Meeting Minutes: Third Working Group webinar on the revision on GeoDCAT-AP (SEMIC - A04.02)

Project:	SEMIC	Date and Time:	23/04/2024 10:00 - 12:00
Meeting Type:	Webinar	Location:	Virtual
Coordinators:	Bert Van Nuffelen Jakub Klímek	Issue Date:	07/05/2024

Agenda of the webinar		
10:00 - 10:10	Introduction	[Slides 1 - 3](#)
10:10 - 10:15	The GeoDCAT-AP ecosystem	[Slides 4 - 10](#)
10:15 - 11:00	GeoDCAT-AP issues	[Slides 11 - 36](#)
11:10 - 11:25	Guest speakers: National experience	[Slides 37 - 84](#)
11:55 - 12:00	Wrap-up & next steps	[Slides 85 - 90](#)

Meeting Slides
[LINK](#)

Participants		
Name	**Initials**	**Organisation**
Age Sild	AS	Estonian Land Board
Anastasia Sofou	AS	SEMIC Team
Anja Loddenkemper	AL	GDI-NI, Germany
Antoine Jacques	AJ	GIM, Belgium
Arthur Schiltz	AS	SEMIC Team

Participants		
Name	**Initials**	**Organisation**
Athina Rapsomaniki	AR	Ministry of Home Affairs, Greece
Aurelija Sturaite	AS	Construction Sector Development Agency, Lithuania
Bert Van Nuffelen	BVN	SEMIC Team
Bjorn Hagstrom	BH	Hagström Consulting AB, Sweden
Byron Cochrane	BC	OGC
Christoph Mikula	CM	Citizens' Energy Communities, Austria
Claudia Attard	CA	Environment and Resources Authority, Matla
Elena Grigoriou	EG	Ministry for Environment and Energy, Greece
Fabio Vinci	FV	Epsilon Italia
Floor Van Damme	FVD	National Geographic Institute, Belgium
Francois Prunayre	FP	Cerfrance Haute-Loire, France
Geraldine Nolf	GN	Digital Flanders, Belgium
Gery Nicolay	GN	Spacebel, Belgium
Ine de Visser	IdV	Geonovum, Netherlands
Inês Quintanilha	IQ	Publications Office
Jakub Klímek	JK	SEMIC Team
Jitka Faugnerová	JF	Czech Environmental Information Agency
Jitse De Cock	JDC	SEMIC Team
Joachim Nielandt	JN	Digital Flanders, Belgium
Joeri Robbrecht	JR	DG ENV
Jordi Escriu	JE	JRC
Josema Alonso	JA	SEMIC Team
Juliana Karciauskiene	JK	The National Center for Remote Sensing and Geoinformatics, Lithuania
Kate Lyndegaard	KL	WeTransform, Germany

Participants		
Name	**Initials**	**Organisation**
Kůsová Veronika	KV	State Administration of Land Surveying and Cadastre, Czech Republic
Lars-Inge Arnevik	LIA	Norwegian Mapping Authority
Lena Hallin-Pihlatie	LHP	Finnish Land Survey
Lenka Rejentova	LR	Czech Environmental Information Agency
Marceau Louis	ML	Public Service of Wallonia, Belgium
Markus Jobst	MJ	Citizens' Energy Communities, Austria
Mart Randmäe	MR	Estonian Land Board
Mathieu Chaussier	MC	GIM, Belgium
Matthias Palmér	MP	Metasolutions AB, Sweden
Mihai Paunescu	MP	Publications Office
Ouns Kissiyar	OK	Digital Flanders, Belgium
Paloma Abad	PA	National Geographic Information Center, Spain
Pavla Kramolisova	PK	Czech Geological Survey
Pawel Soczewski	PS	GISPartner, Poland
Silvie Semerádová	SS	T. G. Masaryk Water Research Institute, Czech Republic
Štěpán Kafka	SK	Charles University, Czechia
Stijn Tallir	ST	Digital Flanders, Belgium
Uwe Voges	UV	Conterra, Germany
Vendula Dastychová	VD	Czech Environmental Information Agency
Veronika Kusova	VK	State Administration of Land Surveying and Cadastre, Czech Republic
Jesper Zedlitz	JZ	State Chancellery of Schleswig-Holstein, Germany

Points discussed and decisions taken

Topic discussed	Outcome
GeoDCAT-AP	
Issue [#100](#)	A resolution will be proposed during the webinar of May 14.
Issue [#107](#)	Approved.
Issue [#96](#)	Approved.

