Decision log regarding v2 Core Metadata schema Health-RI

Template:

|  |  |
| --- | --- |
| *Decision* | *[Short title for the decision]* |
| *Date of decision* | *[Date]* |
| *Context* | *[Description of the problem/opportunity and relevant constraints]* |
| *Options Considered* | * *[Option 1] - Pros: [...] - Cons: [...]* * *[Option 2] - Pros: [...] - Cons: [...]* * *[Option 3] - Pros: [...] - Cons: [...]* |
| *Decision* | *[The selected approach and brief explanation]* |
| *Consequences* | *[Impacts this decision will have on the schema and its users]* |
| *References* | *[Link to standards, discussions, or other resources]* |

|  |  |
| --- | --- |
| *Decision* | Choice to adhere to one flavor of access rights in HealthDCAT-AP |
| *Date of decision* | 29-11-2024 |
| *Context* | In fall 2024, HealthDCAT-AP introduced 3 types of cardinalities, based on which access rights apply to data, open, sensitive and protected. The more restrictions on a dataset, the more fields are mandatory in the Dataset class. |
| *Options Considered* | * [Implement all flavors] - Pros: Adhere to HealthDCAT-AP (draft) completely. - Cons: Implementation is a lot of work; FDP can currently not handle different versions of SHACLS. * [Implement 1 flavor: SENSITIVE] - Pros: This is the most strict one, meaning implementing this now and the other two later will enable backward compatibility (except any other changes in HealthDCAT-AP until official release of it) - Cons: Huge burden for data holders to provide all this metadata, it will be a huge effort when converting to v2. Also, some used vocabularies are not in place yet, so full compliance to HealthDCAT-AP is currently not possible. * [Implement 1 flavor: PROTECTED] - This option was not considered. * [Implement 1 flavor: OPEN]  - Pros: Least burden for data holders for now to supply the required fields. Switching to HealthDCAT-AP fields is already a challange, and we don't want to require data holders to supply fields now that potentially might not end up in the final version of HealthDCAT-AP.  - Cons: Backward compatibility issues when we introduce the other 2 flavors which require more mandatory fields. |
| *Decision* | It was decided to adhere to the OPEN version to be compliant to in v2. |
| *Consequences* | No full compliance to HealthDCAT-AP current draft (or snapshot taken for v2 HRI). It is important to note that at the moment of writing this/release of v2, HealthDCAT-AP is still not finished, and changes can still occur in the schema. |
| *References* | HealthDCAT-AP: <https://healthdcat-ap.github.io/#application-profile-diagram>  Notes from modelling sessions: [RollingMinutes\_ModellingSessions.docx](https://healthri.sharepoint.com/:w:/r/sites/hri-team022/Shared%20Documents/05%20-%20FAIR/Workplan%20%26%20Activities/2024/1%20Metadata%20standards/P2%20metadata%20schema/Meeting%20notes%20and%20slides/RollingMinutes_ModellingSessions.docx?d=w95105f6aa4174bcf91a4e5cd6c2b3237&csf=1&web=1&e=uu5oUe) |

|  |  |
| --- | --- |
| *Decision* | Choice to not include several fields from HealthDCAT-AP in the v2 model |
| *Date of decision* | 11-12-2024 |
| *Context* | Some properties can currently not be filled but are mandatory in HealthDCAT (eg. Applicable legislation (ELI of EHDS not ready), healthCategory (controlled vocabulary is not public yet). |
| *~~Options Considered~~*  *Which fields does this concern? And what is the specific reason not to take up?* | Class Dataset:   * healthdcatap:hdab – There is no HDAB appointed yet in NL. * healthdcatap:healthCategory – This field refers to the categories as defined in art. 51 EHDS. It requires the use of a controlled vocabulary which is currently not yet created. See table of HealthDCAT-AP vocabularies below table in section linked [here](https://healthdcat-ap.github.io/#controlled-vocabularies-to-be-used).   Class Agent:   * healthdcatap:publisherType - This field refers to the types of publishers as discussed in the TEHDAS2 program. It requires the use of a controlled vocabulary which is currently not yet created. See table of HealthDCAT-AP vocabularies below table in section linked [here](https://healthdcat-ap.github.io/#controlled-vocabularies-to-be-used). |
| *Decision* | Decision to not take up fields that can currently not be filled. |
| *Consequences* | No full compliance to HealthDCAT-AP current draft (or snapshot taken for v2 HRI). It is important to note that at the moment of writing this/release of v2, HealthDCAT-AP is still not finished, and changes can still occur in the schema. |
| *References* | HealthDCAT-AP: [https://healthdcat-ap.github.io/](https://healthdcat-ap.github.io/#application-profile-diagram)  Notes from modelling sessions: [RollingMinutes\_ModellingSessions.docx](https://healthri.sharepoint.com/:w:/r/sites/hri-team022/Shared%20Documents/05%20-%20FAIR/Workplan%20%26%20Activities/2024/1%20Metadata%20standards/P2%20metadata%20schema/Meeting%20notes%20and%20slides/RollingMinutes_ModellingSessions.docx?d=w95105f6aa4174bcf91a4e5cd6c2b3237&csf=1&web=1&e=uu5oUe) |

