tumour number codes: 10 (first malignant), 20 (second malignant), 00 (first benign), 07 (diagnosed abroad) ICD-10, ICD-03	Disease, ICD) og ICD-O (International Classification of Diseases for Oncology).
Coding system	Identifies standardised coding schemes used to encode data elements within the dataset (e.g. ICD 10 CM, SNOMED CT, DRGs), where applicable.	Coding systems in use (ex: ICD-10-CM, DGRs, SNOMED=CT).	This property allows you to specify the coding systems used within your dataset. For example, if a dataset uses ICD-10 for disease classification, this property allows data users to search for datasets with the same coding system. As a machine-actionable property, it also facilitates automated processes, making dataset discovery more efficient. Specify the standardised coding systems used in the dataset (e.g. ICD-10-CM, SNOMED-CT). The coding systems controlled vocabulary (https://hdeu-dcat.acceptance.data.health.europa.eu/resource/authority/coding-system/) MUST be used. RDF example: healthdcatap:hasCodingSystem	ICD-10 (International Classification of Diseases, 10th Revision) — for tumour topography and cause of death; ICD-O-3 (International Classification of Diseases for Oncology, 3rd Edition) — for tumour topography (ksicd0) and morphology (ksmeingerd/fsmeingerd); Multiple ICD versions historically: ICD-7, ICD-8, ICD-9, ICD-10 (tracked per record in fsdanorsicdkerfi)	
Contact point	Provides contact details for enquiries related to the dataset.	Contact information that can be used for sending comments about the Dataset.	This property informs about a contact point (Individual, Organization, Location, Group) that can answer questions about the dataset. Details on how to describe these are provided under class vcard:Kind. RDF example: dcat:contactPoint	mottaka@landlaeknir.is	

Conforms to	Identifies a standard or technical specification that defines the structural or syntactic organisation of the dataset as a whole.	An established standard to which the described resource conforms.	If your data conforms to an established standard or specification, use this property to indicate which one. The Health Standards controlled vocabulary (https://hdeu-dcat.acceptance.data.health.europa.eu/resource/authority/standard/) MUST be used. RDF example: <code>dct:conformsTo</code>		
Custodian	Identifies the organisation responsible for holding, managing and maintaining the dataset over time (health data holder).	Party that accepts accountability and responsibility for the data and ensures appropriate care and maintenance of the resource [ISO-19115].	Use this property to identify the entity that technically holds and maintains the dataset. In the EHDS context, this field represents the legally defined health data holder, (Article 2(6) EHDS Regulation), i.e. the entity responsible for holding and maintaining the dataset. This entity may differ from the publisher. RDF example: <code>geodcatap:custodian</code>	Icelandic Cancer Society (Krabbameinsfélag Íslands) — legally responsible for maintaining and managing the registry on behalf of the Directorate of Health (Embætti landlæknis)	

Description	Provides a free text summary describing the content and purpose of the dataset.	A free-text information about the Dataset.	Brief description of the dataset. This property can be repeated for including several language versions of the description. RDF example: dct:description	The Icelandic Cancer Register is a nationwide health registry that systematically collects data on all cancer diagnoses in Iceland. It is overseen by the Directorate of Health and operated by the Icelandic Cancer Association. Established with electronic records dating back to 1955, the register's primary purpose is to support healthcare quality control and enable scientific and epidemiological research. It compiles data from hospitals, healthcare practitioners, laboratories, and public health authorities, ensuring comprehensive national coverage. The registry includes standardized information—aligned with international frameworks such as the European Network of Cancer Registries and the International Association of Cancer Registries—covering diagnosis, pathology, treatment, and patient outcomes. Its data contribute to national statistics, international cancer databases, and ongoing research, making it a key resource for understanding cancer trends and improving care in Iceland.	Skráin nær til meir en 99% allra meina sem greinast hjá þjóðinni og er í flokki bestu krabbameinsskráa heims ásamt öðrum Norrænum skráum.
Dataset distribution	Provides a technical representation or format in which the dataset is made available for secondary use.	An available Distribution for the Dataset.	For Health DCAT Application Profile a distribution is mandatory independently of the access rights. A distribution must include information on the Health Data Access Body supporting data access. RDF example: dcat:distribution	Restricted access; applications submitted to the Directorate of Health	Publicly available cancer statistics from the Icelandic Cancer Register can be freely accessed and reused with proper citation, while customized statistical queries may incur a fee.

				Iceland	Publication of data in media requires prior contact with the registry. Access to identifiable data for research purposes requires a formal application to the Directorate of Health, including supporting documents and a specified list of requested variables.
Geographical coverage	Indicates the geographic area to which the dataset relates (e.g. Member State, region or EU wide), where applicable.	A geographic region that is covered by the Dataset.	The EU Vocabularies Name Authority Lists MAY be used for continents, countries and places that are in those lists; if a particular location is not in one of the mentioned Named Authority Lists, Geonames URIs MAY be used. RDF example: dct:spatial		National coverage Certain data on all cancers diagnosed in Iceland since 1955. Skráin nær til meir en 99% allra meina sem greinast hjá þjóðinni og er í flokki bestu krabbameinsskráa heims ásamt öðrum Norrænum skráum.
Personal data / Has personal data	Indicates whether the dataset contains personal data within the	Key elements that represent an individual in the dataset, Indicates	This property is used to indicate the types of personal data contained in the dataset itself (i.e. the actual data, not the metadata record).	Yes	PIN, Birth year, Gender, residence when diagnosed, diagnosis date, tumour site,

meaning of Article 4, point (1), of Regulation (EU) 2016/679 of the European Parliament and of the Council².

association with Personal Data.

The Data Privacy Vocabulary for Personal Data at <https://w3id.org/dpv/dpv-pd> MUST be used for this property.

Examples of allowed values include:

Gender → <https://w3id.org/dpv/dpv-pd#Gender>

Age → <https://w3id.org/dpv/dpv-pd#Age>

RDF example: `dpv:hasPersonalData`

morphology, basis of diagnosis, pathology results, causes of death, referring physician speciality, treating physician speciality

Health data access body

The competent health data access body responsible for facilitating access to the dataset under the European Health Data Space, where the dataset is subject to restricted or non public access.

Health Data Access Body supporting access to data in the Member State.

