Meeting Minutes: Webinar technical issues on DCAT-AP (SEMIC - A04.02)

Project:	SEMIC	Date and Time:	21/11/2023 14:00 - 16:00
Meeting Type:	Webinar	Location:	Virtual
Coordinators:	Pavlina Fragkou Bert Van Nuffelen Makx Dekkers	Issue Date:	28/11/2023

Agenda of the webinar		
14:00 - 14:10	Welcome & introduction	Slides 1 - 9
14:10 - 14:35	DCAT-AP 3.0 issues	Slides 10 - 19
14:35 - 15:00	Dataset series	Slides 20 - 30
15:05 - 15:15	Catalogues	Slides 31 - 33
15:15 - 15:45	DCAT-AP profile building	Slides 34 - 59
15:45 - 15:55	Some proposals	Slided 60 - 67
15:55 - 16:00	Next steps	Slides 68 - 71

Meeting Slides
LINK

Participants		
Name	Initials	Organisation
Adam Anrdt	AA	Danish Agency for Digitalisation
Agata Majchrowska	AM	European Commission
Alberto Abella	AB	Fiware, Spain
Anastasia Sofou	AS	SEMIC Team

Participants		
Name	**Initials**	**Organisation**
Andrea Perego	AP	European Parliament
Andres Garcia	AG	/
Bart Hanssens	BH	BOSA, Belgium
Bert Van Nuffelen	BVN	SEMIC Team
Bharat Sharma	BS	National Health Service, UK
Charles Andrew Vande Catsyne	CAVC	Sciensano, Belgium
Costas Simatos	CS	Sigma Cubed, Belgium
Csongor Nyulas	CN	Stanford University, US
Darius Amilevicius	DA	Information Society Development Committee, Lithuania
Dietmar Gattwinkel	DG	DG CNECT
Emidio Stani	ES	SEMIC Team
Estelle Maudet	EM	Data.gouv.fr, France
Eva Thelisson	ET	Federal Roads Office, Switzerland
Fabian Kirstein	FK	FOKUS, Fraunhofer, Germany
Fredrik Persäter	FP	Swedish Mapping, Cadastre and Land Registration Authority
Geraldine Nolf	GN	Digital Flanders, Belgium
Hagar Lowenthal	HL	JRC
Isabelle Baber	IB	Federal Statistics Office, Switzerland
Jakub Klímek	JK	Ministry of the Interior of the Czech Republic
Jari Reini	JR	National Land Survey, Finland
Jitse De Cock	JDC	SEMIC Team
Kety Giuliacci	KG	National Institute of Geophysics and Volcanology, Italy

Participants		
Name	**Initials**	**Organisation**
Lars Storgaard	LS	Danish Agency for Data Supply and Infrastructure
Lars-Inge Arnevik	LIA	Norwegian Mapping Authority
Loes Deventer	LD	Digital Flanders, Belgium
Ludger Rinsche	LR	]init[, Germany
Makx Dekkers	MD	SEMIC Team
Marco Combetto	MC	JRC
Marie-Bénédicte Laridant	MBL	Digital Agency, Belgium
Martynas Mockus	MM	Information Society Development Committee, Lithuania
Matthias Grönewald	MG	Federal IT Cooperation, Germany
Matthias Palmér	MP	Metasolutions AB, Sweden
Matthias Throp	MT	Norwegian Mapping Authority
Michael Reichart	MR	Austrian Federal Computing Centre
Michal Kuban	MK	DG CNECT
Michèle Spichtig	MS	Open Government Data Office, Switzerland
Nick Wysshaar	NW	Federal Roads Office, Switzerland
Nuno Freire	NF	Europeana Foundation, The Netherlands
Nuria Estables	NE	Ministry of Environment, Spain
Pascal Derycke	PD	Sciensano, Belgium
Pavlina Fragkou	PF	DIGIT
Peter Bruhn Andersen	PBA	Danish Agency for Digital Government
Pierlou Ramade	PR	Data.gouv.fr, France
Rae Knowler	RK	Liip, Switzerland
Riita Alkula	RA	Digital and Population Data Services Agency, Finland

