

## REUTILIZACIÓN DE LA INFORMACIÓN DEL SECTOR PÚBLICO

# DCAT-AP Y LA NORMA TÉCNICA DE INTEROPERABILIDAD DE REUTILIZACIÓN DE RECURSOS DE INFORMACIÓN (NTI-RISP)



## ÍNDICE

<b>OBJETIVOS DIDÁCTICOS</b> .....	<b>3</b>
<b>REQUISITOS PREVIOS</b> .....	<b>3</b>
<b>CONTENIDOS</b> .....	<b>3</b>
<b>1. INTRODUCCIÓN</b> .....	<b>4</b>
1.1 Conceptos y necesidad de interoperabilidad.....	5
1.2 ¿Qué es DCAT? .....	6
<b>2. DCAT-AP</b> .....	<b>6</b>
2.1 ¿Qué es DCAT-AP?.....	7
2.2 Orígenes y estado actual del proyecto DCAT-AP .....	7
2.3 La especificación DCAT-AP .....	9
2.3.1 CONCEPTOS PRINCIPALES .....	9
2.3.2 VOCABULARIOS CONTROLADOS .....	10
2.3.3 OTROS ASPECTOS.....	12
2.4 Caso de uso: Portal paneuropeo .....	13
<b>3. NTI-RISP</b> .....	<b>15</b>
3.1 Introducción al Esquema Nacional de Interoperabilidad.....	15
3.2 Objetivos de la norma y ámbito de aplicación.....	15
3.3 Preceptos de la norma .....	16
3.3.1 SELECCIÓN DE LA INFORMACIÓN.....	16
3.3.2 CATÁLOGO DE INFORMACIÓN PÚBLICA .....	17
3.4 Esquema de URI .....	18
3.5 Metadatos .....	19
3.6 Vocabularios.....	20
<b>RESUMEN</b> .....	<b>22</b>

## OBJETIVOS DIDÁCTICOS

Comenzamos presentando los **Objetivos didácticos** de esta Unidad:

- ✓ Comprender qué es DCAT-AP y por qué es necesario para la estrategia europea de datos abiertos.
- ✓ Identificar NTI-RISP como parte del ENI y su relación con el DCAT-AP.
- ✓ Conocer las pautas que dicta la NTI-RISP para una iniciativa de reutilización de la información.
- ✓ Ser capaz de emplear las directrices marcadas por NTI-RISP para diseñar el esquema de URIs.
- ✓ Ser capaz de describir un catálogo, un conjunto de datos y una distribución de acuerdo a como dicta la NTI-RISP.

## REQUISITOS PREVIOS

Para poder asimilar los conceptos que vamos a desarrollar en la unidad, deberías contar con:

- Conocimientos técnicos básicos de la utilización o conexión de datos de varias fuentes.
- Experiencia cualitativa con catálogos de datos.
- Conocimiento del intercambio de información entre organizaciones.
- Conocimientos básicos de XML y RDF.

## CONTENIDOS

Resumimos los aspectos a tratar en los apartados de la Unidad:

### 1. INTRODUCCIÓN

Nos introduce en los conceptos y necesidad de **interoperabilidad** y el marco europeo de iniciativas relacionadas con el mismo.

### 2. DCAT - AP

Conocemos este perfil de aplicación o especificación para la **descripción de catálogos** de conjuntos de datos del sector público en Europa.

### 3. NTI - RISP

Se profundiza en la Norma técnica de interoperabilidad, publicada en febrero de 2013, sus objetivos, preceptos y cómo se desarrolla.

## 1. INTRODUCCIÓN

### INTEROPERABILIDAD

La **interoperabilidad** es la capacidad de dos organizaciones de poder intercambiar información entre ellas. Constituye una **necesidad** para aquellas organizaciones que tienen que trabajar o cooperar, aunque solo sea en parte, y afecta a múltiples niveles.

