

Agenda

01 Welcome 02 Recap of last webinar 03 **Detailed issues** 04 Dataset Series and DCAT-AP 3 alignment issues 05 Profiling & extending

Wrap up and next steps

06

Workshop practicalities

Audio

Click on 'connect audio' but please mute your microphones



Chat

You can also share your questions for the Q&A session via the chat



Recording

The workshop will be recorded



Context of the SEMIC assets

Introduction to SEMIC



The objectives of the SEMIC action is to promote Semantic Interoperability amongst the EU Member States by:



Promoting, share and reuse of semantic assets, experience and tools and facilitating agreements in key areas.



Identifying opportunities for alignment on semantic definitions, metadata and reference data sources with special focus on identification and definitions of Core Concepts / Vocabularies.



Raising awareness on the importance of data and metadata management.



Current SEMIC assets



A person's name(s), date and place of birth/death, identifier, addresses, citizenship, etc.

Vocabularies



The legal name, address, identifier, company type, and activities of a legal entity.

CORE LOCATION VOCABULARY

The different ways of describing a location, e.g. via an address, a geographic name, or a geometry, in alignment with INSPIRE.



The requirements and evidence of a procedure or formal process.



The administrative information, hierarchy, identifiers, events and classification of a public organisation.



A public event, its time, audience, location, etc.

Application Profiles

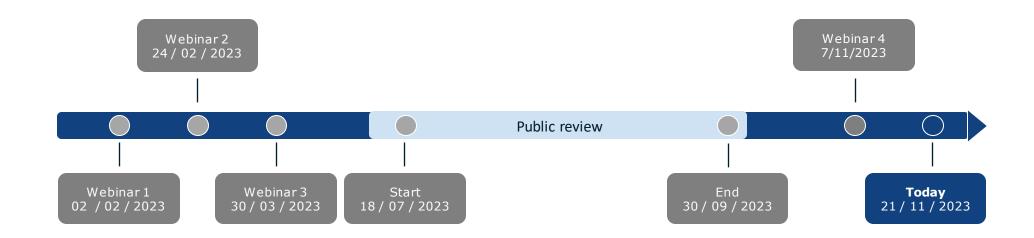


DCAT-AP FOR DATA PORTALS IN EUROPE BRegDCAT-AP FOR BASE REGISTRIES GeoD CAT-AP FOR GEOSPATIAL DATASETS

StatDCAT-AP FOR STATISTICAL DATASETS ADMS
ASSET
DESCRIPTION
METADATA
SCHEMA

Timeline DCAT-AP 3.0 and DCAT-AP HVD





Timeline webinars



Today 21/11/2023

Webinar dedicated to the status and governance of DCAT-AP

Webinar dedicated to the technical issues related to DCAT-AP

Objective of the webinar series



Webinar of 7/11/2023

Discussing the feedback received and proposed resolutions regarding:

DCAT-AP HVD

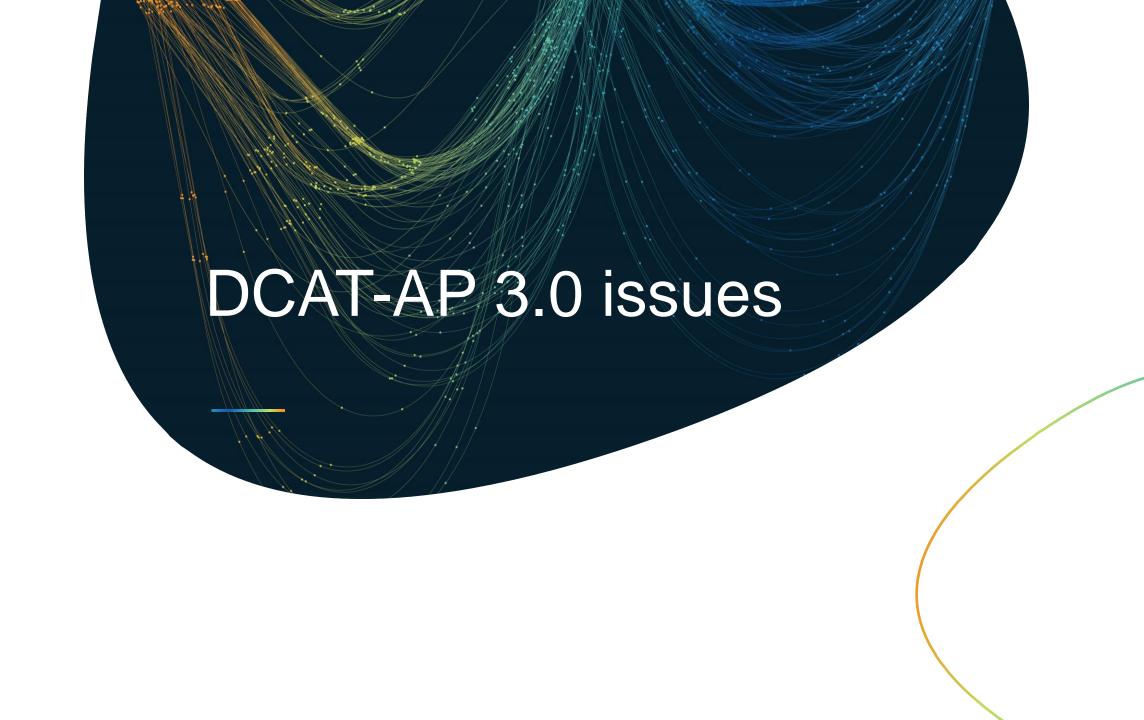
A status update and overview for

- DCAT-AP 3.0
- DCAT-AP extensions
- DCAT-AP ecosystem



Today's webinar

- Resolving the technical issues for DCAT-AP 3.0
 - Detailed topics (ContactPoint, Agent Types, Change Type, Cardinality conformsTo CatalogueRecord, ...)
 - 3.0 Alignment topics (Dataset Series, inverse properties, catalogue composition, etc.)
 - Cross-concern topics (Guidelines for building DCAT-AP profiles, profiling themes and other properties, ...)



Publisher property Data Service

<u>(#319</u> derived from <u>#305</u>)

Description

A Data Service is a catalogued resource at the same level of a Dataset. For Dataset a publisher is recommended. Should a Data Service not have the same "base quality"?

Motivation

- According to DCAT a Data Service is a resource at the same "level" of Dataset.
- DCAT-AP does not recommend any relationship with an Agent

Proposition

Make publisher a recommended property for Data Service

Usage of vCard (issue #283)

Description

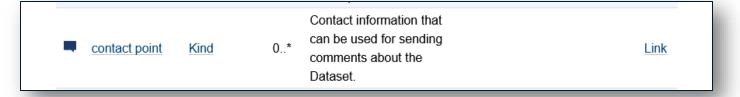
Should DCAT-AP recommend or impose the usage of only the concrete subclasses of the class vCard:Kind? Or are instances undetermined instances of vCard:Kind allowed?

Motivation

- DCAT(-AP) specifies as range for contact point of a Catalogued Resource (dcat:contactPoint) vCard:Kind.
- vCard:Kind has 4 subclasses Group, Individual, Location and Organization.
- DCAT(-AP) uses explicit relationships with Agents (publisher/creator/etc.) and also there no subclass requirement is expressed.
- A contact point is used in DCAT-AP to share contact details (tel., email, etc.). The nature behind those contact details are considered less important as it will be connected closely to the expressed relationships with the Agents.

Proposition

Do not make any explicit recommendation.



Definition of property type of Agent (issue #299) (1)

Description

In DCAT 1 the sole agent that was introduced was the publisher. Therefore the definition of the *Agent type* refers in its definition to this role.



Motivation

Removing the role information make the definition of the property circular.

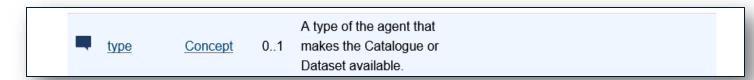
Proposition

• Change definition to: "the nature of the agent."

Definition of Agent type (#320 derived from #299) (2)

Description

In DCAT 1 the sole agent that was introduced was the publisher. Therefore the definition of the *Agent type* refers in its definition to this role.



Motivation

- The recommended controlled vocabulary also includes "publisher" role in the name (adms-skos, which is not actively maintained.)
- In the context of Core Business Vocabulary SEMIC and PO work a more precise code list legal form type based on GLEIF

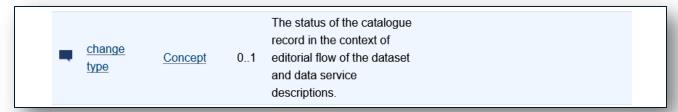
Proposition

Change the recommended controlled vocabulary to legal entity type based on GLEIF maintained by Publication Office.

Property change type of Catalogue Record (issue #295)

Description

The label of the property change type seems not to match with the definition and the mapping of the uri adms:status.