Participants		
Name	**Initials**	**Organisation**
Simon Dutkowski	SD	FOKUS, Fraunhofer, Germany
Simon Stueur	SS	Publications Office
Terje Sylvarnes	TS	The Norwegian Digitalisation Agency
Thomas Tursics	TT	Federal IT Cooperation, Germany
Vincent Bombaerts	VB	Wallonie Public Service, Belgium
William Verbeeck	WB	SEMIC Team

Points discussed and decisions taken

Topic discussed	Outcome
DCAT-AP 3.0 Issues	
Issue [#319](#) & [#305](#): Publisher property Data Service	Approved.
Issue [#283](#): Usage of vCard	Approved with conditions. - Additional usage notes for the four subclasses are to be added, i.e. Group, Individual, Location, and Organisation.
Issue [#299](#): Definition of property *type* of Agent	Approved.
Issue [#320](#): Definition of property *type* of Agent	Not approved. - Further discussion to be held on [GitHub](#).
Issue [#295](#): Property *change type* of Catalogue Record	This issue will be resolved based on the community input that will be provided via the GitHub [issue](#).
Issue [#284](#): Property *application profile* (*dct:conformsTo*) of Catalogue Record	Approved.
Issue [#301](#): Property *access rights* of Dataset	Approved with conditions. - The term 'publicly accessible' is to be used instead of 'public/open accessible'.
Issue [#302](#): Property *access rights* of Data Service	Not approved. - Additional clarification will be posted on [GitHub](#).
Dataset Series Issues	
Issue [#275](#): A Dataset Series can belong to another Dataset Series	Approved.
Issue [#278](#): Properties of *Dataset member* of a *Dataset Series*	Approved.
Issue [#300](#):	Approved.

Inverse properties in W3C DCAT	
Issue [#296](#): Property *series member* or *in series*	Not approved. - Approved instead: remove both properties, *prev* and *next*.
Issue [#300](#): Property *prev* or *next*	Not approved. - Further discussion to be held on [GitHub](#).
Catalogues Issues	
Issue [#292](#): Catalogue organisation	Not approved. - Further discussion to be held on [GitHub](#).
Issue [#272](#): More properties for Data Service	Results of the vote will be posed on [GitHub](#).

Full Meeting Minutes

Welcome & Introduction [Slides 1 - 9](#) **Speaker:** Pavlina Fragkou	**Introduction** PF welcomes everyone and gives a general introduction on SEMIC and the SEMIC assets. This is followed by a short timeline on the activities conducted in the context of DCAT-AP HVD and DCAT-AP 3. Today all issues are related to DCAT-AP version 3. BVN explains the 3 levels of topics that are being addressed in this webinar: - Detailed topics, - DCAT-AP 3.0 alignment topics, - Cross concern topics. **Agenda**: - Welcome - Recap of last webinar - Detailed issues - Dataset Series and DCAT-AP 3 Alignment issues - Profiling & extending - Wrapup and next steps
Detailed Issues [Slides 10 - 19](#) **Speaker:** Bert van Nuffelen **Moderator:** Makx Dekkers	1) **Issue [#319](#) & [#305](#):** This issue is related to whether a Data Service should have a *publisher* property. **SEMIC Proposition:** SEMIC proposes to make *publisher* a recommended property for Data Service. **Resolvement:** The Working Group agrees with this resolution.

2) **Issue [#283](#):**
This issue is related to the usage of vCard and whether SEMIC should recommend and/or impose the usage of only the concrete subclasses of the class vCard:Kind. Or, whether instances of undetermined instances of vCard:Kind are allowed.

SEMIC Proposition:
SEMIC proposes to not make any explicit recommendation.

Resolvement:
The Working Group agrees with this, under the condition that additional usage notes for the four subclasses are added, i.e. Group, Individual, Location, and Organisation.

Discussion:
MP mentions it would be interesting to add a recommendation on the usage. However, BVN challenges this as the SEMIC proposition is to leave it open to the implementer. MP then suggests to mention the four subclasses that are available within vCard:Kind.

3) **Issue [#299](#):**
This issue is related to the definition of property *type* of Agent and its reference to the publisher of a Dataset as was defined in DCAT 1.

SEMIC Proposition:
SEMIC proposes to change the definition of Agent:type to: "the nature of the agent."

Resolvement:
The Working Group agrees with this resolution, but additional clarification will be posted on GitHub.

Discussion:
Some participants post a 0 in the chat, meaning that they are indifferent to this proposition. GN then proposes to change the naming to responsible party instead of agent. However, CN does not agree with this as responsible party is already a role, nature of agent is therefore more generic.