En la UE para las administraciones públicas, se ha definido el **Marco europeo de interoperabilidad** [Fuente: isa\\_annex\\_ii\\_eif\\_en.pdf](#)  el cual define estos **cuatro niveles**:

- **Interoperabilidad Jurídica:** Las actuaciones de una y otra organización que se comunican, tienen un **marco legal equivalente** y que da soporte a la comunicación o a la colaboración.
- **Interoperabilidad Organizativa:** Que se puedan establecer objetivos comunes para las organizaciones mediante la utilización de procesos coordinados. P.e. el desarrollo de una base de datos de personas de interés, con criterios comunes para dos organizaciones.
- **Interoperabilidad Semántica:** Vamos a profundizar en ella en esta unidad a través del estudio de sus dos regulaciones más recientes e importantes:
  - *Europea:* El Perfil de aplicación de catálogos de datos DCAT-AP.
  - *Nacional:* La norma técnica de interoperabilidad para la reutilización de recursos de información.
- **Interoperabilidad Técnica:** Que exista una planificación común de las herramientas de gestión de sistemas de información para posibilitar el traspaso de información.

### MARCO EUROPEO

La comisión europea ha impulsado a través del **Programa ISA<sup>1</sup>** la acción 1.1 (*Improving semantic interoperability in European e-Government systems*), que desarrolla **5 tipos de iniciativas**:

- ✓ Desarrollo de vocabularios para administración electrónica (**e-Government Core Vocabularies** [Fuente: https://joinup.ec.europa.eu/node/102288](https://joinup.ec.europa.eu/node/102288) )
- ✓ Esquema de metadatos para la descripción de activos digitales (**Asset Description Metadata Schema** [Fuente: https://joinup.ec.europa.eu/node/145985](https://joinup.ec.europa.eu/node/145985) )
- ✓ La iniciativa de gestión de los estándares de datos (**European Interoperability Reference Architecture EIRA** [Fuente: https://joinup.ec.europa.eu/node/99464](https://joinup.ec.europa.eu/node/99464) )

---

<sup>1</sup> Interoperability Solutions for European Public Administrations

- ✓ Exploración de los datos enlazados abiertos del ámbito público (**Linked Open Government Data** [Fuente: https://joinup.ec.europa.eu/node/145998](https://joinup.ec.europa.eu/node/145998) )
- ✓ Desarrollo y promoción del perfil de aplicación de catálogos de datos, DCAT-AP (**Development & promotion of the DCAT-AP**). Esta iniciativa es de especial relevancia, y en la que profundizamos en esta unidad.

### + MÁS INFORMACIÓN

Además, la Unión Europea ha puesto en marcha una conferencia anual, “**SEMIC Conference**” [Fuente: https://joinup.ec.europa.eu/node/145999](https://joinup.ec.europa.eu/node/145999) con el objetivo de crear un espacio para compartir los avances en la interoperabilidad y la gestión de la información pública. La primera de estas conferencias se celebró en 2011 en Bruselas.

## 1.1 Conceptos y necesidad de interoperabilidad

La **interoperabilidad semántica**<sup>2</sup> es la posibilidad de interacción a nivel de datos entre distintos sistemas, garantizando que el intercambio de información, respeta la coherencia de los contenidos intercambiados.



La interoperabilidad presenta **importantes ventajas**, entre ellas apuntamos las siguientes:

- **Reducción de tiempos y costes al intercambio de información:** El aumento de la interoperabilidad entre entidades (públicas o privadas) acelera las transacciones y reduce los costes; si bien, la necesidad de transformadores de datos puede introducir algún problema de compatibilidad, así como retardos en la ejecución de la transferencia de información.
- **Posibilita la reutilización de datos:** La reutilización de información genera negocios de valor añadido que son extremadamente dependientes del **coste**<sup>3</sup> de aprovisionamiento de información. La falta de interoperabilidad puede suponer unos costes que hagan inviables negocios innovadores.

<sup>2</sup> Por ejemplo, si nos preguntan por una medida en metros respondemos con esa medida en metros, y no en pulgadas, kilómetros, etc.