Motivation

The property indicates a change in the context of the editorial flow: the metadata of the Catalogued Resource has been created/updated/removed.

History:

- property was present since DCAT-AP 1
- DCAT 2 provided a generic definition and in DCAT-AP 2, both have been aligned.

Proposition

- Request for suggestions
- Be backwards compatible
- Proposed change in the issue is to adapt the human-readable label to *status*

Property application profile (dct:conformsTo) of Catalogue Record (issue #284)

Description

The cardinality of the conforms to for a Catalogue Record is set to max 1.

Р	roperty	Range	Card	Definition	Usage	DCAT
	application orofile	Standard	01	An Application Profile that the Dataset's metadata conforms to.		<u>Link</u>

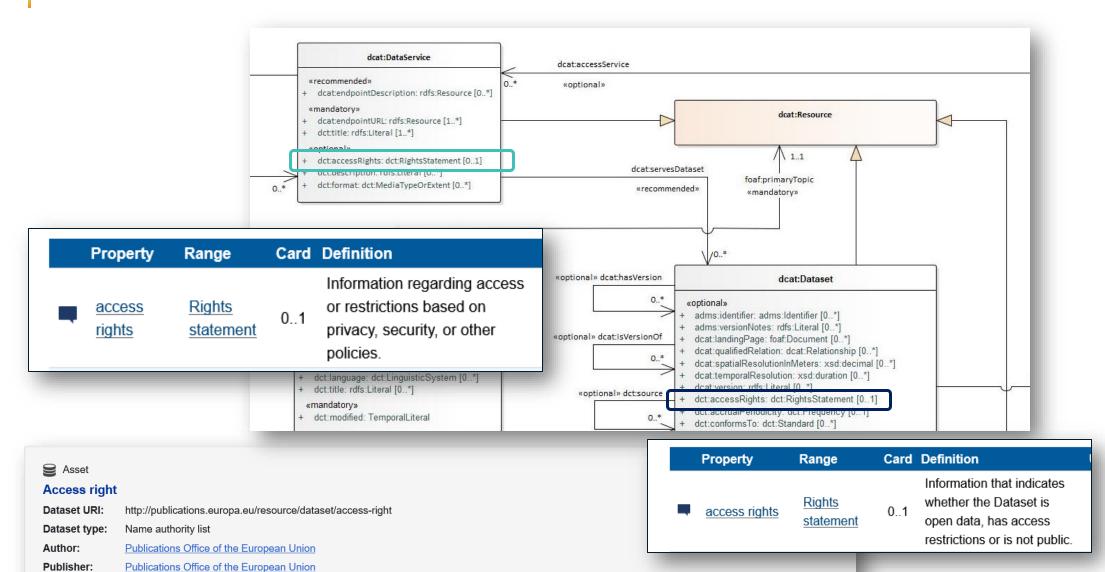
Motivation

A Catalogued Resource can conform to multiple profiles e.g. DCAT-AP, DCAT-AP HVD and MobilityDCAT-AP.

Proposition

Lift the cardinality.

Property access rights (issue #301 and #302)

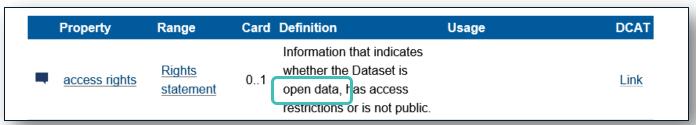


Description: The access-right authority table is a controlled vocabulary listing the access rights or restrictions to resources. It is designed for but not limited to DCAT descriptions of datasets. This authority table is maintained by the Publications Office of the European Union on the EU Vocabularies website.

Property access rights of Dataset (issue #301)

Description

The definition in DCAT-AP is using the term "Open Data". This term is comes with a lot of expectations. While the codelist uses "PUBLIC" with definition "Publicly accessible by everyone".



Motivation

- Open Data is a policy term: it comes with a set of requirements (license, access limitations, presence of distributions and data services, etc.)
- Example: In Belgium the Dataset can be regarded as Open Data even though the Data Service requires an API key. In Czechia the Data Service cannot have such restrictions in order to be categorised as Open Data.

Proposition

- Limit the definition to the "level of accessibility", avoid the term Open Data.
- Proposed new definition: "Information that indicates whether the Dataset is public/open accessible, has access
 restrictions or is not public."

Property access rights of Data Service (issue #302)

Description

The codelist associated with access rights overloads the public access for datasets and data services.

access-right:PUBLIC

Publicly accessible by everyone. Usage note: Permissible obstacles include registration and request for API keys, as long as anyone can request such registration and/or API keys.

Motivation

- Public access for a Dataset = no legal base to limit the access
- Public access for a Data Service = no limitations to use the service, besides technical that ensure the operational stability (like rate restriction, API keys, registration, etc.)

Leads to a conflict: as for one use case there are no limitations allowed, while for the other there are.

Proposition

- Separate the two usage scopes with distinct concepts that indicate better for a data service what is public/non-public
- Two separate controlled vocabularies

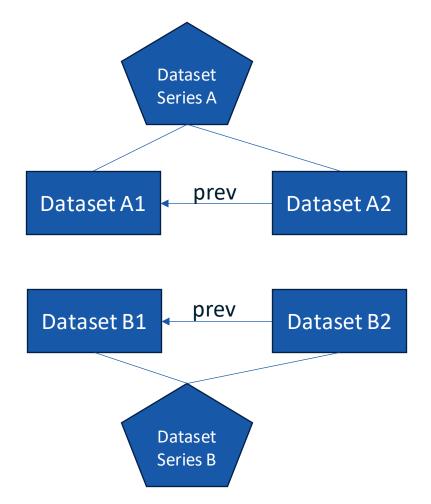


Datasets belonging to multiple Dataset Series

Description

The community made the remark that it is impossible in DCAT to distinguish the membership in 2 ordered Dataset Series.

Motivation

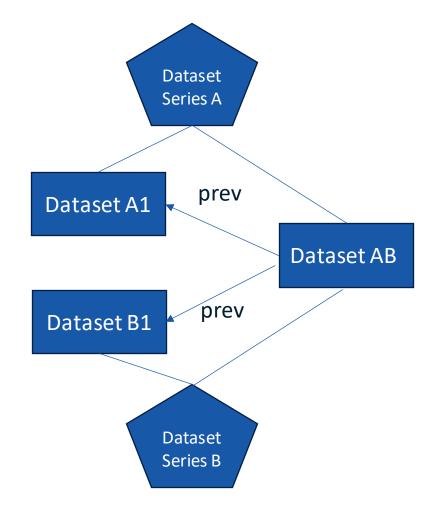


Datasets belonging to multiple Dataset Series

Description

The community made the remark that it is impossible in DCAT to distinguish the membership in 2 ordered Dataset Series.

Motivation



The ordering information *prev* is indistinguishable w.r.t. Dataset Series

Datasets belonging to multiple Dataset Series

Description

The community made the remark that it is impossible in DCAT to distinguish the membership in 2 ordered Dataset Series.

Motivation

- The ordering properties of DCAT prev (and next) are not scoped to the Dataset Series, but to the Dataset.
- If a Dataset belongs to two Dataset Series then the Dataset (may) has 2 prev Datasets, but one cannot distinguish the ordering.
- Example:
 - the 'prev' for Dataset AB in the first Dataset Series 1 is Dataset A1
 - the 'prev' for Dataset AB in the second Dataset Series 2 is Dataset B1.

Proposition

- 1. Add conclusion in a usage note:
 - 1. Datasets can only be part of a single Dataset Series.
 - 2. When Datasets belong to multiple Dataset series then the behavior of the Catalogue is underdetermined.
- 2. Inform W3C about the case.

A Dataset Series can belong to another Dataset Series (Issue #275)

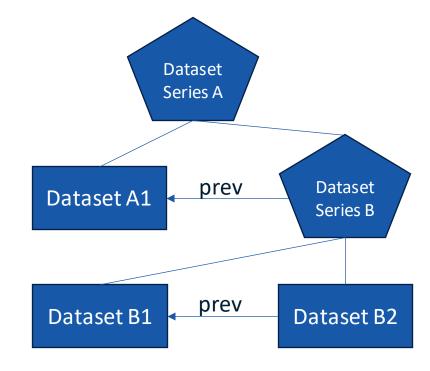
Description

Can a Dataset Series belong to another Dataset Series?

Motivation

In DCAT there is no explicit mention of this case.

DCAT-AP does not mention it either.

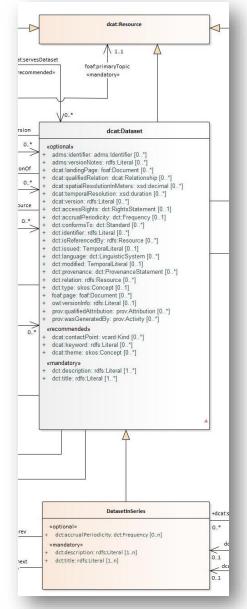


Proposition

As it is not explicitly excluded, it is allowed.