4) **Issue [#320](#):**
This issue is related to the definition of property *type* of Agent and its reference to publisher from adms-skos, which is not actively maintained.

SEMIC Proposition:

SEMIC proposes to change the recommended controlled vocabulary to legal entity type based on GLEIF maintained by the Publications Office.

Resolvement:
The Working group does not agree with this proposition. Instead the possibility to post comments & input through [GitHub](#) is requested.

Discussion:
MP would like to know more about GLEIF as his knowledge on the topic is limited. BVN replies that the code list is mainly focused on the business domain. It is, however, actively maintained and will be republished by OP. There is also a discussion ongoing whether it would be interesting to form a working group on GLEIF. Currently, academia and another domain are missing, BVN does not remember the exact name of the second one. MP wonders about the added value of the code list as the impact will be very large. The impact is substantial, BVN agrees, therefore this is a separate issue. It is about the nature of the agent and the real legal status of the agent.

MP raises the question whether there would be any technical problems to make this change to the audience. The European Data Portal will not make a conversion, SD replies. In general there are no issues with which vocabulary is used. However, when they are doing validation using SHACL shapes, they are not able to determine which DCAT-AP version is used. Therefore they always use the latest SHACL shapes, and older versions of DCAT-AP that use other vocabularies will not be compliant.

5) **Issue [#295](#):**
This issue addresses the fact that the label of the property *change type* does not match the definition and the mapping of the uri ads:status.

SEMIC Proposition:
SEMIC requests suggestions from the Working Group. In the issue on GitHub it was proposed to change the label to *status*.

Resolvement:
This issue will be resolved based on the community input that will be provided on the GitHub [issue](#).

6) **Issue [#284](#):**
This issue addresses the fact that the property *application profile* (*dct:conformsTo*) of Catalogue Record has a maximum cardinality of 1.

SEMIC Proposition:
SEMIC proposes to lift the cardinality.

Resolvement:
The Working Group agrees with this proposition.

7) **Issue [#301](#):**
This issue is regarding the definition of property *access rights* in DCAT-Ap using the term 'Open Data', where the codelist uses the term 'PUBLIC' defined by 'Publicly accessible for everyone'.

SEMIC Proposition:
SEMIC proposes to limit the definition to the 'level of accessibility', avoiding the term Open Data. The definition would then become: 'Information that indicates whether the Dataset is public/open accessible, has access restrictions or is not public'.

Resolvement:
The Working Group agrees with this proposition if the term 'publicly accessible' is used.

Discussion:
AA thinks that 'level of accessibility' is even worse than open data, he argues we are mixing accessibility and access rights. However, MD challenges this and argues that we are not mixing the two terms in this case. Open Data is too sensitive of a term and it should be avoided. Many other participants agree that the term 'open' should be avoided in this case as it creates too much confusion and ambiguity.

GN adds that the codelist of PO is limited to the concept of 'dataset' and not of 'services' today. In addition, the values are not distinctive/unambiguous. In addition, code list value *accessible* is not yet available in linked data format.

BVN mentions that expectations vary from one country to another and that there are different interpretations. That is the reason for adapting the definition.

CN thinks that using "publicly/openly accessible" instead of "public/open accessible" is more correct. To which DG adds that it is in fact not proper English. Correct language is necessary to reduce ambiguity.

LR adds that currently only 3 values for dct:accessRights are allowed and "Open Data" is not one of them. Therefore , the description cannot be correct. LR suggests adding "Use one of