<sup>3</sup> El coste no se refiere solo a la adquisición del derecho a utilizar la información, que en muchos casos es gratuita, sino al coste de incorporar la información a un sistema común y las necesidades tecnológicas asociadas.

- **Favorece la creación de conocimiento:** La creación de conocimiento en muchos casos surge de la combinación de informaciones dispersas en muchas fuentes, dado que la falta de interoperabilidad puede imposibilitar esa conexión y por tanto su generación.
- **Permite el desarrollo fiable de soluciones técnicas:** El desarrollo de la interoperabilidad elimina incertidumbres de mercado a los desarrolladores de soluciones técnicas ya que conocen de antemano cuáles serán los mecanismos de conexión e intercambio de información y acelera por tanto su creación.

## 1.2 ¿Qué es DCAT?



**DCAT** es una especificación (vocabulario) para la descripción de catálogos de datos en la red, que no necesariamente son de organismos públicos, si bien son los más comunes. La **especificación oficial**<sup>4</sup> [Fuente: https://www.w3.org/TR/vocab-dcat/](https://www.w3.org/TR/vocab-dcat/) ha sido elaborada por el W3C, basada en algunos proyectos anteriores.

Esta especificación podría ser utilizada por los 3 tipos de repositorios de datos más comunes:

- ✓ Los repositorios de datos científicos (incluyendo bibliotecas).
- ✓ Los portales públicos de datos abiertos.
- ✓ Los datos compartidos de las smart cities.

## 2. DCAT-AP

### INTRODUCCIÓN

Antes de entender qué es DCAT-AP, debemos explicar:

#### ¿QUÉ ES UN AP?

Un perfil de aplicación (AP) es una especificación que reutiliza términos de uno o más de estándares base (en nuestro caso será DCAT), añadiendo especificidades por la identificación de elementos obligatorios, recomendables y opcionales para ser usados en una aplicación específica, así como recomendaciones de vocabularios controlados para ser utilizados.

---

<sup>4</sup> Data Catalog Vocabulary (DCAT) es una W3C Recommendation desde Enero de 2014.

## 2.1 ¿Qué es DCAT-AP?

### DEFINICIÓN

El perfil de aplicación DCAT, o DCAT-AP para portales de datos, es una especificación basada en **DCAT**<sup>5</sup> [Fuente: https://www.w3.org/TR/vocab-dcat/](https://www.w3.org/TR/vocab-dcat/) para describir catálogos de conjuntos de datos del sector público (datasets) en Europa. Está orientada por tanto, a uno de los **3 tipos** de portales de datos abiertos de las administraciones públicas.

### ¿CUÁLES SON SUS OBJETIVOS?

DCAT-AP busca resolver la problemática de la conexión de datos entre portales que publican datos en la UE y específicamente sus objetivos son:

- **Busqueda entre fuentes:** Permitir la búsqueda de conjuntos de datos a lo largo de varios portales, independientemente de fronteras o sectores (o incluso sin conocer su ubicación precisa). Por ej. mediante la federación de los portales u otro tipo de mecanismos de puesta en común de recursos.
- **Reutilización:** Fomentar la reutilización efectiva de la información.
- **Neutralidad tecnológica:** Mantener la interoperabilidad entre las distintas fuentes y con una aproximación tecnológica neutral.

## 2.2 Orígenes y estado actual del proyecto DCAT-AP

La especificación de la DCAT-AP fue una iniciativa conjunta de la **DG CONNECT** [Fuente: https://ec.europa.eu/digital-single-market/en/dg-connect](https://ec.europa.eu/digital-single-market/en/dg-connect), la Oficina de Publicaciones de la UE y el **Programa ISA**<sup>6</sup>.

La especificación fue elaborada por un grupo de trabajo multidisciplinar con representantes de 16 Estados miembros de Europa, incluyendo:

- ✓ Expertos del portal nacional de datos abiertos de España
- ✓ Algunas instituciones europeas
- ✓ EE.UU.