Proposal: investigate the impact when implementers implement this case. Determine whether or not separate guidelines would be required.

Properties of Dataset member of a Dataset Series (Issue #278)



	Property	Range	Card	Definition	Usage	DCA
•	description	<u>Literal</u>	1n	A free-text account of the Dataset.	This property can be repeated for parallel language versions of the description. The description may indicate the dimension values.	<u>Link</u>
•	<u>frequency</u>	<u>Frequency</u>	0n	The frequency at which the Dataset is updated.	The frequency of a dataset belonging to a dataset series is not equal to the frequency of the dataset series. E.g. A dataset budget 2020 with frequency yearly means that this dataset has an update in 2021, 2022, 2023, etc.	Link
-	<u>in series</u>	<u>Dataset</u> <u>Series</u>	0*	A dataset series of which the dataset is part.		Link
-	<u>next</u>	Dataset member of a Dataset Series	0*	The following resource (after the current one) in an ordered collection or series of resources.		
-	<u>previous</u>	Dataset member of a Dataset Series	0*	The previous resource (before the current one) in an ordered collection or series of resources.	Unless the dataset is the last in the chain a dataset in a collection must have a previous one.	Link
-	<u>title</u>	<u>Literal</u>	1n	A name given to the Dataset.	This property can be repeated for parallel language versions of the name. The title may indicate the dimension values.	Link

Legend:

Mandatory for a DCAT-AP Dataset Other usage note than a DCAT-AP Dataset

Only needed for this usage

Properties of Dataset member of a Dataset Series (Issue #278)

Description

Add other properties such as Spatial Coverage and Temporal Coverage

Motivation

Beyond they are useful, what is additional or more specific information that the requesters want to share about them?

Proposition

Request for a more concrete proposal with additional, more specific information per property. Suggestion to create a new issue.

Legend:

Mandatory for a DCAT-AP Dataset
Other usage note than a DCAT-AP
Dataset
Only needed for this usage



Inverse properties in W3C DCAT (Issue #300)

Description

- W3C introduces properties and their inverses.
- W3C imposes the use of a specific direction, the other is optional: "if one uses the optional one, then also the other MUST be shared."
- W3C does not "allow full profile freedom".

Motivation

- Ensure that DCAT exchanges have the same 'directionality' in the knowledge graphs.
- NO impact on existing properties (DCAT-AP 2.x)
 - distribution, is referenced by, primary topic, was generated by
 - hasPart, resource, replaces

Proposition

Add a short usage note section to confirm this.

§ 7. Use of inverse properties

The properties described in <u>6. Vocabulary specification</u> do not include <u>inverses</u> intentionally, with the purpose of ensuring interoperability also in systems not making use of OWL reasoning.

However, recognizing that inverse's are needed for some use cases, DCAT supports them, but with the requirement that they MAY be used only in addition to those described in 6. Vocabulary specification, and that they MUST NOT be used to replace them.

Property	Inverse	
dcat:prev	dcat:next	
dcat:previousVersion	dcat:nextVersion	
dcat:distribution	dcat:isDistributionOf	
dcterms:hasPart	<pre>dcterms:isPartOf</pre>	
dcat:resource	dcat:inCatalog	
dcterms:replaces	dcterms:isReplacedBy_	
dcterms:isReferencedBy	dcterms:references	
dcat:hasVersion	dcat:isVersionOf	
dcat:inSeries	dcat:seriesMember	
foaf:primaryTopic	<pre>foaf:isPrimaryTopicOf</pre>	
prov:wasGeneratedBy	prov:generated	

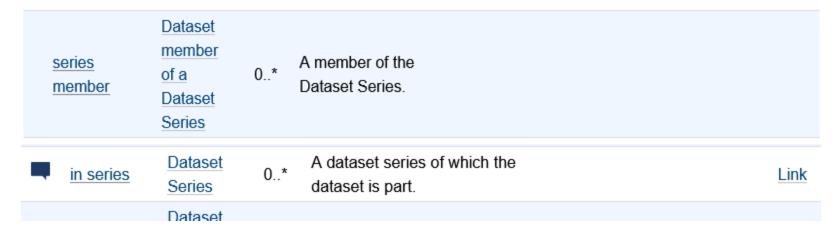
Property series member or in series (Issue #296)

Description

W3C imposes the use of *in series* (cfr inverse properties)

Motivation

Is there need to include and adopt the other as prime property?



Proposition

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

Property *prev* or *next* (Issue #300)

Description

W3C imposes the use of *prev* (cfr inverse properties)

Motivation

Is there need to include and adopt the other as prime property?

Proposition

Proposition: to **remove** property *next* from the profile.





Catalogue organisation (#292)

Description

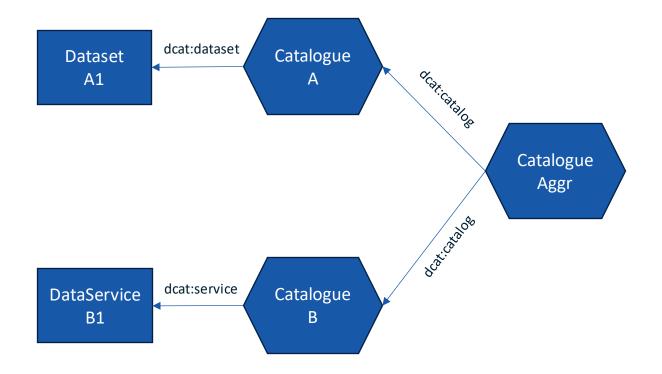
How to express after harvesting that the aggregated catalogue is combined of multiple harvested catalogues?

Motivation

- dcat:catalogis subproperty of dct:hasPart via dcat:resource.
- DCAT-AP proposes the usage of dct:hasPart at the level of Catalogue with the same restrictions as dcat:catalog.
- W3C has "opened up" the usage from dct:hasPart as the generic super property of all "partOf/memberOf" relationships.

Proposition

Replace dct:hasPart with dcat:catalog.



More properties for Data Service (Issue #272)

Description

Data Service has a very reduced amount of properties listed. Make the following properties recommended. The following properties are can be shared optionally already.

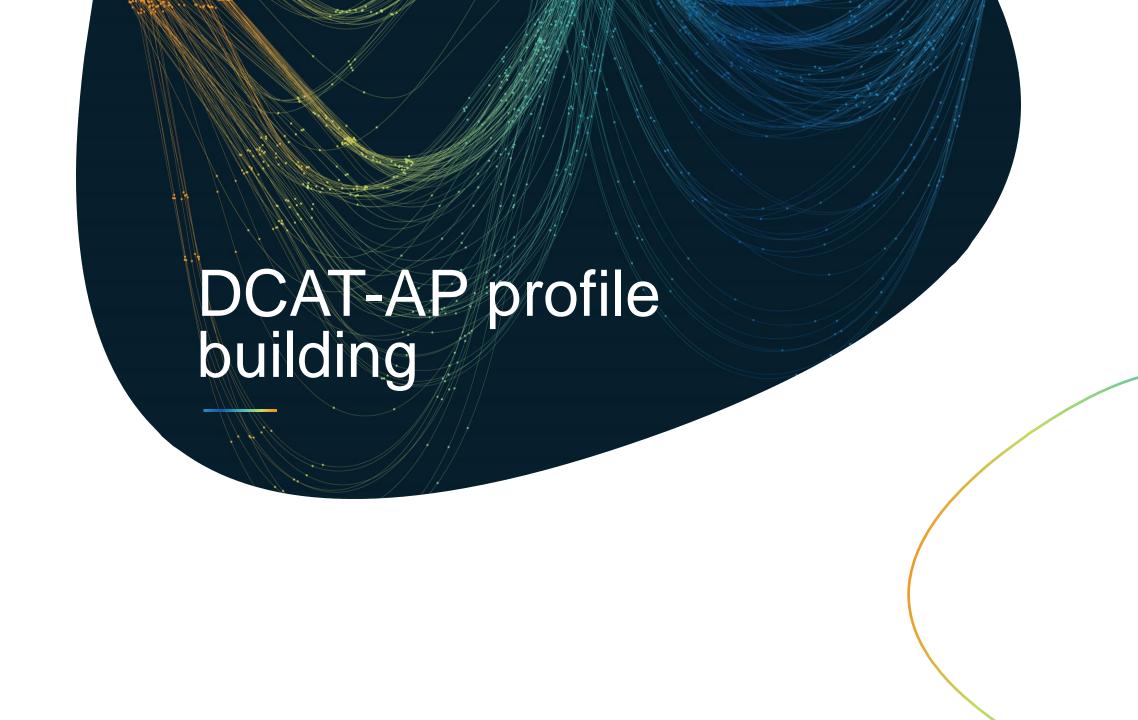
Motivation

- Publisher and documentation earlier discussed
- Contact Point (mandatory in HVD)
- Conforms To (a technical standard or protocol, not a data shape)
- Theme
- Keyword
- Landing page (cfr. documentation/foaf:page)

Proposition

Preference vote on each of them.





