



DCAT, DCAT-AP

[Comprendre les formats et les enjeux](#)

[DCAT](#)

[Chronologie](#)

[Modélisation](#)

[DCAT-AP](#)

[Chronologie](#)

[Modélisation](#)

[Introduction de métadonnées obligatoires, recommandées et facultatives](#)

[Critique de la modularité et incidence sur l'adoption](#)

[DCAT VS DCAT-AP](#)

[Chronologie](#)

[DCAT et Schema.org](#)

[Quel avenir pour DCAT-AP ?](#)

Comprendre les formats et les enjeux

DCAT, DCAT-AP, qu'en est-il vraiment de ses deux formats si proches mais pourtant subtilement différents ?

Si l'on devait résumer, l'on pourrait dire que DCAT est le standard, reconnu par le W3C, qui prends plus de temps à arriver à maturité et qui a une durée de vie de chaque version plus longue; là où DCAT-AP explore plus rapidement de nouvelles extensions de DCAT, pour essayer de découvrir de nouveaux usages et de nouvelles façons de répondre aux besoins grandissant de description des jeux de données.

Voyons maintenant plus en détail chaque format.

DCAT

Data Catalog Vocabulary (DCAT) est un vocabulaire RDF conçu dans le but de faciliter l'interopérabilité entre des catalogues de données publiés sur Internet. Il permet de décentraliser la publication des catalogues et de faciliter la fédération des jeux de données entre différents catalogues.

Chronologie

Développé initialement par l'université nationale d'Irlande il a été ensuite transféré sous le contrôle du W3C parmi le «_eGov Interest Group_» puis sous forme de recommandation par le «_Government Linked Data_» working group du W3C.

DCAT est le fondement de la description des jeux de données dans le secteur public de l'Union Européenne et a été adapté par le programme ISAA ([Interoperability solutions for public administrations, business and citizens](#)¹) de la Commission Européenne.

La version 2.0 du standard sort le 4 février 2020, et c'est un travail conjoint de plusieurs instituts de recherche nationaux européens, ainsi que de la Commission Européenne. Il s'agit là d'une grosse évolution du standard pour supporter, notamment, des ressources de type API en temps réel ; mais également apporter des informations sur la provenance des données et leur qualité. [Data Catalog Vocabulary \(DCAT\) - Version 2](#)²

La version 3.0 du standard est en cours de travail et apporte, pour le moment, le support de *DatasetSeries*, ainsi que la notion de version et de somme de contrôle (checksum) pour des ressources.