La primera versión data de 2013, y la **versión actual 1.1** [Fuente: https://joinup.ec.europa.eu/asset/dcat\\_application\\_profile/asset\\_release/dcat-ap-v11](https://joinup.ec.europa.eu/asset/dcat_application_profile/asset_release/dcat-ap-v11) fue publicada en octubre de 2015.

<sup>5</sup> DCAT es una especificación para la interoperabilidad de catálogos de datos desarrollada por el W3C.

<sup>6</sup> Interoperability Solutions for Administrations: programa de la comisión desde el año 2010.

### + MÁS INFORMACIÓN

La nueva versión de DCAT-AP introduce la clarificación de algunos términos, favorece la interconexión con otras normativas, mejora del multilinguaje y la posibilidad de generalización a más catálogos, entre otras.

En noviembre de 2015, el Portal paneuropeo de datos, que contiene más de un cuarto de millón de conjuntos de datos, fue lanzado cumpliendo este nuevo estándar. **Otros lo han adoptado también** [Fuente:https://joinup.ec.europa.eu/node/145996#Who%20is%20using%20the%20DCAT-AP?](https://joinup.ec.europa.eu/node/145996#Who%20is%20using%20the%20DCAT-AP?)

### PROYECTOS RELACIONADOS

Para que la especificación tenga el mayor alcance posible se han lanzado varios procesos de normalización en ámbitos específicos:

#### GEOGRÁFICO (GEODCAT-AP)

Se trata de una extensión de DCAT-AP para describir los conjuntos de datos, las series y los servicios geoespaciales. Busca la posibilidad de que los datos ya codificados de acuerdo a las normativas ISO 1911X, dirigidos por la directiva **INSPIRE** [Fuente: 13\\_Directiva%20INSPIRE.pdf](#) , sean fácilmente publicados en portales de datos. Este **Subgrupo de trabajo** [Fuente: https://joinup.ec.europa.eu/asset/dcat\\_application\\_profile/asset\\_release/geodcat-ap-v10](https://joinup.ec.europa.eu/asset/dcat_application_profile/asset_release/geodcat-ap-v10), creado dentro de los grupos de trabajo de la norma DCAT-AP, ha publicado ya una **primera versión** [Fuente: 13\\_geodcat-ap\\_version\\_1.0.pdf](#) .

Esta información facilitará la publicación directa de datos geográficos en portales de datos abiertos en formato semántico, y no sustituye a las regulaciones previas de estándares geográficos.

#### INFORMACIÓN ESTADÍSTICA (STATDCAT-AP)

[Fuente: https://joinup.ec.europa.eu/asset/stat\\_dcat\\_application\\_profile/home](https://joinup.ec.europa.eu/asset/stat_dcat_application_profile/home)

El objetivo es ofrecer especificaciones y herramientas que mejoran la interoperabilidad entre las descripciones de los conjuntos de datos estadísticos y entre los datos estadísticos y los portales de datos abiertos. Hasta el momento el grupo ha elaborado un borrador de trabajo, cuyas actualizaciones se podrán conocer a través de su **lista pública de correo** [Fuente: http://joinup.ec.europa.eu/mailman/listinfo/stat\\_dcat\\_application\\_profile](http://joinup.ec.europa.eu/mailman/listinfo/stat_dcat_application_profile).

#### OTROS PROYECTOS

Para facilitar la adopción de DCAT-AP, se están desarrollando otros proyectos como:

- ✓ **Compatibilidad con CKAN** [Fuente: http://ckan.org/](http://ckan.org/): gestor de repositorios de datos abiertos open source desarrollado por **Open Knowledge** [Fuente: https://okfn.org/](https://okfn.org/). Más información sobre ello, y sobre el conector en **github** [Fuente:](#)

<https://github.com/ckan/ckanext-dcat> en la Unidad relativa a diseño de APIs y Linked Data.