Member state-designated authority that facilitates the secondary use of electronic health data. HDABs assess the information provided by the health data applicant and decide on health data requests and access applications, authorise and issue data permits, obtain data from data holders and make data available in Secure Processing Environments. HDABs systematically track the data request and data access applications received and the data permits issued. As per Article 58 of the EHDS, HDABs are required to publicly list information on the data permits issued. (Article 55; Recital 52 of EHDS Regulation).
RDF example: `healthdcatap:hdab`

Health category / Health data category	Identifies the categories of electronic health data contained in the dataset and listed in Article 51(1) of Regulation (EU) 2025/327.	The health category to which this dataset belongs as described in the Commission Regulation on the European Health Data Space laying down a list of categories of electronic data for secondary use, Art.51.	This property indicates the EHDS data category the dataset belongs to, based on Article 51 (https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=OJ:L_20250327#ntr35-L_202500327EN.000101-E0035) of the EHDS Regulation. Use the controlled vocabulary http://13.81.34.152:1101/resource/authority/healthcategories/ to ensure standardised classification across all datasets	K Data on professional status, and on the specialisation and institution of health professionals involved in the treatment of a natural person
			RDF example: healthdcatap:healthCategory	
Health theme	Indicates a high level health topic or theme associated with the dataset (e.g. cardiovascular disease, cancer).	A category of the Dataset or tag describing the Dataset.	A dataset may be associated with multiple health themes. The Health Theme controlled vocabulary (http://13.81.34.152:1101/resource/authority/health-theme) MUST be used. RDF example: healthdcatap:healthTheme	Cancer

Identifier	Provides a unique and persistent identifier for the dataset (e.g. uniform resource identifier (URI), digital object identifier (DOI) or catalogue identifier).	The main identifier for the Dataset, e.g. the URI or other unique identifier in the context of the Catalogue.	For Health DCAT Application Profile the identifier is mandatory and its value must be unique from the origin database of the Dataset. In the process of dataset transmission to the next database, the 'identifier' property is overwritten by the new database identifier, and the original identifier becomes the value of the 'other identifier' property.
-------------------	--	---	---

RDF example: dct:identifier

Structured data / Is structured	Indicates whether the dataset is structured, meaning that it is organised according to a predefined schema or model that defines its variables, attributes or fields.	Indicates whether the Dataset contains structured data for which a machine-readable description of the data variables can be provided.	Use this boolean property to state whether the dataset contains structured data. Set it to true when a machine-readable description of the data variables can be published, and to false when such structured variables do not apply. When the value is true, at least one healthdcatap:hasVariables should be provided.
--	---	--	--

Keyword	Provides keywords or tags describing the dataset and supporting search and discovery.	A keyword or tag describing the Dataset.	Add keywords to increase dataset discoverability. You can include keywords in different languages, submitting each keyword as a separate entry. Example: Physiological measures, Heart Rate, Stress Measures.
----------------	---	--	---

RDF example: dcat:keyword

Yes

The dataset is structured according to a predefined schema

Language	Indicates the language or languages in which the dataset content is expressed.	A language of the Dataset.	For this property the values from the EU Vocabularies Languages Named Authority List must be used. If your Dataset contains multiple languages, this property can be repeated. RDF example: dct:language	Icelandic	
Number of records	Indicates the total number of records contained in the dataset, where applicable.	Size of the dataset in terms of the number of records.	Use this property to indicate the total number of records in the dataset, giving users an idea of the dataset's size and whether it suits their analysis or research needs. You can provide the exact count or an approximate average if the exact figure is unavailable. It reflects the total data entries (e.g., consultations, prescriptions, lab results, hospital admissions) according to the dataset's content. RDF example: healthdcatap:numberOfRecords		
Number of unique individuals	Provides an estimated number of unique natural persons represented in the dataset, where the dataset contains personal data	Number of records for unique individuals.	Use this property to indicate the number of unique individuals represented in the dataset. It helps users understand the dataset's scale and assess whether it meets their needs for secondary use. Provide the total or an approximate count of distinct individuals whose data is included.		

	and where applicable.		RDF example: healthdcatap:numberOfUniqueIndividuals		
Provenance	Describes the origin, source or creation process of the dataset.	A statement about the lineage of a Dataset.	Information about how the data was collected, including methodologies, tools, and protocols used, this property can be repeated for different versions of the provenance (e.g. the provenance in different languages).	<i>Data originate from three source tables within the registry: (1) kskra_kjarskra (core registry) — primary tumour data including ICD-10/ICD-O-3 codes, diagnosis date, age, laterality, basis of diagnosis; (2) kskra_frumskra (source/notification registry) — pathology and cytology lab reports, referring institution, physician, autopsy data, morphology, cause of death; (3) kskra_nafnaskra_safn (name/population register linkage) — personal identifier, date of death, linked via Þjóðskrá Íslands (National Population Register). Mandatory notifications are submitted by hospitals, pathology laboratories, and haematology units across Iceland. Death certificate linkage is performed via the death cause register of the Directorate of Health (Embætti landlæknis). The registry follows ENCR data collection and quality assurance standards. Electronic data available from 1955; benign CNS and haematopoietic tumours recorded from 2003 onwards for some categories.</i>	Upplýsingar til Krabbameinsskrár berast frá rannsóknarstofum í meinafræði og blóðmeinafræði, sjúkradeildum, heilsugæslustöðvum, einstökum læknum og læknaðstofum. Upplýsingar berast að jafnaði innan tveggja mánaða frá greiningu. Upplýsingar um mein þar sem ekki liggur fyrir meinafræðigreinin g koma til viðbótar úr Dánarmeinaskrá og Vistunarskrá heilbrigðisstofnana. Tölfræði krabbameina á ársgrundvelli er birt á heimasíðu Krabbameinsfélagsins www.krabb.is
Publisher	Indicates the entity making the dataset available, which may differ from the creator or custodian.	An entity (organisation) responsible for making the Dataset available.	Use this property to identify the agent that publishes the dataset description in the catalogue. In the EHDS context, this may be the data holder or another organisation such as an intermediation entity.		

RDF example: dct:provenance

RDF example: dct:publisher				Start date: 1911-01-01 (breast cancer records); 1955-01-01 (all other tumour types, electronic records). End date: ongoing — registry updated daily. Note: some tumour categories (e.g. myelodysplastic syndromes D45–D47, MGUS D47.2) have later start dates for electronic registration (from 1999–2003).	Only one temporal coverage can be given currently in HealthDCAT-AP.
Temporal coverage	Indicates the time period covered by the dataset (e.g. years of data collection), where temporal information exists.	A temporal period that the Dataset covers.	The start and end date of the period that the dataset covers. This property makes use of a Period of Time class. This property makes use of the class dct:PeriodOfTime (more details on PeriodOfTime class tab). The start and end of the interval SHOULD be given by using properties dcat:startDate or time:hasBeginning, and dcat:endDate or time:hasEnd, respectively.		
RDF example: dct:temporal				Register	
Type	Indicates the general category or nature of the dataset (e.g. registry, administrative dataset, research dataset, synthetic dataset).	A type of the Dataset.	A Dataset may be associated with multiple types. The EU Dataset-type controlled vocabulary (http://publications.europa.eu/resource/authority/dataset-type) SHOULD be used. RDF example: dct:type		