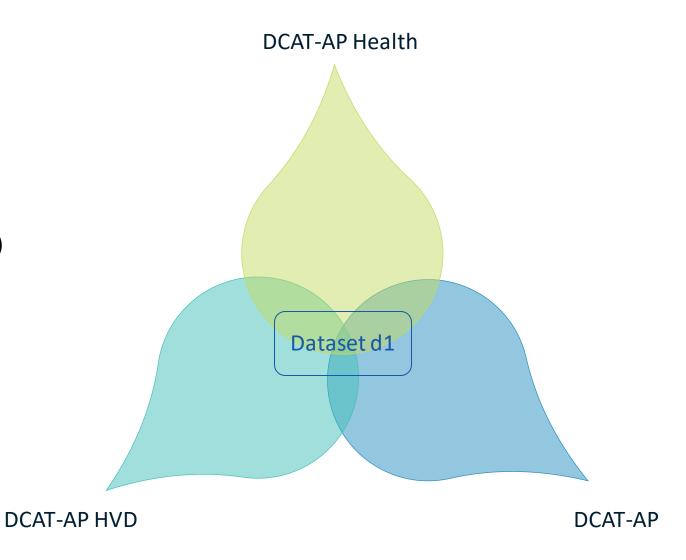
Profile Guidelines

Description

How to create a profile of DCAT-AP?

Balance

- Conciseness (easiness to read, editorial effort)
- Implementation effort
- Once-only effort for publishers of datasets



Profile Guidelines

Description

How to create a profile of DCAT-AP?

Balance

- Conciseness (easiness to read, editorial effort)
- Implementation effort
- Once-only effort for publishers of datasets

Challenges

- A complex topic, with many different opinions
- Abstract

* SEMIC Style guide / The SEMIC Style Guide for Semantic Engineers

The SEMIC Style Guide for Semantic Engineers

This document defines the style guide to be applied to the SEMIC's semantic data specifications, notably to the eGovernment Core Vocabularies and Application Profiles. It provides rules on naming conventions, syntax, artefact management and organisation. It is meant to be complemented with technical artefacts and implementations that enable automatic conformance checking and transformation of conceptual models into formal semantic representations.

The content of these guides is part of the action to promote semantic interoperability amongst the EU Member States, with the objective of fostering the use of standards by, for example, offering guidelines and expert advice on semantic interoperability for public administrations.

This style guide is intended primarily for semantic engineers, data architects and knowledge modelling specialists who are acting as editors or reusers of Core Vocabularies and Application Profiles.

This style guide may constitute a good source of information and explanations for the European Commission's officers, collaborating consultants, and stakeholders involved in interinstitutional standardisation.

- SEMIC Style Guide provides initial anchors for an approach but no complete answer.
- Result of this interaction results in improved guidelines

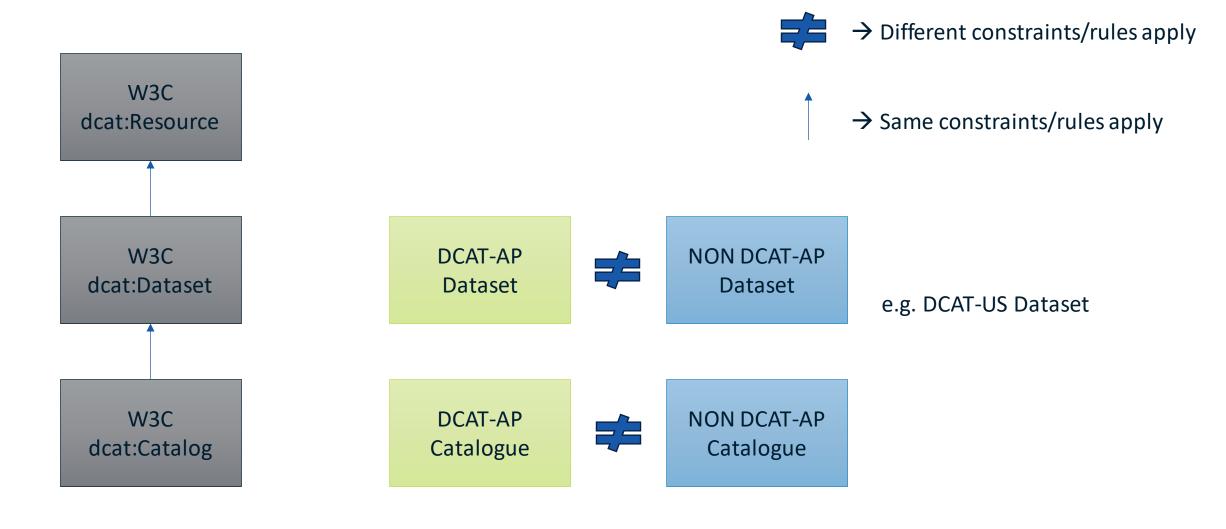
Profile Guidelines

Recurring Questions

- Question1: How to indicate that the properties (constraints) of one class apply to another class? Example: Is a DCAT-AP Dataset Series a DCAT Dataset?
- Question 2: How to reuse a property with conflicting constraints? Example: How to use another controlled vocabulary for dcat: theme?

Question 3: Is profile A coherent/consistent with profile B and what are the changes?

Question 1: indicating that the properties of one class apply to another class



In W3C there are no cardinality constraints and no obligatory codelists as range. Only the scoping/semantics of the properties and classes are fixed.





Same constraints/rules apply

W3C dcat:Resource

DCAT-AP Dataset NON DCAT-AP
Dataset

W3C dcat:Dataset

W3C dcat:Catalog

DCAT-AP Catalogue NON DCAT-AP Catalogue

The collection of all dcat:Resource





Same constraints/rules apply

W3C dcat:Dataset

at:Resource

DCAT-AP Dataset NON DCAT-AP
Dataset

W3C dcat:Catalog

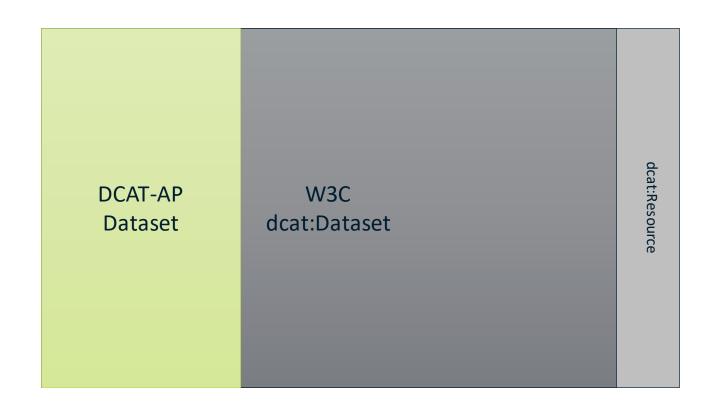
The collection of dcat:Dataset is a subset of the collection of dcat:Resource.

DCAT-AP Catalogue NON DCAT-AP Catalogue





Same constraints/rules apply



NON DCAT-AP
Dataset

W3C dcat:Catalog

DCAT-AP Catalogue NON DCAT-AP Catalogue

The collection of DCAT-AP Dataset is a subset of the collection of dcat:Dataset





Same constraints/rules apply

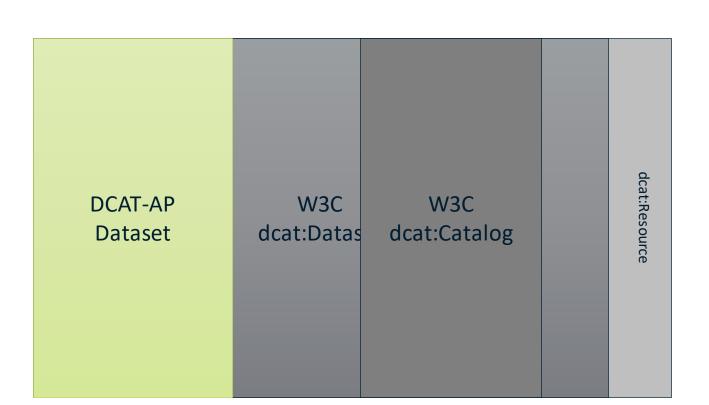


All NON DCAT-AP Datasets are subset of the collection of dcat:Dataset that are not a DCAT-AP Dataset. (Out of scope for DCAT-AP specification.)

