Meeting Minutes: Webinar dedicated to the status and governance of DCAT-AP (SEMIC - A04.02)

Project:	SEMIC: DCAT-AP (HVD)	Date and Time:	07/11/2023 14:00 - 16:00
Meeting Type:	Webinar	Location:	Virtual
Coordinators:	Pavlina Fragkou Makx Dekkers Bert Van Nuffelen	Issue Date:	15/11/2023

Agenda of the webinar		
14:00 - 14:15	Welcome	Slides 1 - 9
14:15 - 15:40	Issues related to DCAT-AP HVD	Slides 10 - 30
15:40 - 15:55	Upcoming & ongoing activities within the DCAT-AP ecosystem	Slides 31 - 42
15:55 - 16:00	Wrap-up and next steps	Slides 42 - 47

Meeting Slides
LINK

Participants		
Name	**Initials**	**Organisation**
Anastasia Sofou	AS	SEMIC Team
Anssi Ahlberg	AA	Finnish Digital Agency
Alexandra Balahur	AB	SEMIC Team
Arthur Schiltz	AS	SEMIC Team

Participants		
Name	**Initials**	**Organisation**
Bart Hanssens	BH	DG Digital Transformation
Bert Van Nuffelen	BVN	SEMIC Team
Casper	/	/
Cécile Guasch	CG	DIGIT
Charles-Andrew Vande Catsyne	CAVC	Sciensano, Belgium
Christian Wittig	CW	GovData, Germany
Claudio Baldassarre	CB	DIGIT
Coen Janssen	CJ	CNECT
Dancia Saponja	DS	Ministry of Digital Transformation, Slovenia
David	/	/
Estelle Maudet	EM	Data.gouv.fr, France
Emidio Stani	ES	SEMIC Team
Feilicat Loeffler	FL	Thüringer Finance Ministry, Germany
Florian Barthelemey	FB	SEMIC Team
Gabriel Ripoche	GR	Ministry of Ecology, Energy & Territories, France
Geraldine Nolf	GN	Digital Flanders, Belgium
Gery	/	/
Hagar Lowenthal	HL	JRC
Jakub Klímek	JK	Ministry of the Interior of the Czech Republic
Jill Saligoe-Simmel	JSS	Esri, United States
Jim J. Yang	JJY	Norwegian Digitalisation Agency (Digdir)
Joachim Nielandt	JN	Digital Flanders, Belgium
Kees Nieuwstad	KN	National Road Traffic Data Portal, Netherlands
Lars-Inge Arnevik	LIA	Norwegian Mapping and Cadastre Authority

Participants		
Name	**Initials**	**Organisation**
Licinio Kustra Mano	LKM	DG SANTE
Ludger Rinsche	LR	]init[, Germany
Maik Roth	MR	Open Government Data Secretariat, Swiss
Makx Dekkers	MD	SEMIC Team
Marceau Louis	ML	Public Service of Wallonia, Belgium
Marie-Bénédicte	/	/
Marta	/	/
Martin Alvarez Espinar	MAE	Huawei Technologies
Matthias Palmér	MP	MetaSolutions AB
Michal Kuban	MK	CNECT
Mihai Paunescu	MP	Publications Office
Morgana Andrade	MA	University of Minho, Portugal
Nuno Freire	NF	Europeana
Oystein Asnes	OA	Norwegian Digitalisation Agency (Digdir)
Pascal Derycke	PD	Sciensano, Belgium
Pavlina Fragkou	PF	DIGIT
Pierlou Ramade	PR	Data.gouv.fr, France
Raphael Sturm	RS	Fraunhofer, Germany
Sander Van Dooren	SVD	Digital Flanders, Belgium

Points discussed and decisions taken

Topic discussed	Outcome
Issues related to DCAT-AP HVD	
Issue #307: Distribution & HVD category	Approved: - Keep HVD category mandatory for Dataset. To be discussed on GitHub: - Keep HVD category mandatory for Data Service. Not approved: - Keep HVD category optional for Distribution. - Approved instead: remove HVD category from Distribution.
Issue #310: Risk of overlap for quality of service with generic documentation on data services	Approved: - Add property documentation mapped to foaf:agent to Data Service in DCAT-AP; - Replace the label in DCAT-AP HVD with documentation; - Add a usage note: in the context of HVD quality of service information is documented here.
Issue #305: Publisher property Data Service	Approved: - SEMIC proposes that this is a discussion to be elevated from the DCAT-AP HVD to the DCAT-AP level.
Issue #283: Incorrect usage of vCard	Approved: - SEMIC proposes that this is a discussion to be elevated from the DCAT-AP HVD to the DCAT-AP level.
Issue #282: What is the exact meaning of the mandatory applicableLegislation on a mixed catalogue	Not approved: - Increase the cardinalities at the Catalogue level, but create a new subclass ReportingCatalogue. - Approved instead: the property should not be kept at the Catalogue level at all.