Variables	A structured description of the variables contained in the dataset, including, where applicable, the variable name, definition, data type, coding system or value domain.	Links the Dataset to a CSVW TableGroup that provides a machine-readable description of the data variables.	Use this property to link a structured dataset to a CSVW TableGroup that documents its data variables, columns, or fields in a machine-readable way. This property becomes mandatory when healthdcatap:hasStructuredData is true. RDF example: healthdcatap:hasVariables
------------------	---	--	--

Variables in the Icelandic Cancer Registry (source: Krabbameinsskrá breytulisti, February 2023):

PATIENT-LEVEL VARIABLES:

- kennitala — National ID (pseudonymised research number; source: kskra_nafnaskra_safn)
- fæðingarár — Year of birth
- kyn — Sex per national register: 1=male/boy, 2=female/girl, 7=non-binary/child (from 2009; source: kskra_nafnaskra)
- kssveitarfelag — Residence at diagnosis (health region/municipality; from 2009; note: poorly recorded for benign tumours 1955–2003; source: kskra_kjarnskra)
- danardagur — Date of death (delivered as month and year; from 2009; source: kskra_nafnaskra_safn)

TUMOUR-LEVEL VARIABLES (CORE REGISTRY — kskra_kjarnskra):

- ksdagsfyrstufaerslu — Date of diagnosis (dd.mm.yyyy; from 2009)
- ksnraexlis — Tumour number: 10=first malignant, 20=second malignant, 00=first benign, 07=diagnosed abroad, A1/A2=diagnosed before 1955
- ksgral — Age at diagnosis (integer)
- kscid10 — Tumour topography (ICD-10; from 2009)
- kscid0 — Tumour topography (ICD-O-3; from 2009)
- kshlid — Laterality: 0=not paired organ, 1=right, 2=left, 4=bilateral, 9=unknown (from 2009)
- ksnagr — Basis of diagnosis: 0=death certificate only, 1=clinical, 2=investigation without histology, 4=biochemical, 5=cytology,

6=histology of metastasis, 7=histology of primary tumour (from 2009)

- ksmeingerd — Tumour morphology (ICD-O-3; from 2009)

NOTIFICATION/PATHOLOGY VARIABLES (kskra_frumskra):

- fssendstofnun — Reporting healthcare institution (text; from 2009)

- fssendlaeknir — Referring physician (delivered as medical specialty; from 2009)

- fssendnrlaeknis — Treating physician registration number (delivered as specialty; source: professional licence register)

- fsdagskomu — Date of sample receipt at laboratory (dd.mm.yyyy; from 2009)

- fsrannst — Laboratory performing the diagnosis (from 2009)

- fsgrkr — Diagnosed at autopsy: 0=no, 1=yes (from 2009)

- fsnrpad — Pathology and/or autopsy number (from 2009)

- fsnraexlis — Tumour number (notification-level; from 2009)

- fsicd10 / fsicd0 — Tumour topography ICD-10 / ICD-O-3 (notification level; from 2009)

- fshlid — Laterality (notification level; from 2009)

- fsnagr — Basis of diagnosis (notification level; from 2009)

- fsmeingerd — Tumour morphology ICD-O-3 (notification level; from 2009)

- fsdanorsicd — Underlying cause of death per death certificate (ICD code; from 2009)

- fsdanorsicdkerfi — ICD version used for cause of death coding: ICD-7, ICD-8, ICD-9, or ICD-10 (from 2009)

ICD-10 TOPOGRAPHY SCOPE:
C00–C97 (malignant), D00–D09 (in
situ), D22, D32–D35, D42–D44, D45–
D47 (benign/uncertain CNS,
endocrine, haematopoietic)

Was generated by	Identifies the activity or process through which the dataset was generated (e.g. registry, clinical study, administrative process).	An activity that generated, or provides the business context for, the creation of the dataset.	The activity associated with generation of a dataset. The Health activity controlled vocabulary (https://hdeu-dcat.acceptance.data.health.europa.eu/resource/authority/health-activity) MUST be used. RDF example: prov:wasGeneratedBy	Population-based cancer registration process established in 1955 (breast cancer from 1911) under the Icelandic Cancer Society. Mandatory notifications submitted by hospitals, pathology laboratories (vefjagreining/histology, frumugreining/cytology), and haematology units across Iceland. Death certificate linkage via the death cause register of the Directorate of Health. Population register linkage via Þjóðskrá Íslands (National Population Register). Registry updated daily. Data consolidation across three internal source tables: core registry (kskra_kjarnskra), notification registry (kskra_frumskra), and population linkage (kskra_nafnaskra_safn).	Upplýsingar til Krabbameinsskrár berast frá rannsóknarstofum í meinafræði og blóðmeinafræði, sjúkradeildum, heilsugæslustöðvum, einstökum læknum og læknaðstofum. Upplýsingar berast að jafnaði innan tveggja mánaða frá greiningu. Upplýsingar um mein þar sem ekki liggur fyrir meinafræðigreinin g koma til viðbótar úr Dánarmeinaskrá og Vistunarskrá heilbrigðisstofnana. Tölfræði krabbameina á ársgrundvelli er birt á heimasíðu Krabbameinsfélagsins www.krabb.is
Theme		A category of the Dataset.	A Dataset may be associated with multiple themes. The EU data-theme controlled vocabulary (http://publications.europa.eu/resource/authority/data-theme) MUST be used. RDF example: dcat:theme	Health	

Norway

HealthDCAT-AP Dataset property name	Property description in implementing act Art. 77 draft	HealthDCAT-AP Dataset property definition (R7)	HealthDCAT-AP Dataset property usage note (R7)	NHDC property name	NHDC property definition	Additional information	Example description on National cancer registry	Comments	Example description on biobank data	Comments2
Title	Provides the official name or title of the dataset.	A name given to the Dataset.	Provide a title(s) for your Dataset, which can be repeated in multiple languages. Example: Healthy Brain Study - Physiological Data RDF example: dct:title	PreferredTerm	Kort beskrivende navn på datakilden, delkilden eller datasamlingen		Kreftregister, Kvalitetsregister for hjerne- og ryggmargssvulster	title of the datasource and datacollection	The Norwegian mother, father and child cohort, data acquired from biobank samples, MAB012 , https://www.fhi.no/contentassets/c28c0ec2d4e24e36bbcb04ceeaba3d05/mab012.pdf	
Access rights	Describe the level of access applicable to the dataset (e.g. public, restricted or non public).	Information that indicates whether the Dataset is publicly accessible, has access restrictions or is not public.	This property allows you to indicate whether the dataset is publicly accessible, has restrictions, or is not publicly available. This is a mandatory property in HealthDCAT-AP. Use one of the following values :public, :restricted, :non-public, which means: NON_PUBLIC — The dataset is not publicly available and access is subject to restrictions.	HjemmelTilgjengeliggjøring	URL til paragraf med hjemmel for tilgjengeliggjøring	This describes the legal basis for making data available, i.e. conditions under which access is allowed. This is the closest direct equivalent	https://lovdata.no/dokument/NL/lov/2014-06-20-43#KAPITTEL_3;%20https://lovdata.no/dokument/NL/lov/2014-06-20-43#KAPITTEL_4	we are going to adopt this property in the future	Access to data from MoBa is regulated through several laws and directives. Especially the directive/regulation about health cohorts is of great importance.	