[Data Catalog Vocabulary \(DCAT\) - Version 3](#)³

¹ https://ec.europa.eu/isa2/home_en

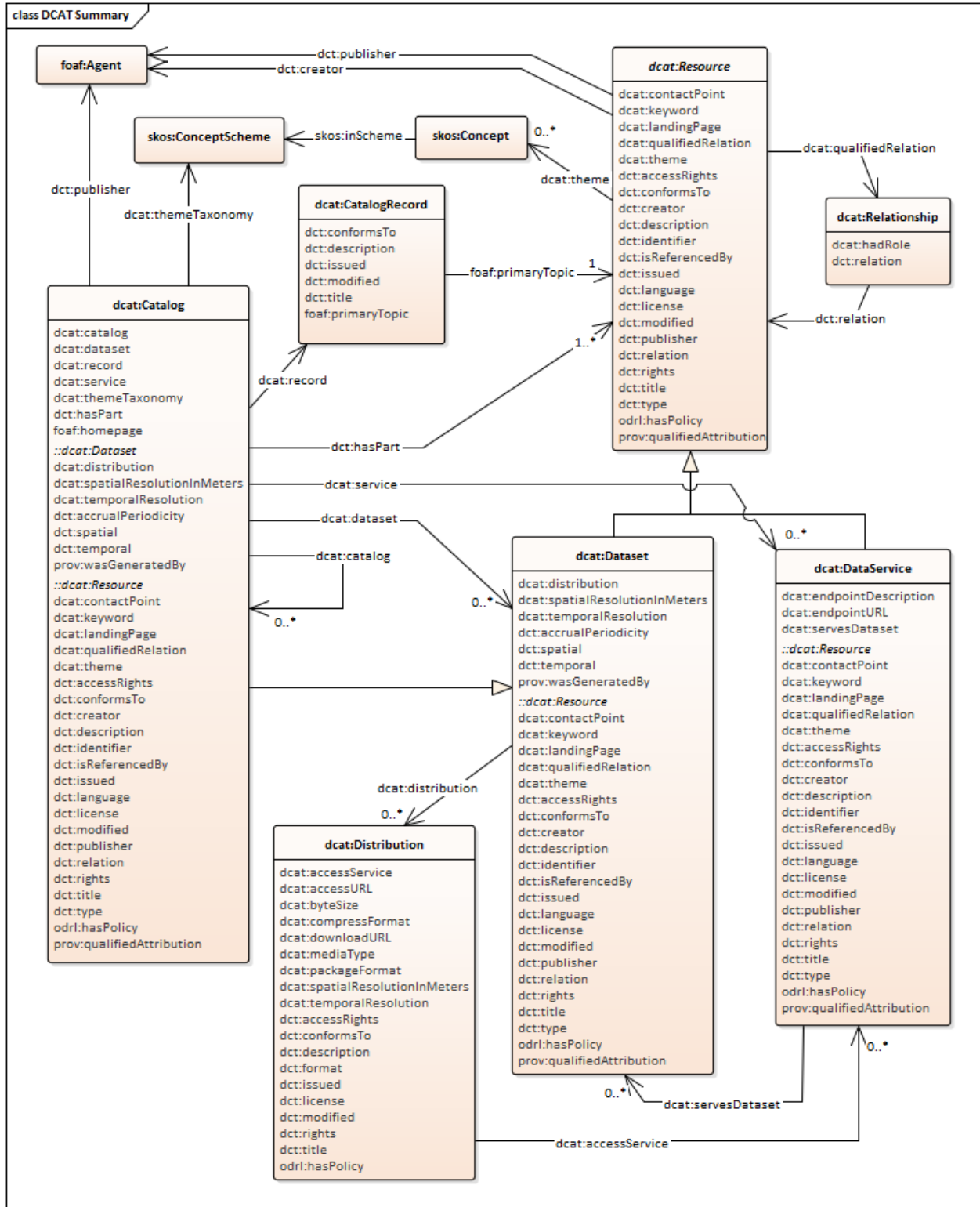
² <https://www.w3.org/TR/vocab-dcat/>

³ <https://www.w3.org/TR/vocab-dcat-3/>

Modélisation

DCAT s'articule autour de six classes principales :

1. **Catalog**
2. **Ressource** qui est une classe métier qui sera étendue par **Dataset**, **DataService** et **Catalog**
3. **Dataset**
4. **Distribution** qui représente une façon d'accéder à un **Dataset** comme un fichier téléchargeable par exemple.
5. **DataService** qui pourrait être une API par exemple.
6. **CatalogRecord** qui représente une métadonnée dans un catalogue. Principalement des informations *techniques* comme qui a ajouté l'élément ou quand, et non *descriptive* comme une donnée d'emplacement géographique.



DCAT-AP

L'idée de pouvoir étendre DCAT à l'aide d'Application Profile est séduisante. Cela permet d'aller « plus vite » que la norme parente, de tester rapidement et de voir ce qui fonctionne. Un peu à l'image des recommandations du W3C pour les navigateurs web et de ce qu'ils supportent en réalité (parfois beaucoup de choses qui ne sont pas encore dans la norme).

Chronologie

En un sens, on peut dire que des DCAT-AP est une extension de DCAT v1, initiée en avril 2013. Cet *Application Profile* commence à être utilisé par l'Union Européenne à partir de septembre 2013.

Après deux ans de travaux et d'amélioration constante, une version 1.1 voit le jour en 2017. Puis, en novembre 2019, la version 2.0 sort, quelques mois avant la publication de DCAT v2.

Ce mécanisme d'extension a permis aussi la création de [GeoDCAT-AP](#)⁴ et [StatDCAT-AP](#)⁵ virent le jour. Ses *Application Profile* spécifiques et indépendants, qui répondent aux besoins de décrire des jeux de données de type spatial ou statistiques.