Full Meeting Minutes

Welcome & Introduction [Slides 1 - 3](#) **Speaker:** Bert Van Nuffelen	BVN welcomes the participants and introduces the topics of the webinar: ● Introduction ● Geospatial issues ● Codelists (postponed due to time constraints) ● Quest speakers: National experience ● Next steps
The GeoDCAT-AP ecosystem [Slides 4 - 10](#) **Speaker:** Jakub Klímek, Bert Van Nuffelen	**GeoDCAT-AP ecosystem** GeoDCAT-AP is primarily a DCAT-AP profile for geospatial datasets. However, the major other inputs are the ISO and INSPIRE standards. Recently, the OGC Standards Working Group (SWG) started working on the GeoDCAT standard. **GeoDCAT** The progress on GeoDCAT is currently limited, but ongoing. To continue the progress, community members are invited to participate. In addition, a hands-on workshop will be organised in autumn but will be further concretised later on. Collaboration will continue between OGC, SEMIC and JRC. **Next webinar** The next webinar will take place on the 14th of May which will cover the remaining issues, High-Value Datasets (HVD) and the existing tools, particularly the XSLT. **Issue overview** An overview of the issues is given. A total of 57 issues are open on GitHub. Attention is raised for the issues with label 'status:resolution-proposed'. If no feedback is received on these issues, they will be closed with the proposed resolution regarded as approved. Issues that receive a green label will not be discussed during the webinar. Such issues are to be discussed on GitHub. The principle, that the issue will be closed unless negative feedback is received, is applied here too.

GeoDCAT-AP issues
 	
Slides 11 - 36
 	
Speaker:
	Issue #100
	Spatial resolution is an optional property on Dataset, Distribution and Data Service. The issue is regarding the relation of the spatial resolution properties among these classes.
	
	The questions raised by the SEMIC team regarding this issue are:
	1. Can a Distribution have a different spatial resolution than the Dataset?
	    1.1. If not, why do we need the property both on a Dataset and a Distribution?
	    1.2. If yes,
	        1.2.1. Should all Distributions of a Dataset have the same spatial resolution as the Dataset?
	        1.2.2. Should the Dataset be tagged by all the different spatial resolutions found in its Distributions?
	2. Do all datasets served by a Data Service have the same spatial resolution as the Data Service?
	
	Resolution
	It is agreed that based on the feedback received during the webinar, a proposition will be drafted to be discussed during the webinar of 14 May.
	
	Discussion
	GN mentions that in ISO the Distribution is included in the Dataset; when mapping to GeoDCAT it will be the same.
	
	JZ adds that he has seen Datasets with Distributions having different spatial resolutions.In his opinion, 1.2.2 looks like a suitable solution. PS and JE also consider 1.2.2 as the best option.
	
	MP proposes as an alternative solution to simply mark the most specific spatial resolution on the dataset level. Currently distributions allow different languages and formats, however, two distributions are considered equivalent, and not complementary. It is unclear to MP whether different distributions are related to one dataset, or different datasets. He argues that two spatial resolutions on different distributions cannot really be related to the same dataset.
	
	SK and AL agree with the arguments expressed by MP.
	
	AL argues that there is only one spatial resolution related to a single dataset. The data is published, and the resolution cannot change via a publication.
	
	PA expresses that the same product may have different resolutions. For example, an address database, for which the portals are at a scale of 1:5,000 and the roads in interurban areas at 1:10,000.
	
	LHP is of the opinion that many resolutions can be specified when

describing a dataset.

However, SK argues that it is not the same dataset if distributions have different resolutions. He adds that in ISO 19115, resolution only exists for a dataset. AL agrees with this.

PS describes the case where one source dataset is distributed with different resolutions. He would regard this as multiple distributions related to the same dataset.

UW raises the idea that the resolution of the dataset could be defined as the default resolution for the distribution. However, a distribution could overwrite it.

Issue [#107](#)
With multiple values for spatial coverage being allowed e.g. for a dataset, it may not be clear enough, what the interpretation of those multiple values should be based on the current usage note.

The current usage note is: "This property refers to a geographic region that is covered by the Dataset."