			PUBLIC — The dataset is openly accessible to everyone. RESTRICTED — The dataset has access restrictions but may be available under specific conditions. RDF example: dct:accessRights			ent to Access rights in HealthD CAT-AP.			
Applicable legislation	Identifies Union legal acts or national legislation that specifically govern the establishment, maintenance or availability of the dataset, where such legal acts or legislation exists.	The legislation that mandates the creation or management of the Dataset.	The ELI of the EHDS Regulation, published in March 2025, can now be used as the applicable legislation where relevant: http://data.europa.eu/eli/reg/2025/327/oj . As multiple legislations may apply to the resource the maximum cardinality is not limited. RDF example: dcatap:applicableLegislation	Lovverk, Forskrift	En eller flere URL-er til lovverk på lovdata.no. F.eks. URL til Lov om helseregistre og behandling av helseopplysninger. Flere URL-er skilles ved å bruke «;».	Applicable legislation = legal acts and regulations governing the dataset	https://lovdata.no/dokument/NL/lov/2014-06-20-43 https://lovdata.no/dokument/SF/forskrift/2001-12-21-1477	In the national specification, Lovverk and Forskrift are defined as one or more URLs to relevant laws and regulations on lovdata.no that govern the dataset.	Forskrift om befolkningsbaserte helseundersøkelser - Lovdata
Code values	Identifies classification systems or	A free-text description of any health classification, terminology, ontology,	Use this field to describe, in free text, the code values of the coding systems used that apply to	Kode	Kodeverdien er koden som står foran en kodelinje, f.eks. 1 for Mann. Kodeverdier kan	The code value is the code that	«2» og «Kvinne».		Several code classification systems are used because of linkage to registers especially for disease specific registers have specific classification

controlled vocabularies reflected in the content of the dataset (e.g. disease classifications, procedure classifications), where applicable.

thesaurus, or coding system used in the dataset.

the dataset (e.g. U07.1, Y59.0). No controlled vocabulary is required at this stage.
RDF example:
healthdatacatap:hasCodeValues

være alfanumeriske og inneholde både mellomrom og forskjellige tegn. Dersom man skal koble en kodeverdi med kodeløst er det derfor viktig at formatene er «en-til-en».

precede s a code description, for example 1 for 'Male'. Code values may be alphanumeric and can include both spaces and various characters. Therefore, when linking a code value with a code description, it is important that the formats correspond on a one-to-one basis.

systems, domain specific controlled vocabularies is also used i.e. for genetic, and other omics techniques.

Coding system	Identifies standardised coding schemes used to encode data elements within the dataset (e.g. ICD 10 CM, SNOMED CT, DRGs), where applicable.	This property allows you to specify the coding systems used within your dataset. For example, if a dataset uses ICD-10 for disease classification, this property allows data users to search for datasets with the same coding system. As a machine-actionable property, it also facilitates automated processes, making dataset discovery more efficient. Specify the standardised coding systems used in the dataset (e.g. ICD-10-CM, SNOMED CT). The coding systems controlled vocabulary (https://hdeu-dcat.acceptance.data.health.europa.eu/resource/authority/coding-system/) MUST be used.	This property allows you to specify the coding systems used within your dataset. For example, if a dataset uses ICD-10 for disease classification, this property allows data users to search for datasets with the same coding system. As a machine-actionable property, it also facilitates automated processes, making dataset discovery more efficient. Specify the standardised coding systems used in the dataset (e.g. ICD-10-CM, SNOMED CT, DRGs). RDF example: <code>healthdcatap:hasCodingSystem</code>	Kodeverk	Kodeverk favner i denne sammenheng bredere enn det man i helsevesenet normalt definerer som et kodeverk. Et kodeverk kan i denne sammenheng både representere en variabels verdsett/svaralterativ eller et standardisert (kontrollert) vokabular for å beskrive en egenskap. Den formelle definisjonen av et kodeverk er «en samling av begreper eller en klassifikasjon hvor hver tilhørende begrepsdefinisjon er tilknyttet en unik kode innenfor samlingen eller klassifikasjonen.	ICD-10 codes referenced in descriptions (e.g. C50, C34)		ICD-10, Minimum information standards /domain specific), etc.	
----------------------	---	--	---	----------	---	---	--	---	--

	RDF example: healthdcatap:has CodingSystem							
Contact point	Provides contact details for enquiries related to the dataset.	Contact information that can be used for sending comments about the Dataset.	This property informs about a contact point (Individual, Organization, Location, Group) that can answer questions about the dataset. Details on how to describe these are provided under class vcard:Kind. RDF example: dcat:contactPoint	Kontaktinformasjon: Epost, Telefonnummer, Hjemmeside	URL til hjemmeside, E-postadresse for henvendelser om datakilden, delkilden eller datasamlingen, Telefonnummer for henvendelser om datakilden, delkilden eller datasamlingen.	krq.data@fhi.no		MoBaadm@fhi.no
Conforms to	Identifies a standard or technical specification that defines the structural or syntactic organisation of the dataset as a whole.	If your data conforms to an established standard or specification, use this property to indicate which one. The Health Standards controlled vocabulary (https://hdeu-dcat.acceptance.data.health.europa.eu/resource/authority/standard/) MUST be used. RDF example: dct:conformsTo	If your data conforms to an established standard or specification, use this property to indicate which one. RDF example: dct:conformsTo	not included		could be considered in the future	could be considered in the future	

Custodian	Identifies the organisation responsible for holding, managing and maintaining the dataset over time (health data holder).	Party that accepts accountability and responsibility for the data and ensures appropriate care and maintenance of the resource [ISO-19115].	Use this property to identify the entity that technically holds and maintains the dataset. In the EHDS context, this field represents the legally defined health data holder, (Article 2(6) EHDS Regulation), i.e. the entity responsible for holding and maintaining the dataset. This entity may differ from the publisher. RDF example: geodcatap:custodian	Dataansvarlig	Navn på virksomheten eller enheten som eier og har det juridiske dataansvaret for datakilden, delkilden eller datasamlingen.	Norwegian Institute of Public Health	Norwegian Institute of Public Health
Description	Provides a free text summary describing the content and purpose of the dataset.	A free-text information about the Dataset.	Brief description of the dataset. This property can be repeated for including several language versions of the description. RDF example: dct:description	Beskrivelse	the same as in the HealthDCAT-AP	Norwegian Registry of Brain and Spinal Cord Tumors contains variables specific for diagnostics, treatment and follow-up of brain- and spinal cord tumors. The information comes from multiple sources, see information about the Cancer Registry.	The Norwegian Mother, Father and Child Cohort Study is a unique study where over 90,000 pregnant women were recruited from 1999 to 2008. More than 70,000 fathers have participated. Blood samples have been acquired from the children at the time of birth, from the mothers during pregnancy and after pregnancy.