- ✓ Desarrollo de un **validador**: Disponible en el portal de la Comisión y para comprobación online hay disponibles:
  - [joinup.ec.europa.eu](http://joinup.ec.europa.eu) Fuente: [https://joinup.ec.europa.eu/software/dcat-ap\\_validator/release/all](https://joinup.ec.europa.eu/software/dcat-ap_validator/release/all)
  - [validator.dcat-editor.com](http://validator.dcat-editor.com) Fuente: <https://validator.dcat-editor.com/>
- ✓ El desarrollo de un **editor completo** Fuente: <https://dcat-editor.com/manager/#view=start> de DCAT-AP (en beta):

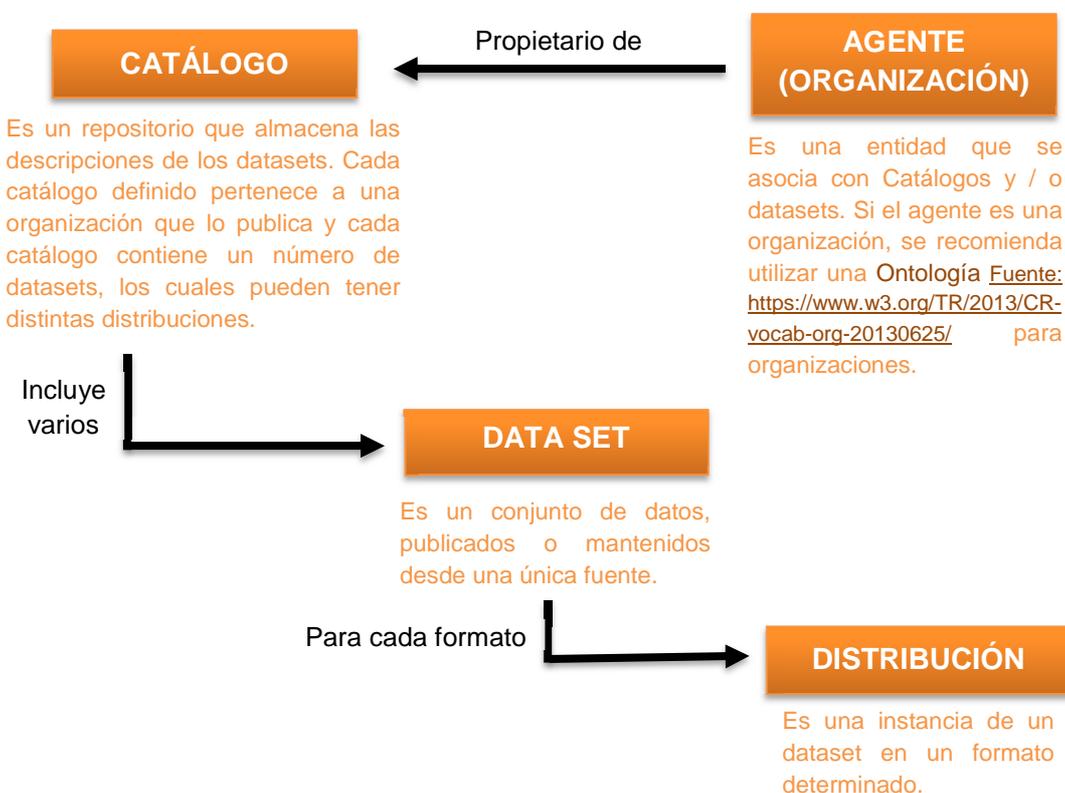
### + MÁS INFORMACIÓN

Otro recurso de interés es una **librería JavaScript** Fuente: <http://rdforms.com/editors/dcat/> para la edición de formatos RDF compatibles con DCAT (*Esta librería es para DCAT y no específica de DCAT-AP*).

## 2.3 La especificación DCAT-AP

### 2.3.1 CONCEPTOS PRINCIPALES

La especificación necesita de una serie de **definiciones** que son utilizadas a lo largo de su descripción, son las siguientes:



- **Categoría:** Es una etiqueta que identifica el tema del dataset.
- **Esquema de categorías:** Es un grupo de conceptos (p.e. un vocabulario controlado) en el que están definidas las categorías.
- **Licencia:** Un documento legal que describe los permisos para la utilización de un recurso (catálogo o distribución).
- **Periodo de tiempo:** Un intervalo de tiempo que es identificado por sus fechas de inicio y fin.