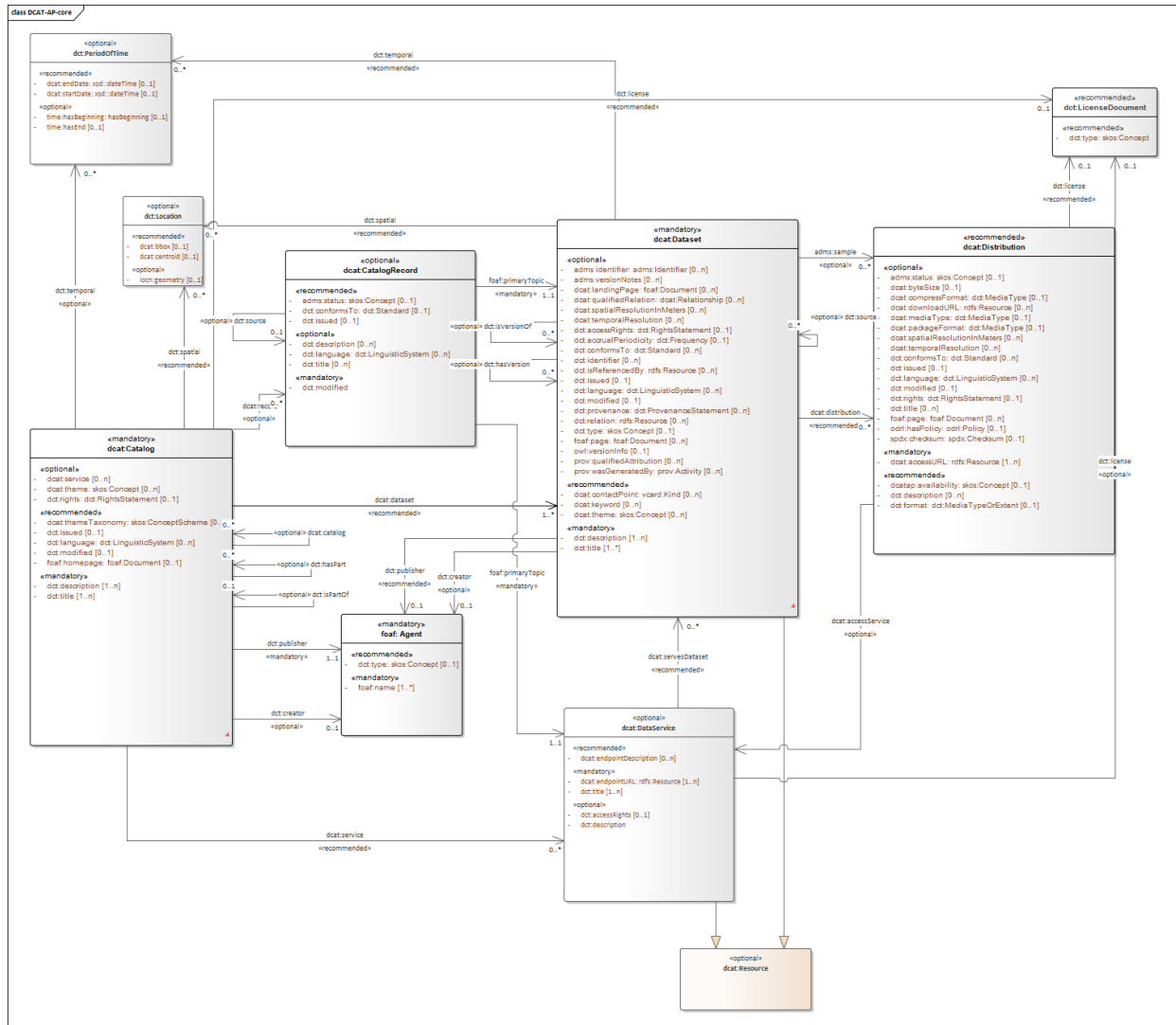
4

<https://joinup.ec.europa.eu/collection/semantic-interoperability-community-semic/solution/geodcat-application-profile-data-portals-europe/release/101>

5

<https://joinup.ec.europa.eu/collection/semantic-interoperability-community-semic/solution/statdcat-application-profile-data-portals-europe/release/101>

Modélisation



Retrouvez page 10 de *dcat-ap_2.1.0.pdf*

Introduction de métadonnées obligatoires, recommandées et facultatives

En plus d'amener plus d'extensibilité en termes de métadonnées, DCAT-AP introduit une notion de métadonnées obligatoires, recommandées (qui pourraient passer obligatoire dans une version future de la norme) et optionnelles.