Dataset distribution	Provides a technical representation or format in which the dataset is made available for secondary use.	An available Distribution for the Dataset.	For Health DCAT Application Profile a distribution is mandatory independently of the access rights. A distribution must include information on the Health Data Access Body supporting data access. RDF example: dcat:distribution	not included			to be implemented, could be considered in the future	The dataset will be distributed in in several formats depending on the technology used to acquire data (DNA, RNA, URINE, whole blood, plasma, serum), and depending on if the reserachers wants raw data or fully curated datasets.	The NIPH is responsible for data distribution
Geographic coverage	Indicates the geographic area to which the dataset relates (e.g. Member State, region or EU wide), where applicable.	A geographic region that is covered by the Dataset.	The EU Vocabularies Name Authority Lists MAY be used for continents, countries and places that are in those lists; if a particular location is not in one of the mentioned Named Authority Lists, Geonames URIs MAY be used. RDF example: dct:spatial	GeografiskAvgrensning	Opptaksområde for datakilden	Norge		All Norwegian Counties	
Personal data / Has personal data	Indicates whether the dataset contains personal data within the meaning	Key elements that represent an individual in the dataset, Indicates association with Personal Data.	This property is used to indicate the types of personal data contained in the dataset itself (i.e. the actual data, not the metadata record). The Data Privacy Vocabulary for	GradAv Personidentifikasjon	Høyeste grad av identifisering det er mulig å søke om fra denne datakilden, delkilden eller datasamlingen.	1 = Direkte personidentifiserbare opplysninger		Yes	

of Article 4, point (1), of Regulation (EU) 2016/679 of the European Parliament and of the Council².

Personal Data at <https://w3id.org/dpv/dpv-pd> MUST be used for this property.

Examples of allowed values include:
 Gender → <https://w3id.org/dpv/dpv-pd#Gender>
 Age → <https://w3id.org/dpv/dpv-pd#Age>
 RDF example:
 dpv:hasPersonalID ata

Health data access body	The competent health data access body responsible for facilitating access to the dataset under the European Health Data Space, where the dataset is subject to restrictive	Health Data Access Body supporting access to data in the Member State.	Member state-designated authority that facilitates the secondary use of electronic health data. HDABs assess the information provided by the health data applicant and decide on health data requests and access applications, authorise and issue data permits, obtain data from data holders and make data available in Secure Processing Environments. HDABs	not included			to be implemented	NIPH	
--------------------------------	--	--	---	--------------	--	--	-------------------	------	--

	d or non public access.		systematically track the data request and data access applications received and the data permits issued. As per Article 58 of the EHDS, HDABs are required to publicly list information on the data permits issued. (Article 55; Recital 52 of EHDS Regulation). RDF example: healthdcatap:hdab				
Health category / Health data category	Identifies the categories of electronic health data contained in the dataset and listed in Article 51(1) of Regulation (EU) 2025/327.	The health category to which this dataset belongs as described in the Commission Regulation on the European Health Data Space laying down a list of categories of electronic data for secondary use, Art.51.	This property indicates the EHDS data category the dataset belongs to, based on Article 51 (https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=OJ:L_202500327#ntr35-L_202500327EN.000101-E0035) of the EHDS Regulation. Use the controlled vocabulary http://13.81.34.152:1101/resource/authority/healthcategories/ to ensure standardised classification across all datasets	not included		to be implemented	Category Q, F and G.

RDF example:
healthdcatap:healthCategory

Health theme	Indicates a high level health topic or theme associated with the dataset (e.g. cardiovascular disease, cancer).	A category of the Dataset or tag describing the Dataset.	A dataset may be associated with multiple health themes. The Health Theme controlled vocabulary (http://13.81.34.152:1101/resource/authority/health-theme) MUST be used. RDF example: healthdcatap:healthTheme	AnbefalteBruksområder	En beskrivelse av hvilke typer formål (eksempelvis forskning, analyse, styringsdata, beredskap osv) og faglige temaer (eksempelvis kreft, hjerte og kar, kirurgi osv), en datakilde, delkilde og/eller datasamling er egnet for.	not really a match but the closest	Kreft, forskning, analyse, kvalitetssikring av helsetjenesten, styring, råd og veiledning om helsehjelp	to be implemented	Norwegian Environmental Biobank, part I: The importance of nutritional status for the effect of heavy metals on the health of mothers and their children (MoBa-ETox)
Identifier	Provides a unique and persistent identifier for the dataset (e.g. uniform resource identifier (URI), digital object identifier (DOI) or catalogue identifier).	The main identifier for the Dataset, e.g. the URI or other unique identifier in the context of the Catalogue.	For Health DCAT Application Profile the identifier is mandatory and its value must be unique from the origin database of the Dataset. In the process of dataset transmission to the next database, the 'identifier' property is overwritten by the new database identifier, and the original identifier becomes the value of the 'other identifier' property. RDF example: dct:identifier	code	Unik id for datasettet		K_KREG.CNS	to be clarified and implemented	MAB012, https://www.fhi.no/contentassets/c28c0ec2d4e24e36bbcb04ceeaba3d05/mab012.pdf

Structured data / Is structured	Indicates whether the dataset is structured, meaning that it is organised according to a predefined schema or model that defines its variables, attributes or fields.	Indicates whether the Dataset contains structured data for which a machine-readable description of the data variables can be provided.	Use this boolean property to state whether the dataset contains structured data. Set it to true when a machine-readable description of the data variables can be published, and to false when such structured variables do not apply. When the value is true, at least one healthdata:hasVariables should be provided.	not included			to be clarified and implemented. however we are not in favour of unstructured metadata/dataset	Is structured, typically related to the technology used, not necessarily to a predefined general schema	
Keyword	Provides keywords or tags describing the dataset and supporting search and discovery.	A keyword or tag describing the Dataset.	Add keywords to increase dataset discoverability. You can include keywords in different languages, submitting each keyword as a separate entry. Example: Physiological measures, Heart Rate, Stress Measures. RDF example: dcat:keyword	Emneord	Beskrivende emneord (tag). Skilles ved å bruke «;». F.eks. Emneord;Nytt emneord;	it is a voluntary property and not provided by the DataSource	to be clarified and implemented	Nutritional, stress, renal and metabolic disease biomarkers	