	the following values (:public, :restricted, :non-public)." in the usage note.
	8) Issue [#302](#): This issue addresses that the codelist associated with *access rights* overloads the public access for datasets and data services. **SEMIC Proposition:** SEMIC proposes to separate the two usage scopes with distinct concepts that indicate better what is public & non-public for a data service. In addition SEMIC proposes to use two separate controlled vocabularies. **Resolvement:** The working group does not agree with this proposition. Additional clarification will be posted on [GitHub](#) to indicate what the differences are. **Discussion:** MG does not see a strong conflict and disagrees with the creation of a second codelist. MP thinks this is splitting hairs and wonders how many values there would be for a data service. MD replies that there will be six. AM argues that a general notice with FIFO notification would be appropriate. MP, however, thinks a clear list would be better to make a decision JK mentions that for the values for a data service we could clearly indicate the technical requirements in the usage note - "API key required", "Rate limited", "Registration required", etc. MD tries to close the matter as the proposal is not mature enough. There is a need for further analysis such that people can see what the implications are. MG thinks the term access rights is unfitting. With an API key restriction one might have public rights but not any means of access, so accessibility might be limited. MP shares the INSPIRE directive for reference. He notes that the articles ad hoc do not make a distinction between dataset and data service, all of them are formulated with sentences containing "spatial data sets and services". He does not say, however, that that kind of granularity is needed. The three values that are currently present are sufficient. The distinction between dataset and dataservice, on the other hand, is redundant, he argues.
Dataset Series	**9) Datasets belonging to multiple dataset series**

and DCAT-AP 3 Alignment issues **Speaker:** Bert van Nuffelen **Moderator:** Makx Dekkers	This issue addresses the remarks from the community that it is not possible to distinguish the membership in two ordered Dataset Series. **SEMIC Proposition:** SEMIC proposes to add a conclusion in the usage note: Datasets can only be part of a single Dataset Series, when Datasets belong to multiple Dataset series then the behaviour of the Catalogue is underdetermined. In addition, W3C has to be informed about the case. **Resolvement:** The Working Group agrees with this proposition. 10) **Issue [#275](#)**: This issue addresses whether a Dataset Series can belong to another Dataset Series. **SEMIC Proposition:** SEMIC proposed that it is allowed, as this is not explicitly excluded. However, the impact of implementation should be investigated to determine whether separate guidelines are required. **Resolvement:** The Working Group agrees with this proposition. **Discussion:** MP mentions that we should remind ourselves what the purpose of DCAT-AP is. MD adds that if people really have a problem with the existing situation, it will be investigated. For the time being complex cases are not being considered. 11) **Issue [#278](#):** This issue addresses the demand for additional properties regarding a *Dataset member* of a *Dataset Series*, such as *Spatial Coverage* and *Temporal Coverage*. **SEMIC Proposition:** SEMIC proposes to create a new issue on GitHub where more concrete proposals with specific information per property can be suggested. **Resolvement:** The Working Group agrees with this proposition. 12) **Issue [#300](#):** This issue is regarding the clarification of the use of inverse properties.

SEMIC Proposition:
SEMIC proposes to add a short usage note section to confirm correct usage.

Resolvement:
The Working Group agrees with this proposition.

Discussion:
MP mentions that this implies that dataset series across catalogues are not allowed. He wonders what should be done in this case, do we conclude that it is out of DCAT scope, or do we address W3C with this issue. BVN replies that it might be good to contact W3C and request a change based on our use case.

13) **Issue #296:**
This issue addresses the fact that W3C imposes the use of *in series*, making the use *series member* seem redundant.

SEMIC Proposition:
Semic proposes to remove the property *series member* from the profile.

Resolvement: to be further discussed during issue
The Working Group does not agree with this proposition. Further discussion will be held under the GitHub issue.

Discussion:
CN wonders whether it is not enough to make *in series* mandatory in DCAT-AP. That means that the implementers will need to provide it.

GN mentions that in ISO a Dataset has to be linked to a Dataset Service. It is only through the GetRelationAPI that the inverse relation can be known.

14) **Issue #300:**
This issue addresses the fact that W3C imposes *prev*, which raises the question whether *next* (which is the inverse property) should be adopted.

SEMIC Proposition:
SEMIC proposes to remove property *next*.

Resolvement:
The Working Group agrees with the proposition from MP to remove both properties, *next* and *prev,* as there are no relevant use cases.

Discussion:

CN would like to leave both. MP, however, argues the need for such properties does not exist from a use case perspective. Other participants argue that having *prev* as an optional property is sufficient.

15) **Issue [#292](#):**
This issue addresses questions on how to express that a harvested, aggregated catalogue is a combination of multiple harvested catalogues.

SEMIC Proposition:
SEMIC proposes to replace dct:hasPart with dcat:catalog.

Resolvement:
The Working Group does not agree with the proposition. To close this issue further discussion will be held on [GitHub](#).