Ses métadonnées sont décrites à partir de dictionnaires et de spécifications existantes, afin de ne pas réinventer la roue et assurer un maximum de compatibilité avec ce que le end-user pourrait vouloir faire.

- adms: <http://www.w3.org/ns/adms#>

- dcat: <http://www.w3.org/ns/dcat#>
- dcatap: <http://data.europa.eu/r5r/>
- dct: <http://purl.org/dc/terms/>
- foaf: <http://xmlns.com/foaf/0.1/>
- locn: <http://www.w3.org/ns/locn#>
- owl: <http://www.w3.org/2002/07/owl#>
- odrl: <http://www.w3.org/ns/odrl/2/>
- prov: <http://www.w3.org/ns/prov>
- rdfs: <http://www.w3.org/2000/01/rdf-schema#>
- schema: <http://schema.org/>
- skos: <http://www.w3.org/2004/02/skos/core#>
- spdx: <http://spdx.org/rdf/terms#>
- xsd: <http://www.w3.org/2001/XMLSchema#>
- vann: <http://purl.org/vocab/vann/>
- voaf: <http://purl.org/vocommons/voaf#>
- vcard: <http://www.w3.org/2006/vcard/ns#>
- time: <http://www.w3.org/2006/time#>

Par exemple, pour la classe **Catalog** vue ci-avant ([Modélisation](#)), les métadonnées obligatoires dans la version 2.1.0 sont:

- Description (type rdfs:Literal)
- Publisher (type foaf:Agent)
- Title (type rdfs:Literal)

Alors que les métadonnées recommandées sont

- Dataset (type dcat:Dataset)
- Service (type dcat:DataService)
- Homepage (type foaf:Document)
- Language (type dct:LinguisticSystem)
- Licence (type dct:LicenseDocument)
- Release Date (type rdfs:Literal typed as xsd:date, xsd:dateTime, xsd:gYear or xsd:gYearMonth)
- Spatial/Geographic (type dct:Location)
- Themes (type skos:ConceptScheme)
- Update/Modification date (type rdfs:Literal typed as xsd:date, xsd:dateTime, xsd:gYear or xsd:gYearMonth)

Enfin, les métadonnées optionnelles doivent être:

- Has part (type dcat:Catalog)
- Is Part Of (type dcat:Catalog)
- Record (type dcat:CatalogRecord)
- Rights (type dct:RightsStatement)
- Catalogue (type dcat:Catalog)

- Creator (type foaf:Agent)

Chaque version de la norme DCAT-AP reprend, pour chaque classe, l'ensemble de ses éléments.

Critique de la modularité et incidence sur l'adoption

Le fait que DCAT-AP soit plus modulable, ou du moins représente une certaine évolution de DCAT est aussi une grande faiblesse. Les rares pays qui se décident à rejoindre l'effort finissent par tous faire leur propre Application Profile, jugeant sans doute que leurs besoins sont plus spécifiques que la norme européenne. Ses pays sont:

- L'Italie, qui a également étendue GeoDCAT-AP
- La Suède
- La Suisse
- L'Allemagne
- La Belgique, qui n'a même pas respecté la convention de nommage qui aurait voulu que ce soit DCAT-AP_BE mais ont publié DCAT-BE.

Dans ce contexte, peu de pays se décident à adopter DCAT-AP. Les utilisateurs sont peu nombreux, malgré le discours rassurant de l'Union Européenne. Pour preuve, [la communauté en ligne de l'UE](#)⁶ ne compte plus de nouveaux messages depuis 2016. La Commission européenne n'hésite pas cependant à lister quatre fois la Suède et trois fois la Belgique, simplement parce que le pays propose différentes façons d'accéder à des données utilisant ce format de description (entre autres).

[Discussion sur Joinup](#)⁶ et [Liste des membres sur Joinup](#)⁷.