Language	Indicates the language or languages in which the dataset content is expressed.	A language of the Dataset.	For this property the values from the EU Vocabularies Languages Named Authority List must be used. If your Dataset contains multiple languages, this property can be repeated. RDF example: dct:language	not included		to be implemented	English	
Number of records	Indicates the total number of records contained in the dataset, where applicable.	Size of the dataset in terms of the number of records.	Use this property to indicate the total number of records in the dataset, giving users an idea of the dataset's size and whether it suits their analysis or research needs. You can provide the exact count or an approximate average if the exact figure is unavailable. It reflects the total data entries (e.g., consultations, prescriptions, lab results, hospital admissions) according to the dataset's content. RDF example: healthdcatap:numberOfRecords	not included		to be implemented		4

Number of unique individuals	Provides an estimate of number of unique natural persons represented in the dataset, where the dataset contains personal data and where applicable.	Number of records for unique individuals.	Use this property to indicate the number of unique individuals represented in the dataset. It helps users understand the dataset's scale and assess whether it meets their needs for secondary use. Provide the total or an approximate count of distinct individuals whose data is included. RDF example: healthdcatap:numberOfUniqueIndividuals	not included		to be implemented	3000	
-------------------------------------	---	---	---	--------------	--	-------------------	------	--

Provenance	Describe the origin, source or creation process of the dataset.	A statement about the lineage of a Dataset.	Information about how the data was collected, including methodologies, tools, and protocols used, this property can be repeated for different versions of the provenance (e.g. the provenance in different languages). RDF example: dct:provenance	innsamlingsmetode	Kort beskrivelse av metode for innsamling av data.
-------------------	---	---	--	-------------------	--

Kreftregisteret får rapportert og henter inn opplysninger fra ulike kilder: kliniske meldinger, patologisvar og kliniske systemer (stråle, medikamentell behandling), samt fra andre registre (NPR, DÅR, Folkeregisteret). Noe data kodes og registreres manuelt, noe importeres fra kildene

Urine samples were collected in urine cups at volumes of 8 ml from pregnant mothers and transferred to urine transport tubes (Becton-Dickinson (BD), Franklin Lakes, NJ, USA). Samples collected prior to 2003 were shipped in tubes without any bacteriostatic additive (BD, Plymouth, UK), while samples collected in 2003 and after were shipped in urine tubes containing chlorhexidine (UAP Vacutainers, BD, Franklin Lakes, NJ, USA). For more information on biological sampling, processing and storage, please refer to the original reference articles for NIPH's biobank by Rønningen et al. 2006 and Paltiel et al. 2014.

Publisher	Indicates the entity making the dataset available, which may differ from the creator or custodian.	An entity (organisation) responsible for making the Dataset available.	Use this property to identify the agent that publishes the dataset description in the catalogue. In the EHDS context, this may be the data holder or another organisation such as an intermediation entity. RDF example: dct:publisher	not included			to be implemented	Patterns and dietary determinants of essential and toxic elements in blood measured in mid-pregnancy: The Norwegian Environmental Biobank - PubMed	
Temporal coverage	Indicates the time period covered by the dataset (e.g. years of data collection), where temporal information exists.	A temporal period that the Dataset covers.	The start and end date of the period that the dataset covers. This property makes use of a Period of Time class. This property makes use of the class dct:PeriodOfTime (more details on PeriodOfTime class tab). The start and end of the interval SHOULD be given by using properties dcat:startDate or time:hasBeginning, and dcat:endDate or time:hasEnd, respectively. RDF example: dct:temporal	Opprettet	Året datakilden, delkilden eller datasamlingen ble opprettet og begynte å samle inn data.	not the best match although we have relevant properties on the datasource and the variable level - "from", "till"	Opprettet 2019	to be clarified and implemented. it is not really fitting our metadata management praxis	2002-2008, follow up data until 2015

Type	Indicates the general category or nature of the dataset (e.g. registry, administrative dataset, research dataset, synthetic dataset).	A type of the Dataset.	A Dataset may be associated with multiple types. The EU Dataset-type controlled vocabulary (http://publications.europa.eu/resource/authority/dataset-type) SHOULD be used. RDF example: <code>dct:type</code>	not included		to be implemented	Research dataset	
Variables	A structured description of the variables contained in the dataset, including, where applicable, the variable name, definition, data type, coding system or value domain	Links the Dataset to a CSVW TableGroup that provides a machine-readable description of the data variables.	Use this property to link a structured dataset to a CSVW TableGroup that documents its data variables, columns, or fields in a machine-readable way. This property becomes mandatory when <code>healthdcap:hasStructuredData</code> is true. RDF example: <code>healthdcap:hasVariables</code>	yes but not as a property	Dato for robotassistert kirurgi.	we have more advance use of variable and we expect more converge with HealthDC AT-AP in the future	Albumin (U-Alb-Mi) Potassium (U-K) Cortisol (U-Korsol) Creatinine (U-Krea) Sodium (U-Na) Uric acid (U-Uraat)	

Was generated by	Identifies the activity or process through which the dataset was generated (e.g. registry, clinical study, administrative process)	An activity that generated, or provides the business context for, the creation of the dataset.	The activity associated with generation of a dataset. The Health activity controlled vocabulary (https://hdeu-dcat.acceptance.data.health.europa.eu/resource/authority/health-activity) MUST be used. RDF example: prov:wasGeneratedBy	Opphav, etc.	we have several properties that are relevant		we expect more converge with HealthDC AT-AP in the future	Research	
Theme		A category of the Dataset.	A Dataset may be associated with multiple themes. The EU data-theme controlled vocabulary (http://publications.europa.eu/resource/authority/data-theme) MUST be used. RDF example: dcat:theme	Emneord	Beskrivende emneord (tag). Skilles ved å bruke «;». F.eks. Emneord;Nytt emneord;	it is a voluntary property and not provided by the DataSource	to be clarified and implemented	to be clarified and implemented	

Sweden

HealthDC AT-AP Dataset property name	Property description in implementing act Art. 77 draft	HealthDCAT-AP Dataset property definition (R7)	HealthDCAT-AP Dataset property usage note (R7)	NH DC pr op ert y na me	NHD C prop erty defi nitio n	Addi tiona l infor mati on	Example description on National cancer registry	Comments
Title	Provides the official name or title of the dataset.	A name given to the Dataset.	Provide a title(s) for your Dataset, which can be repeated in multiple languages. Example: Healthy Brain Study - Physiological Data RDF example: dct:title				National Cancer Register	

Access rights	Describes the level of access applicable to the dataset (e.g. public, restricted or non public).	Information that indicates whether the Dataset is publicly accessible, has access restrictions or is not public.	This property allows you to indicate whether the dataset is publicly accessible, has restrictions, or is not publicly available. This is a mandatory property in HealthDCAT-AP. Use one of the following values :public, :restricted, :non-public, which means: NON_PUBLIC — The dataset is not publicly available and access is subject to restrictions. PUBLIC — The dataset is openly accessible to everyone. RESTRICTED — The dataset has access restrictions but may be available under specific conditions. RDF example: dct:accessRights	Non-public
Applicable legislation	Identifies Union legal acts or national legislation that specifically govern the establishment, maintenance or availability of the dataset, where such legal acts or legislation exists.	The legislation that mandates the creation or management of the Dataset.	The ELI of the EHDS Regulation, published in March 2025, can now be used as the applicable legislation where relevant: http://data.europa.eu/eli/reg/2025/327/oj. As multiple legislations may apply to the resource the maximum cardinality is not limited. RDF example: dcatap:applicableLegislation	The Cancer Register is a health data register that is regulated by the Health Data Register Act (1998:543) and associated regulation (2001:709). The National Board of Health and Welfare's regulation and general advice (HSLF-FS 2016:7), shows in more detail how the healthcare provider's obligation must be fulfilled.