W3C dcat:Catalog

DCAT-AP Catalogue

NON DCAT-AP Catalogue



The collection of dcat:Catalog is a subset of the collection of dcat:Dataset

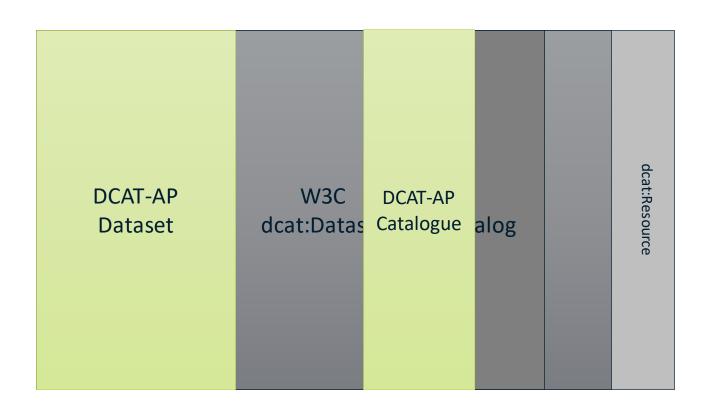


Different constraints/rules apply



Same constraints/rules apply

DCAT-AP Catalogue NON DCAT-AP Catalogue



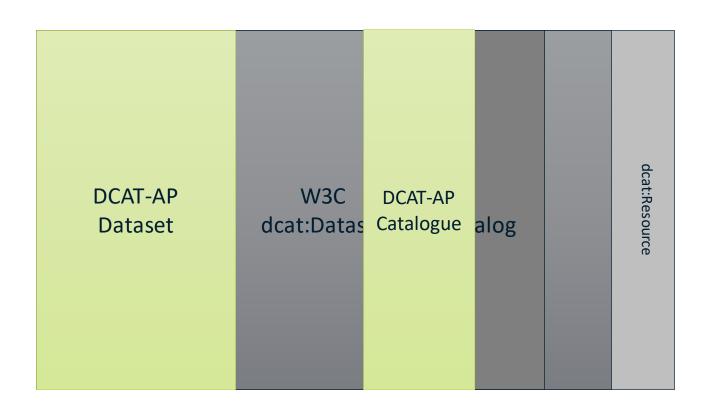
The collection of DCAT-AP Catalogue is a subset of the collection of dcat:Catalog. (NON DCAT-AP catalogues are out of scope for DCAT-AP.)



Different constraints/rules apply



Same constraints/rules apply



The collection of DCAT-AP Catalogue is a subset of the collection of dcat:Catalog. (NON DCAT-AP catalogues are out of scope for DCAT-AP.)



Different constraints/rules apply

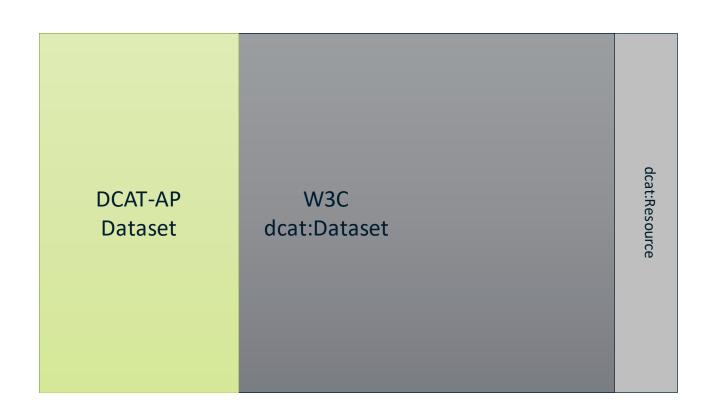


Same constraints/rules apply

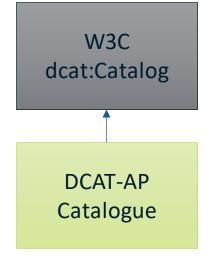


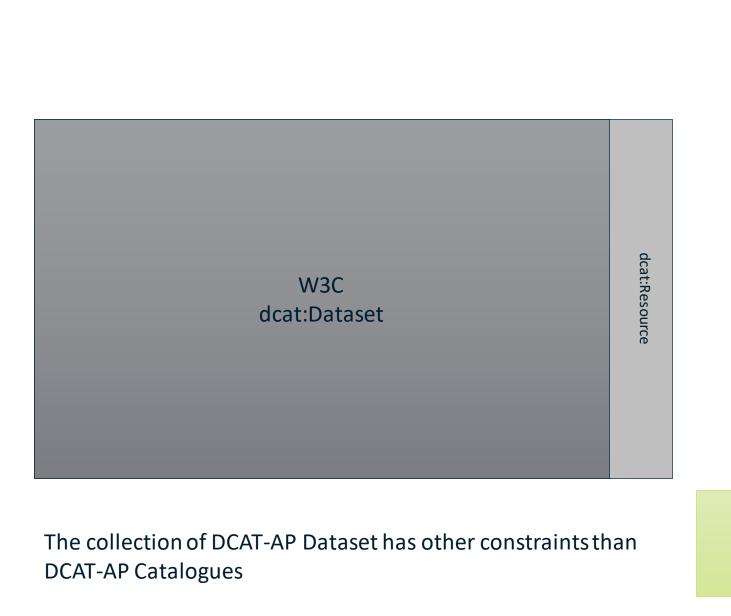


Same constraints/rules apply



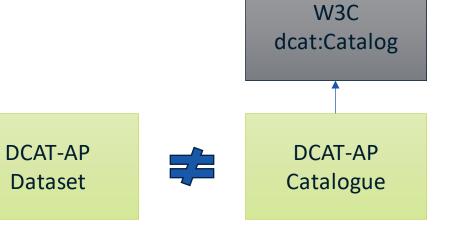
The collection of DCAT-AP Catalogue is a subset of the collection of dcat:Catalog





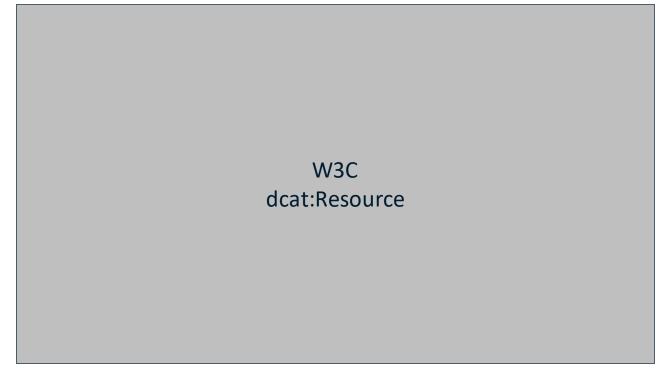




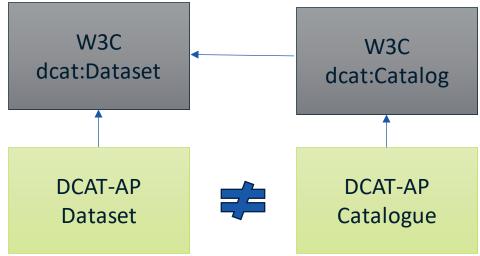




Same constraints/rules apply



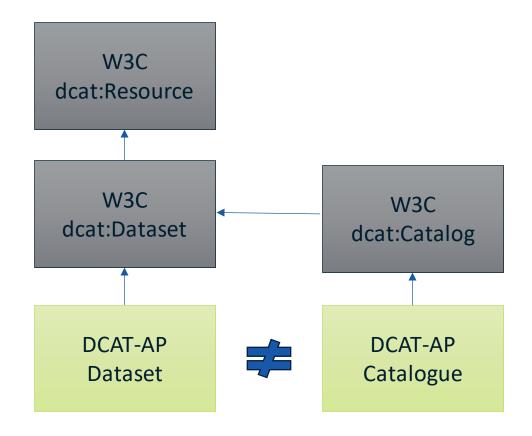
The collection of DCAT-AP Dataset is a subset of the collection of dcat:Dataset