Full Meeting Minutes

Welcome & Introduction Slides 1 - 9 **Speaker:** Pavlina Fragkou	Welcome word from PF and an introduction of AB as the new programme manager for SEMIC. Agenda: - Welcome - Resolutions on DCAT-AP HVD issues - Upcoming & ongoing activities within the DCAT-AP ecosystem - Wrap up and next steps A short introduction and setting of the context is done by PF.
Issues related to DCAT-AP HVD Slides 10 - 30 **Speaker:** Bert Van Nuffelen	BVN presents a timeline of DCAT-AP 3.0 and DCAT-AP HVD. The development started in February 2023, with a large public review over the summer. Today we will start to dive into the output of the public review. The review lasted from 18 July 2023 until 30 September 2023. During this period a total of 102 issues were reported. The objective is to address approximately 80+ issues with the release of DCAT-AP 3.0 and DCAT-AP for HVD. Of these 102 a total of 25 issues were tagged with the label HVD. 19 of them can be easily resolved with editorial fixes. 6 of them were not editorial in nature and will be addressed in this series of webinars. 1) **Issue #307:** Distribution & HVD category The question was raised for which entities it is necessary to express the HVD category. The proposition was motivated by making the HVD category optional for a Distribution. However, the HVD category should be maintained on the level of a Data Service. **SEMIC Proposition:** SEMIC proposes to: - Keep the HVD category mandatory for Dataset; - Keep the HVD category mandatory for Data Service; - Make the HVD category optional for Distribution. **Resolvement:** - It is agreed upon to keep the HVD category mandatory for Dataset. - Whether the HVD category remains mandatory for Data Service needs further investigation on GitHub & will hopefully be resolved by the webinar 21st of November. - Moreover, the Working Group agrees on the removal of the HVD category for Distribution. **Discussion:**

MP suggests removing the HVD Category from Distribution. Casper agrees with MP and mentions it might cause more confusion than solve problems. Also, JK & AA agree with the suggestion for removal from MP.

SVD wonders whether it is even necessary to include an HVD property on Dataset. In the catalogue you can already indicate whether a Dataset is HVD through theme:Taxonomy in a skos:ConceptScheme. BVN responds that these properties have been introduced for two purposes. The first property, ApplicableLegislation, is to identify with which legislation a catalogued resource is compliant. The cardinality is explicitly left open so that it can be used for multiple legislations, other than HVD IR. The second reason is the need to have an HVD category comes from the implementing regulation.

The first property, ApplicableLegislation, is to identify with which legislation a catalogued resource is compliant. The cardinality is explicitly left open so that it can be used for multiple legislations.

HL thinks it should not be mandatory for Data Service. She argues that it is the Dataset itself that is high-value, but not the Data Service because a Data Service can serve many Datasets. Therefore, it creates ambiguity to the existence of the Data Service.

MP adds that this raises the question whether a Data Service exists independently and is described as an entity manipulating a Dataset or to serve a Dataset in a dependent manner. If it is connected to the Dataset then it can be inferred to be HVD, otherwise not and the applicable legislation is also necessary for the Data Service. BVN makes the assumption that we currently view a Data Service as being an existing, independent entity and therefore we add this HVD category. The issue evidently exists when there is a Data Service serving two Datasets where one is HVD and one is not. MD expresses the scope of this issue and that we cannot answer such a question today. It will need further discussion under a new GitHub issue.

GN asks why we cannot handle the HVD category as a dcat:theme subproperty as is done by the INSPIRE categories for example. She argues that it adds another layer of complexity. This need comes directly from the implementing regulation, BVN responds. It is not a thematic topic that is aligned with the existing themes such as INSPIRE.

Several participants agree with Geraldine. They argue it would make technical implementation easier.

MP does not agree to work with dcat:theme, because it will make it harder to know what to look for. He would prefer to keep dct:theme clean. Creating different properties for all specific needs would defy the idea of interoperability.

This discussion will also continue on GitHub under a new issue.

2) **Issue [#310](#):** Within the context of HVD there exists the need to express the quality of Data Services, such as APIs. This creates the risk of overlap for quality of services with generic documentation on Data Services.

SEMIC Proposition:
- Add property *documentation* mapped to *foaf:agent* to Data Service in DCAT-AP;
- Replace the label in DCAT-AP HVD with documentation;
- Add a usage note: that in the context of HVD quality of service information is documented here.

Resolvement:
The proposal is approved by the Working Group.

Discussion:
MP argues that we do not have a regulatory requirement to separate the quality of service from other documentation. BVN responds that there is a requirement for expressing this kind of information, but it is rather vague and nonspecific. The proposed approach is a generic catch-all approach to address this issue. MP agrees to the proposition.