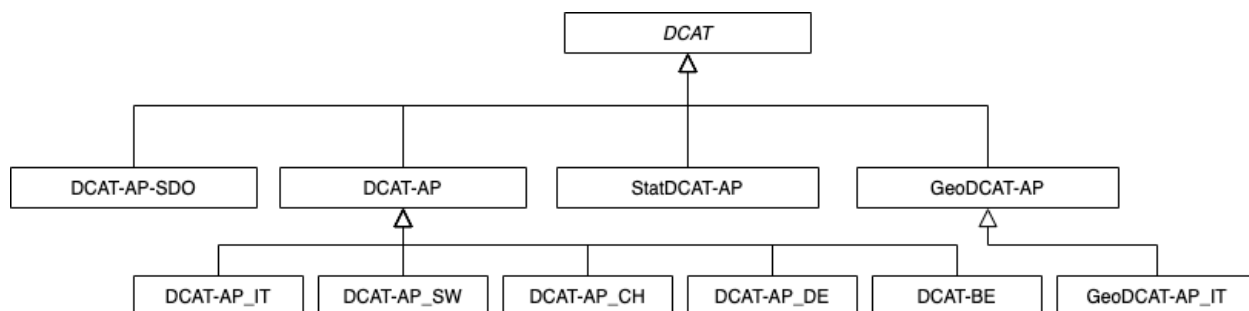
⁶

<https://joinup.ec.europa.eu/collection/semantic-interoperability-community-semic/solution/dcat-application-profile-data-portals-europe>

⁷

<https://joinup.ec.europa.eu/collection/semantic-interoperability-community-semic/solution/dcat-application-profile-data-portals-europe/members>

DCAT VS DCAT-AP



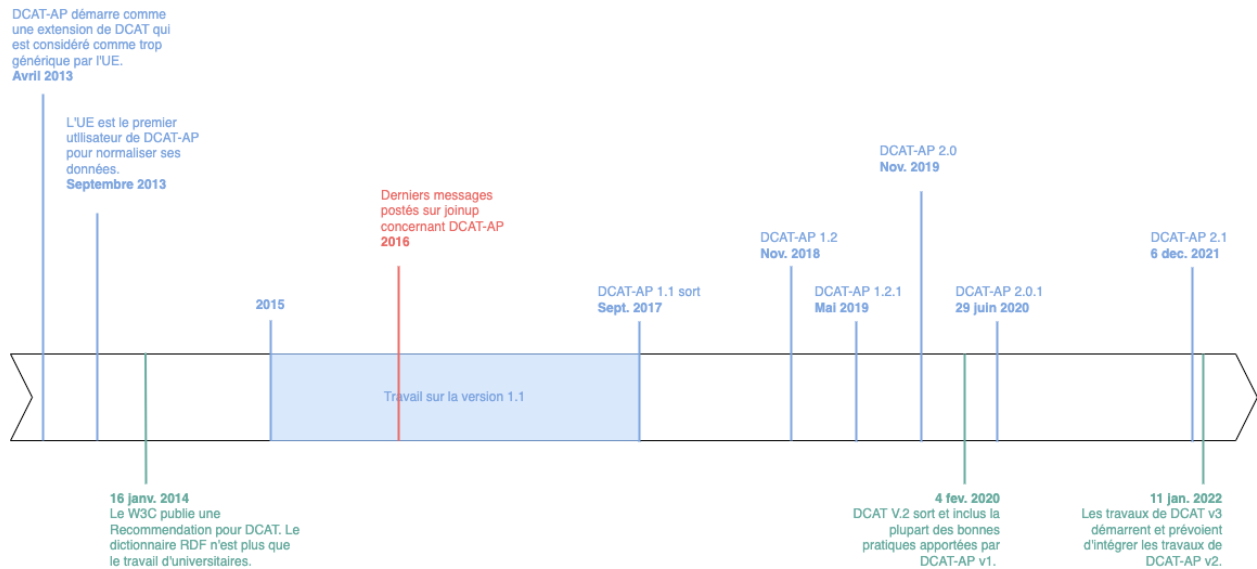
Pour une vue plus détaillée sur la différence entre DCAT et DCAT-AP, consultez l'annexe A - Diagramme classe DCAT / DCAT-AP.

Chronologie

Dès 2014, le W3C intègre les travaux réalisés dans des DCAT-AP à l'intérieur de la version 2 de DCAT. En effet, il leur a paru important d'intégrer les bonnes pratiques issues des applications de profil à l'intérieur de la norme.

De ce fait, une compatibilité avec DCAT v2 implique une compatibilité avec DCAT-AP v1.2.1. C'est ainsi que la France assure avoir une compatibilité avec DCAT-AP alors que udata offre "simplement" une compatibilité avec DCAT v2.