SEMIC Proposition
Explicitly say in a usage note that: "when multiple values are used for spatial coverage, this may be interpreted as a spatial union, or as alternative representations of spatial coverage that might fit, with no explicit spatial relation" to give guidance to implementers as to how to treat the multiple values of spatial coverage. This usage note indicates which kind of interpretations are possible, without the restriction of either one.

Resolution
The proposition is approved.

Discussion
MP believes there are only two cases in DCAT-AP. Namely, URI and boundingBox explicitly. He wonders whether we need to support the two in between.

JZ interprets the different values as being alternatives for the same thing.

GN notes that in ISO, the spatial representation is defined by a bounding box. SK adds that multiple bounding boxes and bounding polygons are also included in ISO. He agrees that both representations should be kept in DCAT-AP.

MP replies that if two bounding boxes are provided, they should be treated as unions.

AL mentions that in Germany the union of extents / geographic areas is specified and that a boundingBox-polygon is a must-have. Additionally, there are persistent spatial identifiers or names of cities for statistical reasons. IdV adds that in the Netherlands this is also the case.

According to JE, the solution should definitely support disjoint bounding boxes which hopefully is available in DCAT-AP based encoding.

Issue [#96](#)
The meaning of having multiple spatial / geographic coverages specified on one Data Service may not be clear enough.

Currently the usage note is as follows: "This property refers to a geographic region that is covered by the Data Service.

SEMIC Proposition
Explicitly say in a usage note that: "If the Data Service provides access to multiple Datasets, then this will be the union of the spatial coverages of these Datasets."

Resolution
The proposition is approved.

Discussion
No additional discussion was held on this issue.

The GeoDCAT-AP Strategy
When implementing a use case two different strategies can be followed, namely the abstraction/aggregation approach where information is condensed into one property and the precise approach where the right property is applied for the use case.

For instance, multiple code-list values can be used for the same subject, this is an aggregation approach. However, for another use case sub-properties are used, in order to have more precise metadata. This would allow using a unique label or applying specific restrictions, while still relating to a higher layer of abstraction.

On the aggregation level the most prominent advantage is that it is just one property to implement. However, as it caters to a more abstract usage the label and definitions become more abstract. Additionally, cardinality constraints and usage notes become more difficult to interpret.

In the precise approach, the major disadvantage is increased implementation efforts. However, labels and definitions are more precise. Additionally, cardinality constraints are unambiguous, usage notes are simpler and automation of constraint generation in artefacts is simpler as well.

Let us consider bringing these implementations to a portal. If you would like a representation where each dataset from a different use case is listed, the precise approach is a natural match. For an aggregated view, the aggregation approach is the natural match. Performing the mapping from a precise view to an aggregation view is fairly simple. In the other direction, however, it becomes hard to distinguish the different code-lists, the labels, etc.

Within the context of GeODCAT-AP, these two approaches translate to two implementation strategies.

Strategy 1:
If GeoDCAT-AP is derived from INSPIRE metadata, then the aggregation approach is sufficient. In this case going from GeoDCAT-AP, that is the aggregation, back to INSPIRE or ISO becomes very difficult.

Strategy 2:
If GeoDCAT-AP is to be used natively for the INSPIRE metadata, then it must express the INSPIRE metadata as precisely as possible, hence it would be the precise approach. In this case two-way conversion would be much easier.

Depending on which strategy is chosen the shift to the other strategy may have different impacts. Shifting from the abstract approach to precise would require large efforts in the form of a major rewrite. However, the precise approach to the abstract approach can be largely automated.

Discussion
GN mentions that if you take Strategy 2 you must take into account the whole ISO standard (and not only INSPIRE) to foresee niche richness in GeoDCAT. AL, PS and MJ agree with this statement. GN adds that in Flanders Strategy 1 is preferred as the geospatial description is done in ISO/INSPIRE.

MP does not agree with the statement of "impossible" in Strategy 1. He argues that not the UI design but the semantics of the subjects should be the driver in choosing a strategy.

IdV argues that it would be good to take a step back and consider the purpose of GeoDCAT-AP. BVN replies that it is not a matter of making GeoDCAT-AP a replacement for ISO/INSPIRE but to make the specification adaptable if GeoDCAT-AP would become a replacement in the future.

JZ believes Strategy 2 opens more possibilities because the two-way conversion becomes possible. PA, JE and PS agree with this.

