

DCAT Application Profile for Data Portals in Europe

Brecht Wyns, Makx Dekkers, Nikolaos Loutas, Vassilios Peristeras, Athanasios Karalopoulos

The **DCAT Application Profile for data portals in Europe** ([DCAT-AP](#)) is a specification based on [W3C's Data Catalogue vocabulary \(DCAT\)](#) for describing public sector datasets in Europe. The [DCAT-AP](#) was developed by a working group of experts following an open collaborative process. Since its initial development in 2013, this process has continued to further develop the specification, leading to the publication of DCAT-AP v1.1 in October 2015.

The basic use case for [DCAT-AP](#) is to enable cross-data portal search for data sets and make public sector data better searchable across borders and sectors. This can be achieved by the exchange of descriptions of datasets among data portals. From the start, [DCAT Application Profile](#) had the purpose of adapting [DCAT](#) to facilitate the reuse of data. In concrete:

- It proposes mandatory, recommended or optional classes and properties to be used for a particular application;
- It identifies requirements to control vocabularies for this particular application;
- It gathers other elements to be considered as priorities or requirements for an application such as conformance statement, agent roles or cardinalities.

[DCAT-AP](#) has been implemented by over 15 open data portals in the European Union, including the [European Data Portal](#). Moreover, some EU Member States have created extensions for the [DCAT-AP](#), such as the [DCAT-AP_IT](#). In order to facilitate the implementation of [DCAT-AP](#), many implementers have developed tools such as validators, harvesters and exporters of [DCAT-AP](#) metadata. An overview of those tools is available via [Joinup](#).

In order to better respond to the information requirements of the statistical and geospatial domains while ensuring consistency with [DCAT-AP](#), the [ISA² Programme of the European Commission](#) has created two extensions to DCAT-AP:

- [StatDCAT-AP](#), developed in close collaboration with [Eurostat](#), brings the statistical and open data communities closer by enhancing the visibility and facilitating the inclusion of statistical data sets in open government data portals; and
- [GeoDCAT-AP](#) describes geospatial datasets, dataset series, and services. It provides an RDF syntax binding for the union of metadata elements defined in the core profile of ISO 19115:2003 and those defined in the framework of the [INSPIRE Directive](#). Its basic use case is to make spatial datasets, data series, and services searchable on general data portals, thereby making geospatial information better searchable across borders and sectors. A specific working group from the [Joint Research Centre](#) and [ISA² programme](#) were at the basis of this extension.

During the initial development of [DCAT-AP](#) in 2013 as well as in the process of revising the specification in 2015, a need was identified to provide more guidance for the application of the profile in practical situations, e.g. to identify existing practices or to formulate advice for implementers who need to map their local metadata to [DCAT-AP](#)-compliant metadata. Therefore, a DCAT-AP working group in the scope of the [ISA² Programme of the European Commission](#) developed implementation guidelines based on contributions from the implementer community. The ISA² Programme and the working group gathered and documented implementation challenges and proposed possible solutions and work-arounds. These [guidelines](#) included:



How to use the MDR data themes vocabulary?

Refer to <https://joinup.ec.europa.eu/node/150359/>



How to model and express provenance?

Refer to <https://joinup.ec.europa.eu/node/150349/>



How to model Dataset series?

Refer to <https://joinup.ec.europa.eu/node/150348>



How to refer to licence documents and licence URIs?

Refer to <https://joinup.ec.europa.eu/node/150347>



How to use identifiers for datasets and distributions?

Refer to <https://joinup.ec.europa.eu/node/150346/>



How to extend DCAT-AP?

Refer to <https://joinup.ec.europa.eu/node/150345/>



How to describe contact information?

Refer to <https://joinup.ec.europa.eu/node/150343/>



How to manage duplicates?

Refer to <https://joinup.ec.europa.eu/node/150341/>



Overview of tools

Refer to <https://joinup.ec.europa.eu/node/150350/>

The activity that led to the guidelines raised several issues that could not be further developed. A main issue that was raised and incited a lot of discussion without reaching a conclusion was the way that data can be accessed through services rather than as a fixed file. Another issue, for which a workaround was provided through the guidelines but no general consensus could be reached without impacting the [DCAT](#) specification itself, relates to the expression of relationships between datasets. These and other technical issues are the subject of current activities to further develop guidelines for [DCAT-AP](#) and to revise the [DCAT-AP](#) specification. A [working group meeting](#) during which the issues will be prioritised will take place on 25/10/2016.