|  |  |
| --- | --- |
| *Decision* | Choice to diverge from specifications of HealthDCAT-AP and/or DCAT-AP NL for specific properties |
| *Date of decision* | Various modelling sessions. |
| *Context* | The v2 model is intended to be both compliant to DCAT-AP NL and HealthDCAT-AP. This already poses a small challenge, since these two models are not 100% compatible. In these cases, the HRI model should adhere to the strictest use of cardinalities where they are not the same in DCAT-AP NL and HealthDCAT-AP. However, this makes the HRI model even more strict. This is why for some cases, the v2 HRI model diverges from either DCAT-AP NL or HealthDCAT-AP restrictions. |
| *~~Options Considered~~*  *Which fields does this concern? And what is the specific reason not to take up?* | Next to the classes named in the previous issue, v2 core Health-RI model diverges from some restrictions/specifications in the following properties of DCAT-AP v3, DCAT-AP NL or HealthDCAT-AP:   * dct:creator (both Catalog and Data Service classes) - DCAT-AP v3/ DCAT-AP NL / HealthDCAT-AP restrict to max 1, but HRI v2 model allows many. * dct:language (Distribution class): DCAT-AP NL restricts to max 1 language per Distribution. HRI v2 model allows many, since there can be multiple languages present in the same Distribution. * dcat:distribution (Dataset class): HealthDCAT-AP requires at least one distribution (also in the OPEN flavor of the profile), but HRI v2 model has this recommended. Not all datasets have a Distribution ready to be described. This is why it was decided to not make this property (and therefore the class Distribution) mandatory. |
| *Decision* | See above per field. |
| *Consequences* | No full compliance to HealthDCAT-AP current draft (or snapshot taken for v2 HRI). It is important to note that at the moment of writing this/release of v2, HealthDCAT-AP is still not finished, and changes can still occur in the schema. |
| *References* | HealthDCAT-AP: [https://healthdcat-ap.github.io/](https://healthdcat-ap.github.io/#application-profile-diagram)  Notes from modelling sessions: [RollingMinutes\_ModellingSessions.docx](https://healthri.sharepoint.com/:w:/r/sites/hri-team022/Shared%20Documents/05%20-%20FAIR/Workplan%20%26%20Activities/2024/1%20Metadata%20standards/P2%20metadata%20schema/Meeting%20notes%20and%20slides/RollingMinutes_ModellingSessions.docx?d=w95105f6aa4174bcf91a4e5cd6c2b3237&csf=1&web=1&e=uu5oUe) |

|  |  |
| --- | --- |
| *Decision* | *Decision on modelling Dataset Series* |
| *Date of decision* | *04-02-2025* |
| *Context* | *In Version 1 of the metadata model, Dataset Series was modelled according to DCAT, where it’s stated that Dataset Series is a subclass of Dataset. This means Dataset Series inherits all the properties from Dataset. We decided that this is not a good way to do it for the version 2, since some new Dataset properties don’t make sense to include in Dataset Series. This is why we decided to model it according to DCAT-AP, where Dataset Series is a subclass of DCAT:Resource.* |
| *Options Considered* | * *[Model Dataset Series according to DCAT] - Pros: [Compatible to DCAT, less time consuming for implementation, since the properties are the same as in Dataset] - Cons: [Does not make sense with the new properties from HealthDCAT-AP, it’s also not compliant to DCAT-AP and all its extensions.]* * *[Model Dataset Series according to DCAT-AP] - Pros: [Dataset Series wouldn’t inherit Dataset’s properties, which makes more sense and is easier to fill in by data providers. It also makes it compatible to DCAT-AP and its extensions] - Cons: [Not compliant to original DCAT]* |
| *Decision* | *We decided to model Dataset Series according to DCAT-AP, so that it is a subclass of dcat:Resource, instead of dcat:Dataset. This makes more sense considering newly added properties from healthDCAT-AP and makes it compliant to DCAT-AP and its extensions.* |
| *Consequences* | *Possible issues with compliancy to the version 1.* |
| *References* | [*See modeling session minutes*](https://healthri.sharepoint.com/:w:/s/hri-team022/EWpfEJUXpM9LkaTlzWwrMjcBRhC1VcuE5fk7B4LytkOO4w?wdLOR=c6183BF6B-15A9-8A46-A089-63094A7D65D1) |