GN mentions that in the INSPIRE world they express a certain level of quality through a conformsTo whether the Data Service adheres to this particular level of quality. From a regulatory standpoint such a minimum level of quality does not exist, BVN responds. It is domain specific and therefore could be addressed within a domain.

3) **Issue [#305](#):** Within the HVD context, a Data Service is a catalogued resource at the level of a Dataset, where there is a need for a publisher. The question is whether a Data Service should not have the same property. This is, however, not a regulatory requirement.

SEMIC Proposition:
SEMIC proposes that this is a discussion to be held in DCAT-AP, not DCAT-AP HVD. Once resolved in DCAT-AP, DCAT-AP HVD will follow the same resolution. The issue on the level of DCAT-AP HVD is closed.

4) **Issue [#283](#):** This issue concerns the incorrect usage of the range of vCard in terms of contact points. A second part of this issue concerns the usage of abstract versus concrete subclasses of vCard.

   SEMIC Proposition:
   SEMIC proposes to solve this issue on the level of DCAT-AP, once resolved there DCAT-AP HVD will adopt the resolution. The issue on the level of DCAT-AP HVD is closed.

5) **Issue [#282](#):** This issue addresses the exact meaning of the mandatory property applicableLegislation on a Catalogue.
   A use case to consider is that the Catalogue is a reporting catalogue. This is a Catalogue that is supplied to the auditing authority and contains all information about the Member State's HVD. Another case is that the Catalogue is a harvesting catalogue, which is used to indicate that some information in the Catalogue is in the scope of the HVD implementing regulation. This information is useful when harvesting operations are performed, which is why it received this name.

   SEMIC Proposition:
   SEMIC proposes to increase the cardinalities at the Catalogue level, but create a new subclass ReportingCatalogue where HVD reporting is mandatory. This allows an indication that HVD information is included at the Catalogue level, but does not indicate that all information in the Catalogue is within HVD scope.

   Resolvement:
   The Working Group does not approve of the incorporation of ReportingCatalogue. The Working Group instead agrees that the property should not be kept at the Catalogue level at all.

   Discussion:
   GN wonders if this cannot be done in the HVD Category that was previously discussed. Otherwise, the legislation has to be more than HVD only, like for example containing the legislations of INSPIRE, Open data, and all the other directives. If it has all directives and regulations, then we can use this because it will adhere to the once-only principle and in harvesting & reporting it could be filtered on. BVN responds that the ReportingCatalogue is a special case, it could be a query on the top level.

   Why not model several links between legislation and catalogue, asks CG. Qualifying the links as Catalogue complies with Legislation, Catalogue fulfils obligation, could help in describing the logic of the applicable legislation. BVN replies that currently it does not address these subtle distinctions and that it needs further investigation if such a need exists.

MP mentions that if you have a theme or a keyword that applies to the whole Dataset & the Distributions, then from the same perspective if there is an applicableLegislation of a Catalogue it also applies to the Datasets within. It would make more sense to add applicableLegislation on the catalogue level manually when it is a Catalogue that is solely constructed from a reporting perspective of HVD. For the harvesting approach, there has to be a rule to detect something, such as whenever there is a dataset inside a catalogue that is an HVD you should also add the property on the level of the catalogue. Such a rule would require implementation. He argues that the ReportingCatalogue at the Catalogue level would be better than SEMIC's proposition and the Catalogue is restricted to reporting only.

The sense that it has to be at the level of the resources, such as MP argues, is something GN agrees with. However, this property is at the level of the resource, that is the Dataset, already. This discussion is about the meaning at the level of the Catalogue, MD responds.

LKM inquires how this can be applicable to datasets from the eHealth Data Space and if they can use it. BVN & MD confirm that the HVD context can be used there as well.

Whether the reporting catalogue could be something that is instantiated under health regulation is a question CG adds to the discussion. Because now the reporting catalogue is considered as a class that is only used in the context of HVD because of the implementing regulation. However, in other domains such a need might exist from other regulations as well. However, such a need for reporting within the health domain does not exist, BVN replies. It is however true that ReportingCatalogue is only applicable to HVD that are related to the HVD implementing regulation.

SVD suggests adding another property reportsOnLegislation to a Catalogue, instead of applicableLegislation, that indicates that this is only necessary in the case of reporting.

The question whether an HVD mandates a specific type of Catalogue is raised by CG, otherwise we should take this discussion to the resource level.

BVN expresses that the implementing regulation requires that every dataset needs specifically expressed metadata, the existing metadata in the Open Data portals can be augmented with quality checks and information that make it compliant with HVD regulations. Then there is the part on reporting, where it should be clear what the difference between metadata supplied

as a report and generic metadata that is related, but not necessary for the reporting, is. This is the reason for the existence of the ReportingCatalogue.