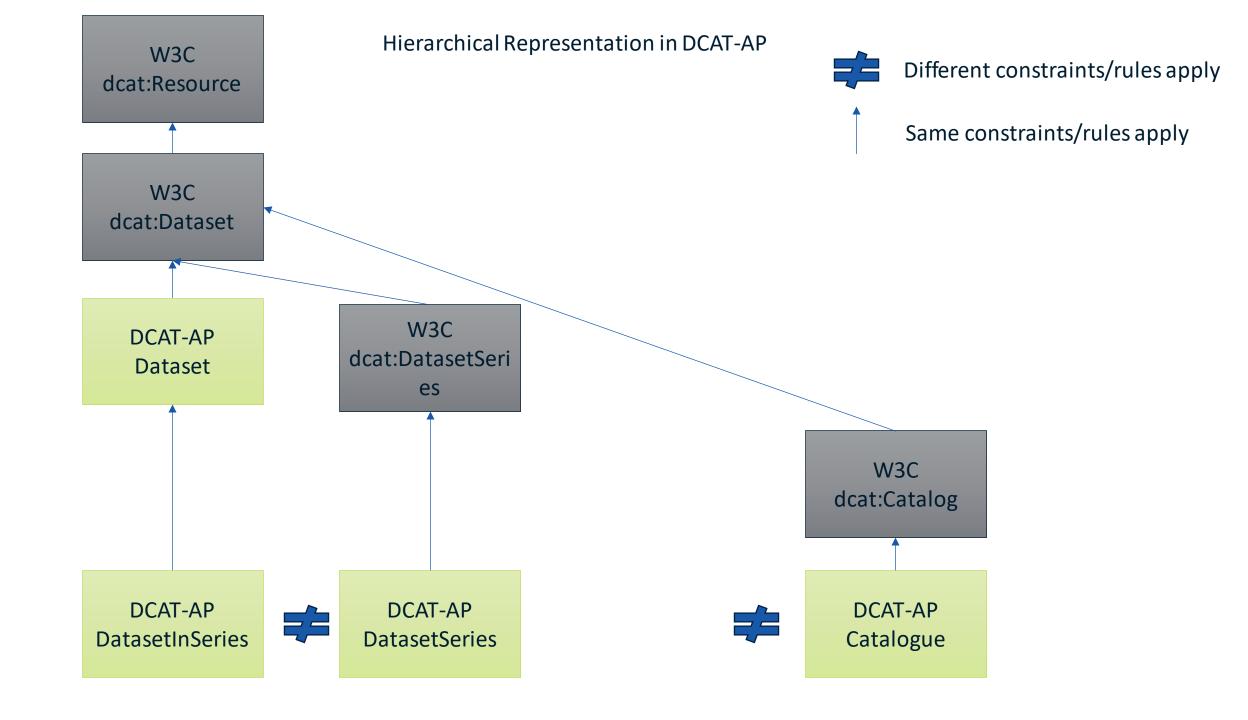


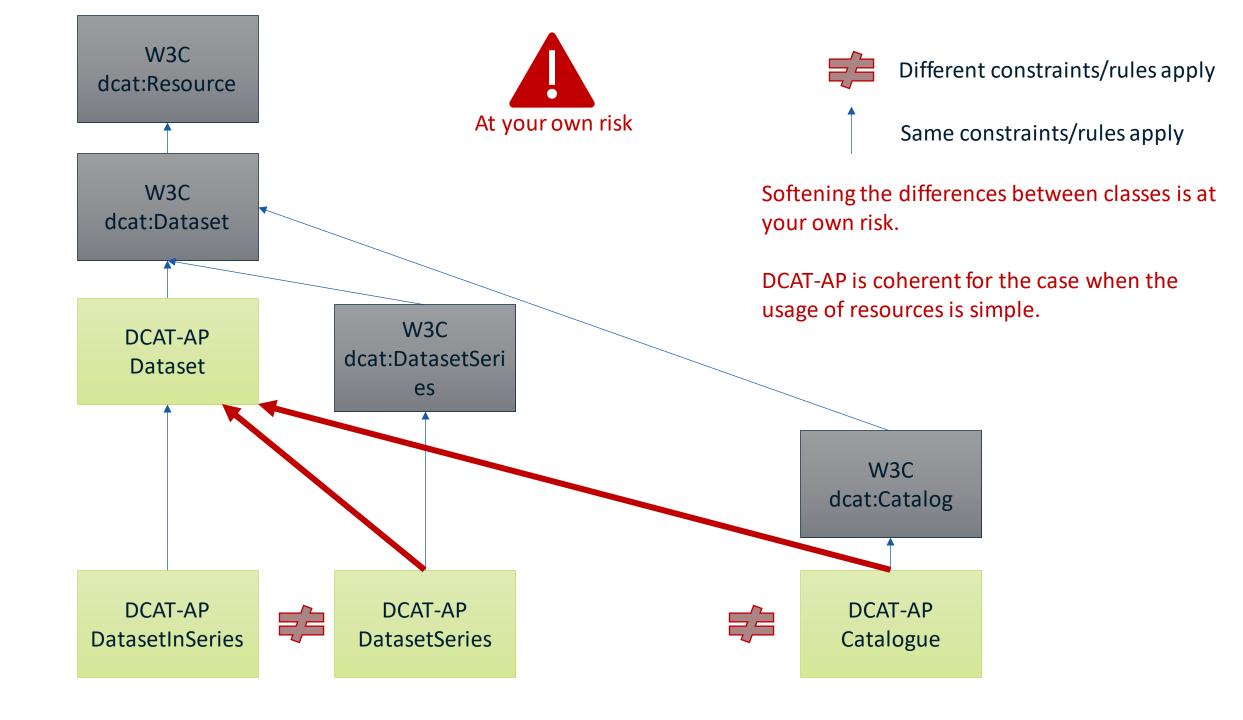


Same constraints/rules apply

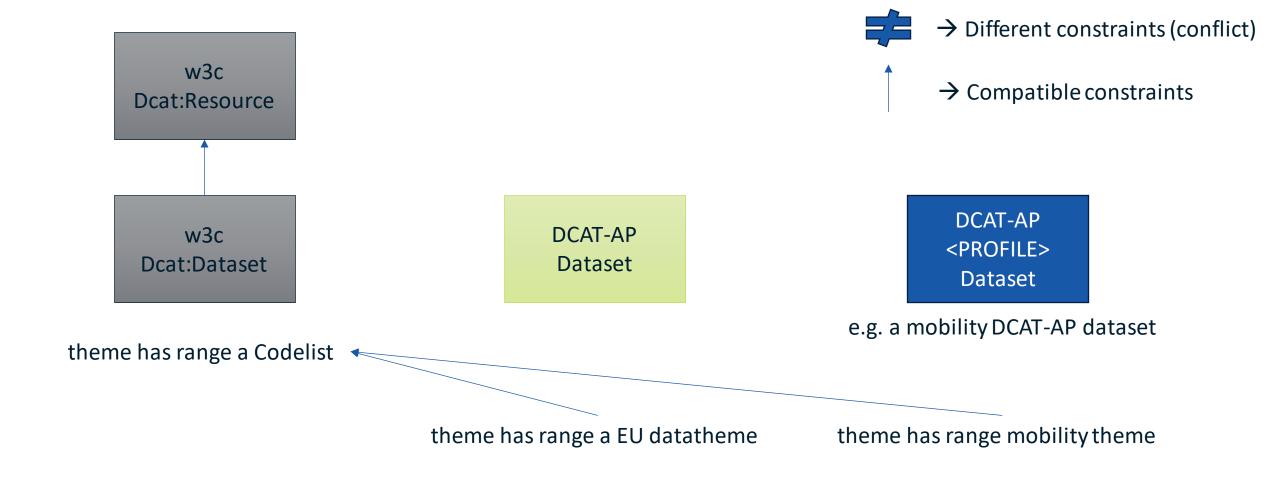


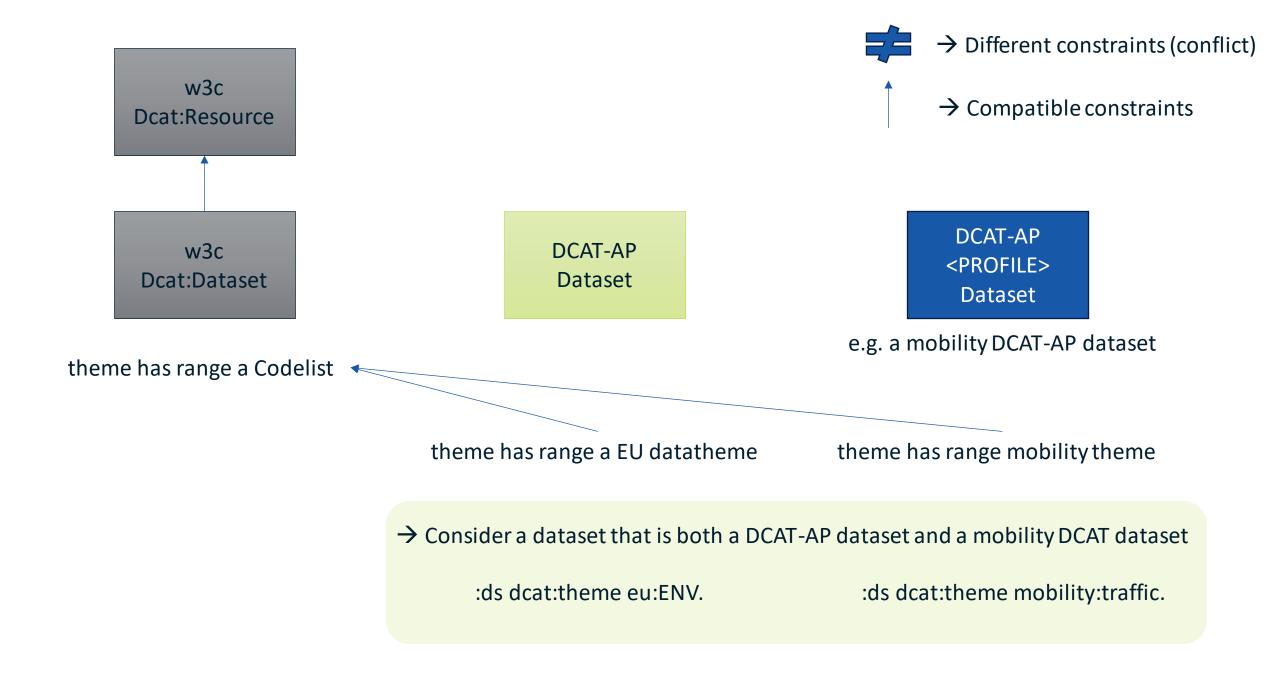
The collection of all dcat:Dataset is a subset of the collection of dcat:Resource