Code values	Identifies classification systems or controlled vocabularies reflected in the content of the dataset (e.g. disease classifications, procedure classifications), where applicable.	A free-text description of any health classification, terminology, ontology, thesaurus, or coding system used in the dataset.	Use this field to describe, in free text, the code values of the coding systems used that apply to the dataset (e.g. U07.1, Y59.0). No controlled vocabulary is required at this stage. RDF example: healthdcatap:hasCodeValues	Relevant code values from the codings system listed in Coding System	
Coding system	Identifies standardised coding schemes used to encode data elements within the dataset (e.g. ICD 10 CM, SNOMED CT, DRGs), where applicable.	Coding systems in use (ex: ICD-10-CM, DGRs, SNOMED=CT).	This property allows you to specify the coding systems used within your dataset. For example, if a dataset uses ICD-10 for disease classification, this property allows data users to search for datasets with the same coding system. As a machine-actionable property, it also facilitates automated processes, making dataset discovery more efficient. Specify the standardised coding systems used in the dataset (e.g. ICD-10-CM, SNOMED-CT). The coding systems controlled vocabulary (https://hdeu-dcat.acceptance.data.health.europa.eu/resource/authority/coding-system/) MUST be used. RDF example: healthdcatap:hasCodingSystem	ICD-7, ICD-9, ICD-O/2, ICD-O/3, UICC TNM, TOPO, PAD/C24.1/hist SHOULD BE FROM CT ONLY: ICD-O-3	https://hdeu-dcat.acceptance.data.health.europa.eu/resource/authority/coding-system/

Contact point	Provides contact details for enquiries related to the dataset.	Contact information that can be used for sending comments about the Dataset.	This property informs about a contact point (Individual, Organization, Location, Group) that can answer questions about the dataset. Details on how to describe these are provided under class vcard:Kind. RDF example: dcat:contactPoint	cancerregistret@socialstyrelsen.se
Conforms to	Identifies a standard or technical specification that defines the structural or syntactic organisation of the dataset as a whole.	An established standard to which the described resource conforms.	If your data conforms to an established standard or specification, use this property to indicate which one. The Health Standards controlled vocabulary (https://hdeu-dcat.acceptance.data.health.europa.eu/resource/authority/standard/) MUST be used. RDF example: dct:conformsTo	NA
Custodian	Identifies the organisation responsible for holding, managing and maintaining the dataset over time (health data holder).	Party that accepts accountability and responsibility for the data and ensures appropriate care and maintenance of the resource [ISO-19115].	Use this property to identify the entity that technically holds and maintains the dataset. In the EHDS context, this field represents the legally defined health data holder, (Article 2(6) EHDS Regulation), i.e. the entity responsible for holding and maintaining the dataset. This entity may differ from the publisher. RDF example: geodcatap:custodian	The Swedish National Board of Health and Welfare

Description	Provides a free text summary describing the content and purpose of the dataset.	A free-text information about the Dataset.	Brief description of the dataset. This property can be repeated for including several language versions of the description. RDF example: dct:description	The national cancer registry is an incidence based registry that contains information about diagnosed primary tumors. The registry contains information about all malign tumors and a subset of the benign tumors and tumor like conditions.	
Dataset distribution	Provides a technical representation or format in which the dataset is made available for secondary use.	An available Distribution for the Dataset.	For Health DCAT Application Profile a distribution is mandatory independently of the access rights. A distribution must include information on the Health Data Access Body supporting data access. RDF example: dcat:distribution	HDAB-SE	Access to data is conditional on the issuance of a permit by the HDAB after submission of a data request application
Geographical coverage	Indicates the geographic area to which the dataset relates (e.g. Member State, region or EU wide), where applicable.	A geographic region that is covered by the Dataset.	The EU Vocabularies Name Authority Lists MAY be used for continents, countries and places that are in those lists; if a particular location is not in one of the mentioned Named Authority Lists, Geonames URIs MAY be used. RDF example: dct:spatial	Sweden	
Personal data / Has personal data	Indicates whether the dataset contains personal data within the meaning of Article 4, point (1), of Regulation (EU) 2016/679 of the European Parliament and of the Council2.	Key elements that represent an individual in the dataset, Indicates association with Personal Data.	This property is used to indicate the types of personal data contained in the dataset itself (i.e. the actual data, not the metadata record). The Data Privacy Vocabulary for Personal Data at https://w3id.org/dpv/dpv-pd MUST be used for this property.	Gender, Age, Medical Health, Marital Status, Name, Region, Locality, Birth Date, PIN Code	

Examples of allowed values include:
Gender →
<https://w3id.org/dpv/dpv-pd#Gender>
Age →
<https://w3id.org/dpv/dpv-pd#Age>
RDF example:
dpv:hasPersonalData

Health data access body	The competent health data access body responsible for facilitating access to the dataset under the European Health Data Space, where the dataset is subject to restricted or non public access.	Health Data Access Body supporting access to data in the Member State.	Member state-designated authority that facilitates the secondary use of electronic health data. HDABs assess the information provided by the health data applicant and decide on health data requests and access applications, authorise and issue data permits, obtain data from data holders and make data available in Secure Processing Environments. HDABs systematically track the data request and data access applications received and the data permits issued. As per Article 58 of the EHDS, HDABs are required to publicly list information on the data permits issued. (Article 55; Recital 52 of EHDS Regulation). RDF example: healthcatap:hdab	Health Data Access Body Sweden (healthcatap:hdab-se)
--------------------------------	---	--	--	--