Comme les travaux de DCAT v3 sont seulement en cours, on peut raisonnablement espérer que DCAT v3 apportera une compatibilité avec DCAT-AP 2.1 à minima.



Les deux normes évoluent donc côte à côte, avec des itérations plus courtes et de nouveaux changements à supporter plus nombreux pour DCAT-AP. Avec, et c'est la promesse faite, une façon plus riche de décrire les jeux de données et donc permettre de plus facilement véhiculer cette richesse de métadonnées entre les différents portails.

Mais voilà, encore faut-il que tout le monde utilise DCAT-AP, dans la même version que vous (Cf [Critique de la modularité et incidence sur l'adoption](#)). Ce qui n'est pas forcément le cas, et, en pratique, on retombe souvent sur un support à minima de DCAT puis si possible de DCAT-AP (version x ou y).

DCAT et Schema.org

En préambule, il est à noter que DCAT et Schema.org ne poursuivent pas tout à fait le même but. Si le premier essaie de décrire explicitement les données, le second a pour mission première de rendre les données découvrables sur Internet.

Cela étant, des travaux sont en cours pour utiliser le mécanisme d'extensions d'*application profiles* de DCAT afin de pouvoir décrire les métadonnées dans un format similaire à ce sur quoi travaille schema.org et ainsi rendre découvrable les données par les moteurs de recherche sur le web. En ce sens, Schema.org a déjà intégré le vocabulaire DCAT dans sa norme.

[DCAT-AP to Schema.org Mapping](#)

Cependant, il reste impossible d'explicitement un jeu de données via DCAT depuis des données schema.org uniquement. La *conversion* ou l'équivalence n'étant possible que dans l'autre sens.

Ses travaux sont donc préliminaires et n'apportent, pour le moment, aucune information sur comment exposer ses données ou comment les moteurs de recherche vont les exploiter.

Quel avenir pour DCAT-AP ?

Si DCAT finit toujours pas réintégrer dans sa norme les avancées faites via les *applications profiles* et que Schema.org avance sur la découvrabilité des données, on peut donc vraisemblablement se demander si DCAT-AP est important. S'il est important, pour un pays comme le Grand Duché de Luxembourg d'investir du temps et des ressources dans le support d'une norme qui change rapidement et, parfois, radicalement.

Cependant, de bonnes pratiques sont à prendre de DCAT-AP, comme les données vitales pour décrire un jeu de données. En ce sens, la cellule open-data de l'Etat peut se fixer pour but de supporter les métadonnées *mandatory* de DCAT-AP, sans aller dans le support poussé et la description avancée des jeux de données. Ceci constitue une première étape vers des jeux de données décrits de manière qualitative sur la plateforme. Critère qui, rappelons-le, est amené à prendre de plus en plus d'importance dans l'évaluation annuelle faite par la Commission Européenne.



Annexe A

Diagramme classe DCAT / DCAT-AP

dcat:Catalog
dcat:catalog
dcat:dataset
dcat:record
dcat:service
dcat:themeTaxonomy
dct:hasPart
foaf:homepage
::dcat:Dataset
dcat:distribution
dcat:spatialResolutionInMeters
dcat:temporalResolution
dct:accrualPeriodicity
dct:spatial
prov:wasGeneratedBy
::dcat:Ressource
dcat:contactPoint
dcat:keyword
dcat:landingPage
dcat:qualifiedRelation
dcat:theme
dcat:accessRights
dct:conformsTo
dct:creator
dct:description
dct:identifier
dct:isreferencedBy
dct:issued
dct:language
dct:license
dct:modified
dct:publisher
dct:relation
dct:rights
dct:title
dct:type
odrl:hasPolicy
prov:qualifiedAttribution

dcat:Catalog
dct:title
dct:description
dct:publisher
::dcat:Dataset
dct:language
foaf:homepage
dcat:dataset
dcat:themeTaxonomy
dct:spatial
dct:temporal
dct:issued
dct:modified
dcat:catalog
dcat:record
dcat:service
dct:hasPart
dcat:theme
dct:creator
dct:rights
dct:isPart
dct:source

Propriété ajoutée par rapport à DCAT

Propriété supprimée dans DCAT-AP

Propriété mandatory dans DCAT-AP

Propriété recommended dans DCAT-AP

Propriété optional dans DCAT-AP