Discussion:
MP mentions that dct:hasPart is for subsets and that is the way he and his team interpret it. HL argues that there is confusion on when to use which property. There are no use cases for dcat:catalog, but there are many for dct:hasPart. Both HL and several Swedish participants have a need for dct:hasPart.

16) **Issue [#272](#):**
Members of the Working Group are invited to vote on the inclusion of certain properties as recommended regarding Data Service. The results will be shared on the [GitHub](#) page.

Profiling & extending

[Slides 34 - 59](#)

Speaker: Bert van Nuffelen

Moderator: Makx Dekkers

Question 1: How to indicate that the properties (constraints) of one class apply to another class?
MG mentions that this is exactly the scenario we worked on in the field of highly specific research data. He does not argue they found a final solution, but he would like to point to a prior [work on the modelling of profiles](#) in a modular fashion in overlapping use-cases.

DG wonders whether the guidelines in the new section also deal with country specific profiles like the german DCAT-AP.de. PF responds that in general they apply to all profiles regardless whether they are domain or national specific ones.

Question 2: How to reuse a property with conflicting constraints?
This section is skipped due to time constraints.

Profile guidelines proposals (1)
DCAT-AP is the master profile for all profiles in the DCAT-AP ecosystem. By consequence, if DCAT-AP poses a constraint, then the other profiles should follow suit.

Profile guidelines proposals (2)

If a profile would like to use a property with different, conflicting constraints than expressed in DCAT-AP, a new property in the profile namespace has to be made. This property must be a subproperty of the DCAT vocabulary property on which the DCAT-AP property is based. This makes it semantically correct with respect to DCAT, it's easier for implementers, profile editors and interoperability between profiles.

Profile guidelines proposals (3)

These profile guidelines suggest the following:
- A subclass relationship in the diagram means 'inherits' all properties (and their constraints) from the superclass.
- A mapping on a URI means 'semantics' are shared. The usage of the property is inline with the semantics associated with the URI.
- Avoid changes to definitions, use the usage notes to the scope.

These guidelines are in line with the reuse guidelines of the SEMIC Style Guide.

MP wants to confirm that this is a diagrammatic rule and has nothing to do with RDFS or any formal application profile expression. This is confirmed by BVN.

Profile guidelines proposals (4)

These profile guidelines suggest the following:
- If a property is reused (from its superclass) and has no changes then repetition is not needed.
- If repetition is done, indicate clearly that there are no changes.

This applies 'cross-profile' too:
- If a property is reused from its superclass in the other profile and has no changes then repetition is not needed.
- If repetition from another profile is done, indicate clearly that there are no changes.

It should be noted that repetition is not mandatory.

MP argues that when they are using a data set in a national profile they still call it a DCAT dataset and there is no subclassing happening. CN agrees with this statement. BVN responds that if they were to put this dataset visually on a diagram it would be represented as a subclass of DCAT-AP. MP does not agree, it would not be a class in the diagram, therefore there is a problem with how we draw diagrams. It is in fact not a class, but a profile. BVN agrees that we have to mitigate the visual language of a class versus profiling. BVN chose to use a class diagram, he does not see a better alternative. CN adds that there is a way to express the profile language and avoid the superclass structure. They have a very concrete proposal on how to use this by using another arrow. In addition, the profile visuals are different from the class visuals. CN will be able to share an article on this with interested stakeholders some time soon.

	BVN highlights the repetition of properties in a profile. With this proposal it should be highlighted that repetition is not needed and if you do it, you should explicitly flag it. MP agrees with the intention, but not with the language that is used to express this. It causes problems for people conflicting the idea classes and profiles. **Profile guidelines proposals (5)** This proposal touches on the scoping of for example the subclass dataset in series, which is a local scoped dataset in DCAT-AP with a particular usage. The question is whether it is allowed. This way of doing it allows clear scopes to be expressed and it is easier to read. Discussion will be created to have further discussion on github for these proposals.
Closing & Next steps Slides 68 - 71 **Speaker:** Bert van Nuffelen **Moderator:** Makx Dekkers	BVN concludes the webinar and explains the next steps. A question is raised by AM on the obligation to maintain a certain "quality of data" by maintaining previous errors in next stages. PD adds that both the quality of data and the utility of data need to be maintained. MD closes this discussion by stating that a new GitHub issue will be opened where this discussion can continue. Closing words by Pavlina, asking for participants' contribution on GitHub such that SEMIC can improve itself.