Health category / Health data category	Identifies the categories of electronic health data contained in the dataset and listed in Article 51(1) of Regulation (EU) 2025/327.	The health category to which this dataset belongs as described in the Commission Regulation on the European Health Data Space laying down a list of categories of electronic data for secondary use, Art.51.	This property indicates the EHDS data category the dataset belongs to, based on Article 51 (https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=OJ:L_202500327#ntr35-L_202500327EN.000101-E0035) of the EHDS Regulation. Use the controlled vocabulary http://13.81.34.152:1101/resource/authority/healthcategories/ to ensure standardised classification across all datasets RDF example: healthdcatap:healthCategory	Data from population-based health data registries such as public health registries (k, PHDR) (alternatively also Data from medical registries and mortality registries (l, MRMR)) Let's discuss w eg Norway.	
Health theme	Indicates a high level health topic or theme associated with the dataset (e.g. cardiovascular disease, cancer).	A category of the Dataset or tag describing the Dataset.	A dataset may be associated with multiple health themes. The Health Theme controlled vocabulary (http://13.81.34.152:1101/resource/authority/health-theme/) MUST be used. RDF example: healthdcatap:healthTheme	Cancer	
Identifier	Provides a unique and persistent identifier for the dataset (e.g. uniform resource identifier (URI), digital object identifier (DOI) or catalogue identifier).	The main identifier for the Dataset, e.g. the URI or other unique identifier in the context of the Catalogue.	For Health DCAT Application Profile the identifier is mandatory and its value must be unique from the origin database of the Dataset. In the process of dataset transmission to the next database, the 'identifier' property is overwritten by the new database identifier, and the original identifier becomes the	URI	We do not have an identifier as of yet. When a catalogue is built, this shall be present. NBHW uses OIDs for classifications and terminologies, but have not given their data unique identifiers.

			value of the 'other identifier' property. RDF example: dct:identifier		
Structured data / Is structured	Indicates whether the dataset is structured, meaning that it is organised according to a predefined schema or model that defines its variables, attributes or fields.	Indicates whether the Dataset contains structured data for which a machine-readable description of the data variables can be provided.	Use this boolean property to state whether the dataset contains structured data. Set it to true when a machine-readable description of the data variables can be published, and to false when such structured variables do not apply. When the value is true, at least one healthdatacatap:hasVariables should be provided.	True	
Keyword	Provides keywords or tags describing the dataset and supporting search and discovery.	A keyword or tag describing the Dataset.	Add keywords to increase dataset discoverability. You can include keywords in different languages, submitting each keyword as a separate entry. Example: Physiological measures, Heart Rate, Stress Measures. RDF example: dcat:keyword	Cancer, Tumour, Malignancy, TNM , Patology, Morfology, Topology, Tumor Stage, Cancer Incidence	
Language	Indicates the language or languages in which the dataset content is expressed.	A language of the Dataset.	For this property the values from the EU Vocabularies Languages Named Authority List must be used. If your Dataset contains multiple languages, this property can be repeated. RDF example: dct:language	Swedish	

Number of records

Indicates the total number of records contained in the dataset, where applicable.

Size of the dataset in terms of the number of records.

Use this property to indicate the total number of records in the dataset, giving users an idea of the dataset's size and whether it suits their analysis or research needs. You can provide the exact count or an approximate average if the exact figure is unavailable. It reflects the total data entries (e.g., consultations, prescriptions, lab results, hospital admissions) according to the dataset's content.
RDF example:
healthdcatap:numberOfRecords

4,927,237 (2024)

Number of unique individuals

Provides an estimated number of unique natural persons represented in the dataset, where the dataset contains personal data and where applicable.

Number of records for unique individuals.

Use this property to indicate the number of unique individuals represented in the dataset. It helps users understand the dataset's scale and assess whether it meets their needs for secondary use. Provide the total or an approximate count of distinct individuals whose data is included.
RDF example:
healthdcatap:numberOfUniqueIndividuals

3,242,090 (2024)

Provenance	Describes the origin, source or creation process of the dataset.	A statement about the lineage of a Dataset.	Information about how the data was collected, including methodologies, tools, and protocols used, this property can be repeated for different versions of the provenance (e.g. the provenance in different languages). RDF example: dct:provenance	Healthcare providers regularly send data about newly diagnose tumours to the 6 regional cancer centres for coding and registration and after quality control these data are forwarded yearly to National Board of Health and Welfare.
Publisher	Indicates the entity making the dataset available, which may differ from the creator or custodian.	An entity (organisation) responsible for making the Dataset available.	Use this property to identify the agent that publishes the dataset description in the catalogue. In the EHDS context, this may be the data holder or another organisation such as an intermediation entity. RDF example: dct:publisher	The Swedish National Board of Health and Welfare
Temporal coverage	Indicates the time period covered by the dataset (e.g. years of data collection), where temporal information exists.	A temporal period that the Dataset covers.	The start and end date of the period that the dataset covers. This property makes use of a Period of Time class This property makes use of the class dct:PeriodOfTime (more details on PeriodOfTime class tab) The start and end of the interval SHOULD be given by using properties dcat:startDate or time:hasBeginning, and dcat:endDate or time:hasEnd, respectively. RDF example: dct:temporal	1958–

Type	Indicates the general category or nature of the dataset (e.g. registry, administrative dataset, research dataset, synthetic dataset).	A type of the Dataset.	A Dataset may be associated with multiple types. The EU Dataset-type controlled vocabulary (http://publications.europa.eu/resource/authority/dataset-type) SHOULD be used. RDF example: dct:type	Register or Statistical	
Variables	A structured description of the variables contained in the dataset, including, where applicable, the variable name, definition, data type, coding system or value domain.	Links the Dataset to a CSVW TableGroup that provides a machine-readable description of the data variables.	Use this property to link a structured dataset to a CSVW TableGroup that documents its data variables, columns, or fields in a machine-readable way. This property becomes mandatory when healthdcatap:hasStructuredData is true. RDF example: healthdcatap:hasVariables	name (variabeletikett), definition (variabelbeskrivning), data type (datatyp), coding system (*see comment below. ICD-7, ICD-9, ICD-O/2, ICD-O/3, UICC TNM, TOPO, PAD/C24.1/hist), value domain (värdeområde) *NB: The disease classification is presented alongside the individual variables (e.g. variable name "ICD7") and the coding system is described in the related value domain (värdeområde)	https://www.socialstyrelsen.se/globalassets/sharepoint-dokument/dokument-webb/statistik/register-variabelforteckning-cancerregistret.xlsx
Was generated by	Identifies the activity or process through which the dataset was generated (e.g. registry, clinical study, administrative process).	An activity that generated, or provides the business context for, the creation of the dataset.	The activity associated with generation of a dataset. The Health activity controlled vocabulary (https://hdeu-dcat.acceptance.data.health.europa.eu/resource/authority/health-activity) MUST be used. RDF example: prov:wasGeneratedBy	HOSPITAL_RECORDS, OBSERVATIONAL_DATA	

Theme

A category of the Dataset.

A Dataset may be associated with multiple themes.
The EU data-theme controlled vocabulary (<http://publications.europa.eu/resource/authority/data-theme>) MUST be used.
RDF example: `dcat:theme`

HEAL