MP argues that we can actually generate this report from the metadata that is in the Catalogue, therefore we do not need this ReportingCatalogue. Several participants agree with this. To which, CG adds that we can derive which datasets need such regulation and reporting with a simple query.

6) **Issue #312:** Which additional information is provided, too many properties listed?
Many of the properties listed in DCAT-AP HVD do not add any additional requirement compared to DCAT-AP 2.x and 3, this makes the specification longer than necessary.

This issue relates to two audiences:
   - The HVD policy officers
   - The DCAT-AP catalogue maintainer

The reason is that HVD policy officers may not have had previous exposure to DCAT-AP. These officers will look for implementation regulation terms in the metadata and therefore are included in the HVD annex. For DCAT-AP catalogue maintainers, only the additional information compared to DCAT-AP is relevant.

SEMIC Proposition:
SEMIC proposes to align with the profiling guidelines of the next webinar and improve connecting the terms with the HVD implementation regulation. At the same time the goal is to keep the list of properties to a minimum to limit interdependency and keep the annex version-agnostic as much as possible. The goal is to make an editorial that is stable and needs as little as possible alignment when it comes to different DCAT-AP versions and extensions. After the next webinar the rewrite will start, considering the comments of the community on GitHub issue #312.

7) The last 'issue' is a question from SEMIC towards the community regarding domain versus general metadata requirements. The HVD implementing regulation requests the metadata according to the best practices in each domain. However, DCAT-AP HVD only includes generic metadata and not domain specific metadata. Because each domain must conform to the HVD implementing regulation which requires expression of metadata according to DCAT-AP HVD, this might pose problems with respect to current best practices.

SEMIC Proposition:

	SEMIC proposes to use DCAT-AP HVD as a reference point and attempt to align and combine it with best practices. DCAT-AP HVD should be seen as the aggregation language for all metadata to be supplied for reporting. **Discussion:** GN suggests creating guidelines separated by domain, and then you can express whether you conform through a conformity statement. That way you do not have to put everything in the metadata, but you can have it in the guidelines. This could be a solution in this case. It is a form of self-declaration. BVN suggests discussing this further on GitHub. If the community prefers this method it could be incorporated. For the finalisation of the annex: - All input will be consolidated & processed; - Deadline to provide additional input via GitHub on the discussed issues is 01/12/2023; - Based on the above, we will prepare the SEMIC recommendations of DCAT-AP for HVD, which will be presented to the PSI Expert Group on 14/12/2023; - Next steps after the PSI Expert Group are listed in the 'Upcoming & ongoing activities on other DCAT-AP extension'.

Upcoming & ongoing activities on other DCAT-AP extensions
 	
Slides 31 - 42
 	
Speaker: Bert Van Nuffelen
	State of play
	SEMIC aims to align and make metadata reusable in multiple contexts.
	
	Next webinar
	53 issues of the total of 102 issues on DCAT-AP are technical in nature. 16 of those will be discussed in the next webinar.
	
	They are categorised in three levels:
	- Cross concerns issues
	  - Guidelines for building DCAT-AP profiles
	  - Guidelines for profiling themes and other properties to enhance the readability and the interoperability of profiles.
	- Detailed issues
	  - ContactPoint, AgentTypes
	  - Change Type property
	  - Cardinality conformsTo CatalogueRecord
	  - Dataservice properties
	- Alignment DCAT 3.0 issues
	  - Dataset series: ordering, containment
	  - Inverse properties
	  - Catalogue composition
	
	GeoDCAT-AP
	GeoDCAT-AP 3.0 is announced as a collaboration of the INSPIRE

	community and SEMIC supported by DG ENV, JRC and SEMIC.
	A new alignment activity has been taken up both in terms of HVD and GeoDCAT-AP.
	StatDCAT-AP StatDCAT-AP has ongoing internal activities of SEMIC to assess how StatDCAT-AP and DCAT-AP 3.0 and DCAT-AP HVD can be aligned. However, no immediate actions are planned.
	BRegDCAT-AP BRegDCAT-AP will be integrated as an annex similar to DCAT-AP HVD as was agreed upon on the 21st of April 2023.
	MLDCAT-AP MLDCAT-AP is a pilot for describing metadata related to machine learning models, datasets associated with them, algorithms, etc. DCAT-AP was used to make a base data model for this purpose.
Wrap-up and next steps Slides 42 - 47 **Speaker:** Bert Van Nuffelen, Pavlina Fragkou	Participants are invited to the next webinar and a reminder to continue discussions on the GitHub issues to be discussed in the next webinar is extended. PF thanks the working group for the fruitful discussion and she also extends an invitation for the next webinar.