MP would like to mention that it is possible to check the URI patterns

which makes conversion in Strategy 1 possible. BVN has a feeling that it is a bit more complex and that implementers would have to do this themselves as SEMIC currently does not offer tooling for this. MP argues that SHACL can be used as a starting point for the conversion.

IdV argues that Strategy 2 is more difficult for ISO metadata than for INSPIRE metadata. The reason being that it covers different domains, so we have to make properties for all different domains. The variety of different domains with different code-lists will require many different properties to cover all the different domains and code-lists. For a national portal it will be difficult to make this possible.

MP replies that in Sweden they simply allow uploading terminologies to the data portal. Therefore, the content the data portal has access to gets expanded over time. If a couple of these properties with an open range of concepts exists, one can simply upload the values that data providers use.

GN believes it is not about code-lists or sub properties but about whether GeoDCAT-AP is a replacement for ISO/INSPIRE. She argues that ISO is semantically richer than GeoDCAT-AP and then either information is lost in data exchange between both profiles or GeoDCAT-AP will need significantly more complex usage notes.

BVN admits that he does not have an answer. The idea of choosing a strategy is more in terms of a design philosophy. In the current aggregation strategy for example, all INSPIRE topics would be mapped to dct:subject with a usage note containing information on all the different topics, while the visual representation is more akin to the precise mechanism. This is confusing and therefore a decision on the strategy is important.

GN replies that it is not about the code-lists. An aggregation approach or a precise approach can still be used in both strategies, i.e. whether GeoDCAT-AP replaces ISO/INSPIRE or not. She argues that these two approaches and the strategies have to be decoupled.

JR from DG EVN shares his thoughts on the discussion of the strategy. He mentions that by 9 February 2025 MS must report on HVD implementation and provide DCAT-AP metadata. The metadata describing the datasets within the scope of the INSPIRE data themes shall contain at least the metadata elements set out in Commission Regulation (EC) No. 1205/2008. The main concern is the full reuse of INSPIRE implementation for Open Data / HVD obligations to limit implementation and reporting burden. Hence, a full mapping of INSPIRE metadata elements is needed. MS should have the freedom to decide how they set up the governance/harvesting process for INSPIRE and Open Data catalogues (e.g. direct harvesting of all catalogues by data.europa.eu, national open data portals harvesting national geoportals etc.). GeoDCAT-AP should satisfy the legal

	obligations (all INSPIRE metadata elements for HVD data in scope of INSPIRE) to be useful for regulatory reporting purposes and at the same time support maximal flexibility to MS and not limit/prescribe implementation. SEMIC mentions that in discussion with DG ENV and JRC it was concluded that strategic approach 2 (precise) would be a more appropriate approach in case GeoDCAT-AP may become a replacement for ISO/INSPIRE in the future. This affects the issues that were postponed from the last webinar.
Guest speakers: National experience Slides 37 - 84 **Speaker:**	**The National Land Survey of Finland** Finland has a national data catalogue that they provide to the INSPIRE metadata providers. They can describe their metadata there. The templates that they feature are available in three languages, Finnish, Swedish and English and are all based on the ISO profile. Around 100 organisations are currently using the service. The National Metadata Catalogue is harvested by the INSPIRE Geoportal. It is also used in the Finnish National Geoportal which has a metadata search where this metadata is integrated. The National Land Survey (NLS) has close collaboration with the open data portal of Finland. It is based on CKAN and uses an extension of DCAT-AP. The NLS has built a system with a virtual CSW interface based on an open data keyword. They are then further filtered by licence and those with the correct licence are mapped to DCAT-AP and provided to data.europa.eu. The mapping started with mapping ISO and INSPIRE to DCAT-AP. The mapping is not yet complete, but is constantly being updated. The user can always go from the open data portal to the original geoportal to find richer metadata. In the future, the NLS would like to include a solution for the data providers to publish so-called high-value datasets. Additionally, the GeoNetwork software for the geoportal and the CKAN software of the open data portal will be updated to the latest versions. **Digital Flanders** The metadata nodes and portals are presented in different levels. Namely, regional/federal, national, and European. In the first level the metadata nodes for the Flemish, Walloon, Brussels, and Federal region are found. All these, except for the Walloon region, already harvest ISO/INSPIRE standards through CSW harvesting. All regions have one or more open data portals.

The regional portals, and other portals that are not in the regional layer, are harvested in the national portal. Through the Belgian national node, data is provided to data.europa.eu through the Belgian EU node. These datasets include datasets that are filtered on open data keywords.