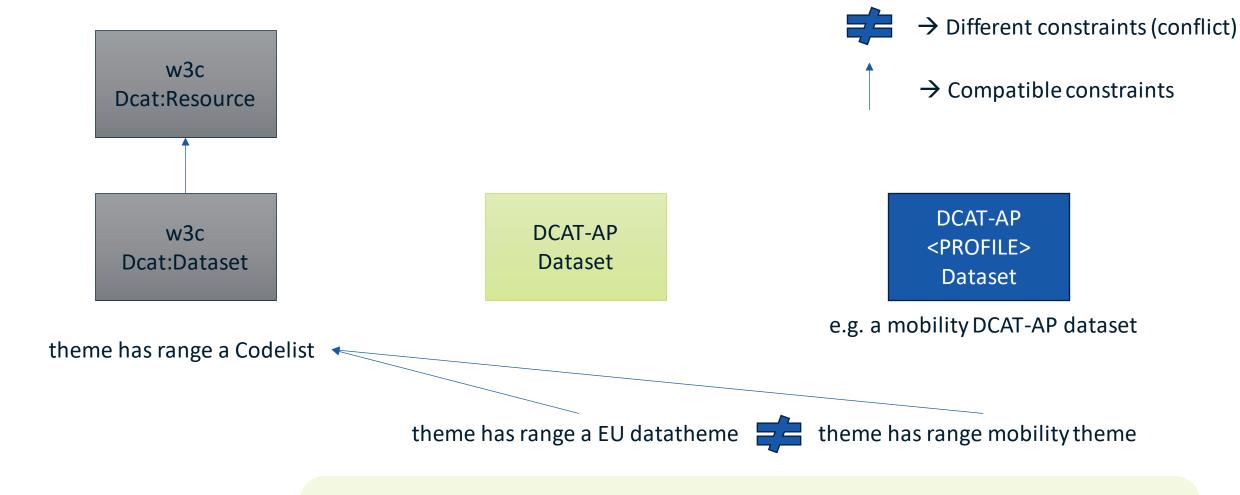




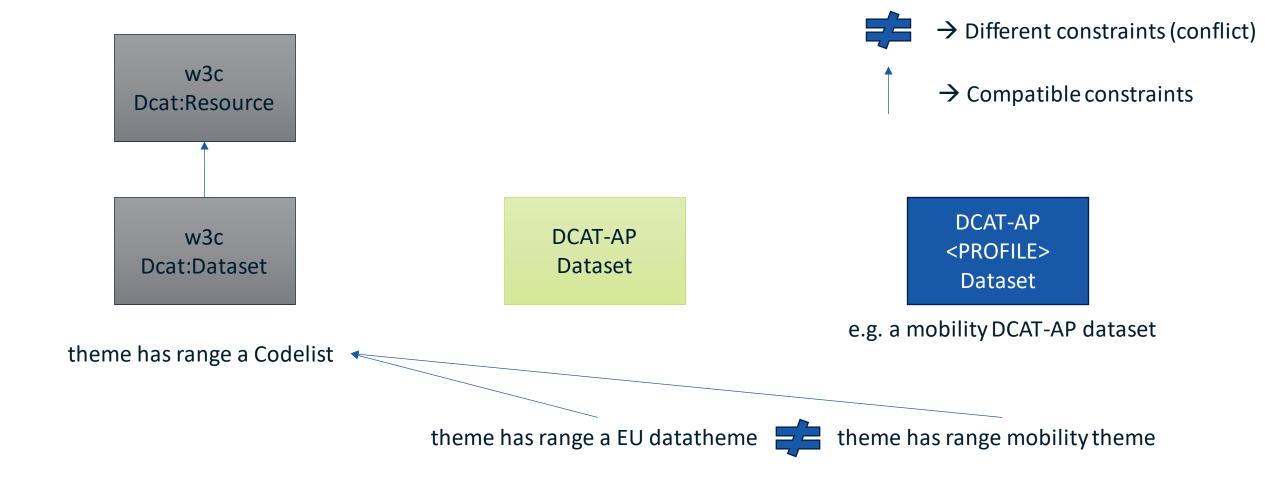
Question 2: reusing a property with conflicting constraints





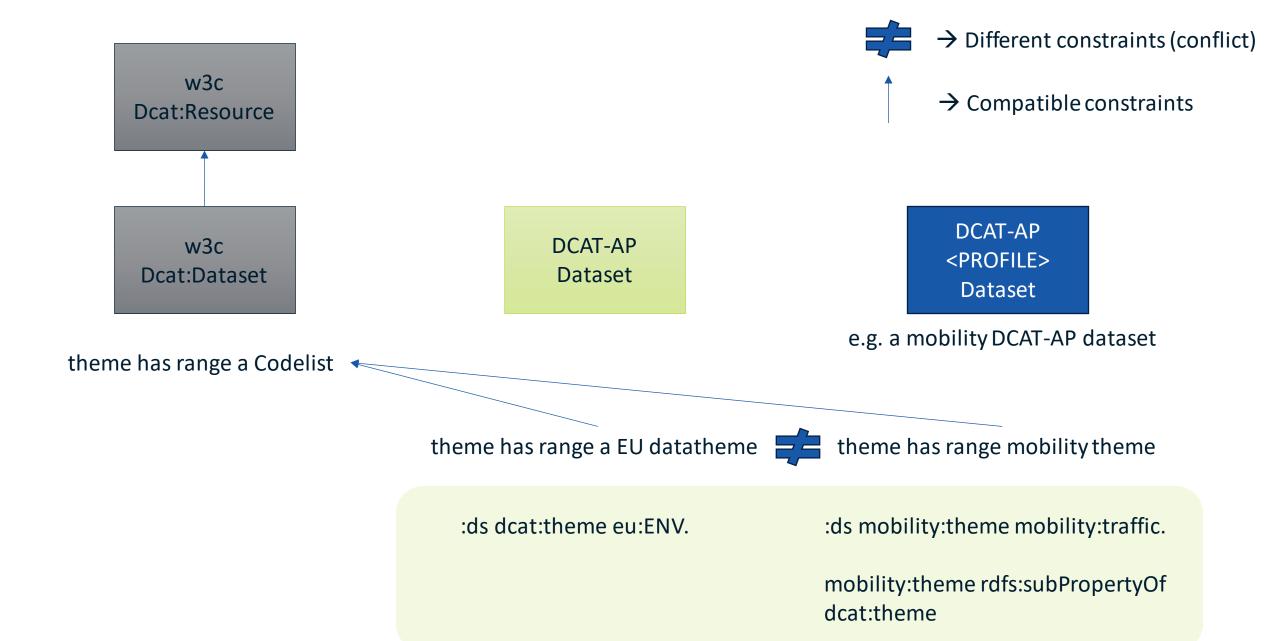


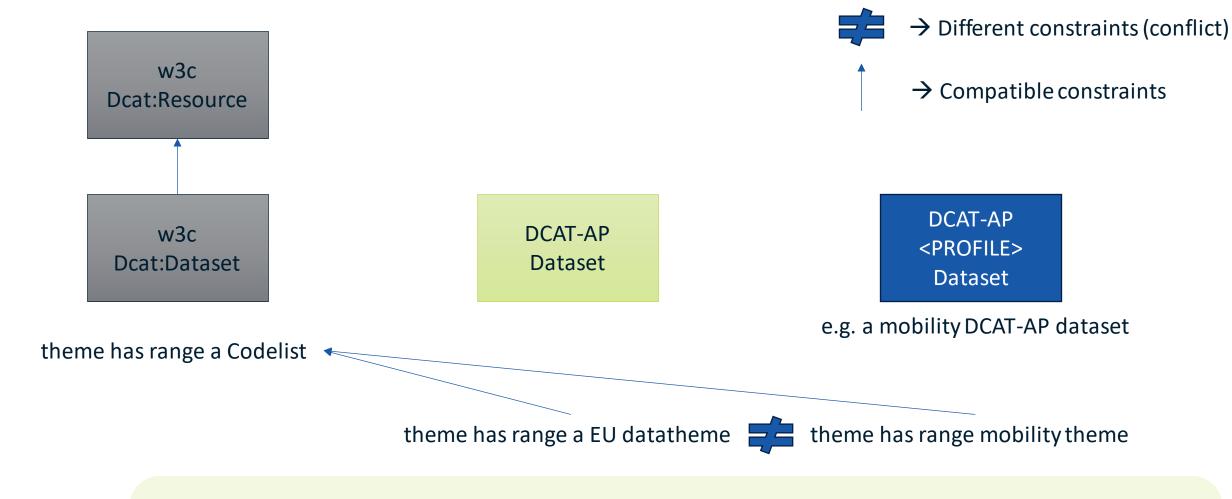
- Max. card 1 in a DCAT-AP means no more than one value.
- Mandatory codelist in DCAT-AP means no other values than those from this codelist.
- → These constraints block the dataset compatible with DCAT-AP and mobility DCAT-AP.