|  |  |
| --- | --- |
| *Decision* | Base the metadata model on HealthDCAT-AP draft, version of 16-12-2024 |
| *Date of decision* | *11-12-2024* |
| *Context* | *Since HealthDCAT-AP is still a draft and being updated, we decided to take a snapshot of the version on a specific date and adhere to that.* |
| *Options Considered* | * *[Decide to use the version of HealthDCAT-AP of a specific date] - Pros: [We don’t need to chase after every change made in the HealthDCAT-AP documentation] - Cons: [Our model is not up to date and possibly not compliant to the latest version of HealthDCAT-AP.]* * *[Continuously update our model according to the changes in HealthDCAT-AP draft] - Pros: [We stay compliant to HealthDCAT-AP] - Cons: [A lot of work, since there is not yet a decision log or proper versioning of HealthDCAT-AP. Could also lead to redundant work, since they could add a change and retract it later.]* |
| *Decision* | *We decided to take a snapshot of the HealthDCAT-AP specification on* 16-12-2024 and base our model on that. This means having more stability and reducing meaningless work. |
| *Consequences* | *Our schema is lacking the updates made in HealthDCAT-AP after the decided date.* |
| *References* | [*HealthDCAT-AP*](https://healthdcat-ap.github.io/)  [*Rolling minutes*](https://healthri.sharepoint.com/:w:/s/hri-team022/EWpfEJUXpM9LkaTlzWwrMjcBRhC1VcuE5fk7B4LytkOO4w?wdLOR=c6183BF6B-15A9-8A46-A089-63094A7D65D1) |

|  |  |
| --- | --- |
| *Decision* | Keep dct:identifier in foaf:Agent a literal and allow identifiers from different systems. |
| *Date of decision* | *23-04-2025* |
| *Context* | *The foaf:Agent class contains the mandatory property dct:identifier to specify an identifier with the Agent (Publisher, Creator,...). Since the range is a literal, this is not great for interoperability. To increase interoperability, it was suggested to change the range to an URI, since most identifiers (like ROR, ORCID) are URIs anyways.*  *In addition, some institutes do not work with the so far proposed identifier system (ORCID), so there was a need to explore other options.* |
| *Options Considered* | * We tried using *xsd:anyURI* as data type for the dct:identifier. However, this led to issues in the FDP when entering metadata. Therefore, it was decided to revert this change and keep the *dct:identifier* as a literal as before. * Concerning alternatives for ORCID, we did some research for alternative sources and came up with some options (see next point).   + To allow for multiple identifiers, the cardinality should be changed to 1..n   + Since it is not originally included in DCAT/NL/Health, we can relax the max nr of values to n, since there are no restrictions from the original schemas |
| *Decision* | * *dct:identifier in foaf:Agent class will stay a literal* * *Concerning alternatives to ORCID, these are the proposed alternatives:*   + [ORCID](https://orcid.org/) - for people; commonly used by researcher but no identity check; it's possible to create multiple IDs for one person   + [ISNI](https://isni.org/) - for organizations and people   + [ROR](https://ror.org/) - for research organizations   + [TOOI](https://standaarden.overheid.nl/tooi/waardelijsten/) - Dutch public organizations / governmental bodies   + [Wikidata](https://www.wikidata.org/wiki/%5BWikidata%5D(https://www.wikidata.org/wiki/Wikidata:Main_Page):Main_Page) - for anything; last resort; temporary solution until the organization creates an ID |
| *Consequences* | * *For now, the identifier is still not interoperable, but we can at least use it in FDP* * *We provide alternative options for ORCID and allow multiple identifiers per agent* |
| *References* |  |