The geospatial datasets that are filtered on INSPIRE keywords are harvested directly from the regional layer to the European layer, which is the INSPIRE geoportal.

In the regional and national portal GeoNetwork and OpenDataSoft are the most used technologies.

On the Flemish level the selected approach was that the geospatial dataset and services follow the ISO/INSPIRE based profiles and are exported as such in the Flemish geoportal. However, all metadata is handled in one metadata management system, including open data which is DCAT-AP based. Therefore, the geospatial data from the ISO/INSPIRE metadata standards can be exported in GeoDCAT-AP and provided to the Flemish Open Data portal.

The underlying technologies for this integrated system are GeoNetwork 4.4 with a DCAT-AP 2.0 plug-in built on top of the system to support the Open Data side.

The geospatial approach was done using ISO and INSPIRE standards. These were adapted to the SDI-Flanders Best Practices on Metadata 2.0, which then were used to describe geospatial datasets and services. On the generic side of things DCAT and DCAT-AP are used as input for DCAT-AP VL, which is used to describe everything Open Data. To connect both worlds, geospatial and open, a mapping from the best practices to GeoDCAT-AP VL is used. In that sense every geospatial dataset is available with generic, open data descriptions as well.

For each layer, that is geospatial data and services, open data, open services, closed services and closed data, the same levels exist. The starting point is an existing metadata standard, next is the metadata management system, then the publication via a portal or catalogue, and lastly the data itself. The region of Flanders has an aggregated website where all data from each layer can be found.

On the portal you can create a new dataset and choose according to which profile you would like to describe. Next, a template can be chosen such as a dataset, a service, etc. The website also has built- in validation.

The Metadata Vlaanderen node can be used directly, or external systems can transfer data to the node. If the data is in DCAT-AP it can be harvested automatically overnight.

For each portal, the aggregated portal, open data portal or geoportal, and even data spaces, the metadata is provided through the same metadata layer to guarantee the once-only principle.

Furthermore, collaboration between the different regions in Belgium is becoming closer and closer.

The largest challenge for Flanders is in the interoperable standards used in different sectors. Herein the use of persistent URIs at every level can help avoid duplication.

Spain
The structure of data collection in Spain is highly similar to that of other SDIs. There are national nodes collecting data from different departments, and regional nodes collecting data from different local regions. In total there are 17 regional SDIs and over 43000 resources, which are aggregated in the Spanish geoportal called IDEE. However, another catalogue exists and is focused on high-value datasets.

The open data infrastructure in Spain is similar to that of the geospatial data. In theory, the data should be exchanged from the highest level, that is the geoportal, to the open data portal. However, in reality the situation is different,and data is harvested to the open data portal from metadata nodes on lower levels than the national geoportal. This is causing metadata files to be duplicated, due to the fact the file identifier is dependent on the catalogue where it is published.

This phenomenon also exists on data.europa.eu, but also within the Spanish infrastructure. In Spain the metadata is based on INSPIRE and uses an identifier. Therefore, it is essential that this is reflected in DCAT-AP.

In the end the metadata file should only appear once in the highest level portal, for example data.europa.eu.

Discussion
BVN wonders whether the mapping that was shown was from the INSPIRE themes to the DataPortal themes. LHP replies that it was from ISO categories to DataPortal themes.

JZ notes that the German GDI-DE also has defined a mapping from ISO topics to DCAT-AP themes. GN mentions that Flanders has published them as well.

BVN wonders whether there are other parties supporting the creation of the DCAT-AP plugin for GeoNetwork in the case of Flanders. GN replies they need to have another 'sprint' (or 2 / 3 sprints) with the core community and other interested parties to push their DCAT plug-in version 2 back into the core of GeoNetwork. In that sense, the 'Flemish variants' are excluded and the generic ones are included in the core,

	ready to use for anyone. With regards to identifiers, MP mentions that the file identifiers are identifiers for the metadata record and will not be preserved when copied. The resource identifiers on the other hand such as gmd:Identifier should be preserved and should be copied into the dct:Identifier in GeoDCAT-AP.
Wrap-up & next steps Slides 85 - 90 **Speaker:** Bert Van Nuffelen, Pavlina Fragkou	The participants are thanked for the contributions and encouraged to provide feedback on the GeoDCAT-AP GitHub repository.