Conflict resolution:

- Choose master DCAT profile: DCAT-AP
- Create a subproperty of theme in the other





- SHACL shapes are independent from each other
- Implementations have to add to an existing DCAT-AP only support for a new property mobility:theme
- Aggregation at the level of DCAT, not at the level of individual profiles.



Profile proposals



To go forward in a practical way



Result in a new section "Creating DCAT-AP profiles".

Profile Guidelines proposals (1)

Proposals are made with the objective to create some steps forward

Description

DCAT-AP is the master profile

Motivation

If DCAT-AP imposes a constraint (usage, cardinalities or mandatory codelists) then another profile must follow these.

Impact

- Properties that are typical for an application scope, like theme, are impacted.
- Profiles that have conflicting constraints with DCAT-AP are a DCAT profile instead of a DCAT-AP profile. It is advised to highlight these conflicts.
- DCAT-AP slowly becomes more generic over time.

Profile Guidelines proposals (2)

Proposals are made with the objective to create some steps forward

Description

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.

Motivation

- Semantically correct w.r.t. DCAT
- Easier for implementers.
 - Implementation is adding a new property, instead of adapting existing property
- Easier for profile editors
- Easier for interoperability between profiles
 - Decisions on this property get decoupled from each other

Impact

• Properties that are typical for an application scope, like theme, are impacted.

Example:

profile:theme is subproperty of dcat:theme

Profile Guidelines proposals (3)

Proposals are made with the objective to create some steps forward

Description

- 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.

Motivation

In line with the reuse guidelines of the SEMIC Style Guide.

Impact

No direct impact.

Profile Guidelines proposals (4)

Proposals are made with the objective to create some steps forward

Description

- 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.

Repetition is not mandatory

Motivation

In line with the reuse guidelines of the SEMIC Style Guide.

Impact

Add an extra column (or indicator) to each property to flag the repetition.

Profile Guidelines proposals (5)

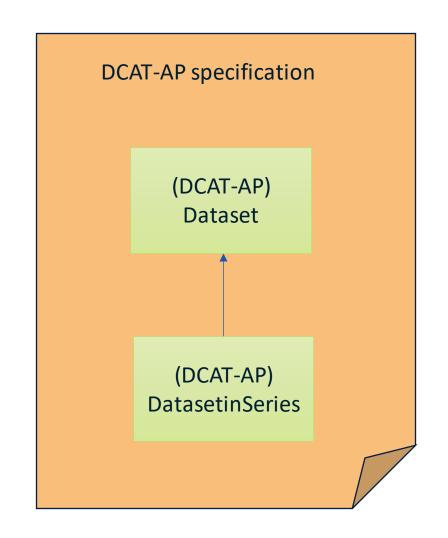
Description

- A Profile is a document which defines a usage scope for all terms mentioned in that data specification.
- The creation of a subclass of a Catalogued Resource is a visual representation of a specific usage scope (more fine grained than the Profile's scope, i.e. a *local scope*).
 - Example Dataset Member in Dataset Series
 - Motivations:
 - Names a usage scope
 - Allows to express additional constraints for that usage scope without impacting the superclass
- Behaves like a profile in a profile.

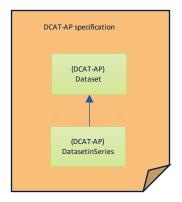
Motivation

Data specification is easier to read:

- no sentences in usages notes like "if this property is used in scope A then apply Y."
- No additional specification to be created.



Profile Guidelines proposals (5)

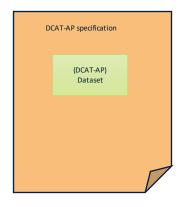


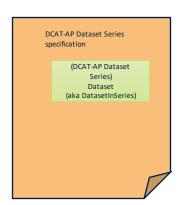
Locally scoped subclasses

- + clear indication of local scope
- + semantics can be specified unambiguously
- implementations guidelines required (Data Exchange Guidelines)

Small and limited local scopes

Proposals are made with the objective to create some steps forward





Local scope in separate specification

- + clear indication of scope
- + semantics can be specified unambiguously
- + no additional implementation guidelines
- more documents to manage

Larger and more extensive scopes



Next steps



Process collected feedback into a new DCAT-AP 3.0 specification

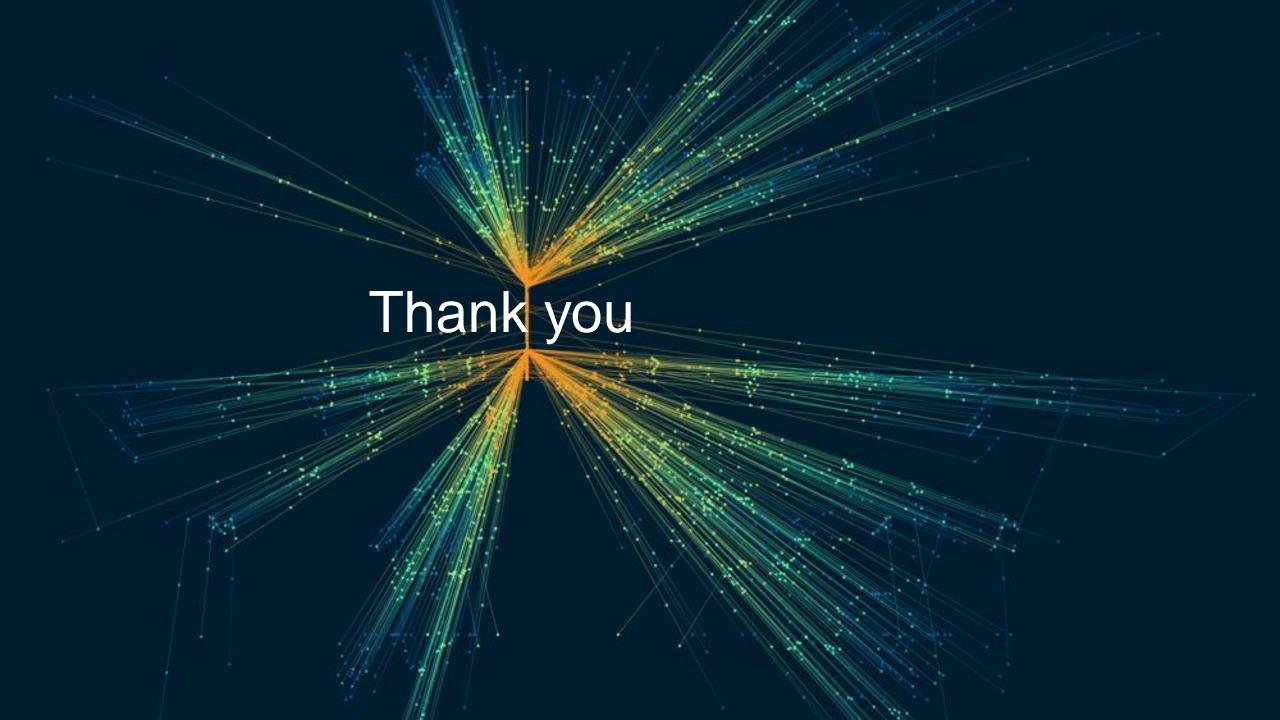


Blog article on profile building



Finalise DCAT-AP for High-Value Datasets by 14/12

Please provide your additional feedback on GitHub.





intercoerable europe

innovation ∞ govtech ∞ community

Stay in touch



(@InteroperableEU) / Twitter



<u>Interoperable Europe - YouTube</u>



<u>Interoperable Europe | LinkedIn</u>



<u>DIGIT-INTEROPERABILITY@ec.europa.eu</u>



https://joinup.ec.europa.eu/collection/interoperableeurope/interoperable-europe