### + MÁS INFORMACIÓN

Se pueden encontrar otros términos específicos en el capítulo 3 de la propia especificación 1.1 de DCAT-AP.

### 2.3.2 VOCABULARIOS CONTROLADOS

Los **vocabularios**<sup>7</sup> permiten describir algunos de los elementos incluidos en la especificación DCAT-AP. Un vocabulario controlado, como los requeridos inicialmente en DCAT-AP, es un conjunto de términos referidos a una materia, y las reglas para su uso en la indexación y recuperación.

Para que sean aceptables como vocabulario en la especificación, deben cumplir una serie de características:

- ✓ Que tengan licencia abierta. **Más información**<sup>8</sup>  
Fuente: <http://opendefinition.org/guide/>
- ✓ Que sean mantenidas por una institución de la UE, un organismo de normalización u otra organización de similar confianza.
- ✓ Que esté documentado adecuadamente (suficientemente).
- ✓ Que permita etiquetar en varios idiomas, idealmente en todos los de la UE.
- ✓ Que tenga un número limitado de términos (10-25).
- ✓ Que tenga términos que estén identificados por URI que lleven a su documentación.
- ✓ Que tengan políticas de permanencia y versionado.



<sup>7</sup> Conjunto de términos y sus relaciones sobre un concepto.

<sup>8</sup> Puedes conocer más sobre licencias abiertas en: <http://opendefinition.org/guide/>

## ELEMENTOS DE VOCABULARIO

La especificación DCAT-AP define una serie de **propiedades**<sup>9</sup> (algunas obligatorias, otras recomendables y otras opcionales) cuyos valores deben pertenecer a **vocabularios controlados**<sup>10</sup> que se identifican a continuación.

Se incluyen aquí los **elementos mínimos**<sup>11</sup> para la puesta en marcha de un catálogo, incluyendo únicamente los elementos totalmente obligatorios:

- **Clase**<sup>12</sup>: Propiedad - **URI**<sup>13</sup>/ Literal – **Cardinalidad**<sup>14</sup> - Notas
- **Agente (formato foaf: agent)** Fuente: [http://xmlns.com/foaf/spec/#term\\_Agent](http://xmlns.com/foaf/spec/#term_Agent): Nombre – Literal - 1..n - Uno por idioma
- **Catálogo (formato: dcat:catalog)**:
  - Título – Literal - 1..n - Uno por idioma
  - Descripción – Literal - 1..n - Uno por idioma
  - Editor - URI (foaf:Agent) – 1 - La organización que publica el catálogo
  - Dataset - URI (dcat:dataset) - 1..n - Uno por cada dataset que componga el catálogo
- **Dataset (formato dcat:dataset)**:
  - **Título**<sup>15</sup> – Literal - 1..n - Uno por idioma
  - **Descripción**<sup>16</sup> – Literal - 1..n - Uno por idioma

## REQUISITOS

Hay una serie de requisitos que deben cumplir tanto los publicadores como los utilizadores de catálogos conforme a DCAT-AP:

- **Para publicadores:**

Para que una propuesta de catálogo sea conforme a DCAT-AP 1.1 tiene que cumplir:

---

<sup>9</sup> Estas propiedades son análogas a los campos de una base de datos.

<sup>10</sup> La lista exhaustiva de vocabularios controlados que componen la norma puede obtenerse en su punto 5.2.

<sup>11</sup> Este es el mínimo de información para que podamos construir un catálogo cumpliendo DCAT-AP. Ni siquiera estaría proporcionando URL para su descarga, y de hecho la existencia de distribuciones de un dataset no es obligatoria (aunque sin ellas tiene poco sentido un catálogo).

<sup>12</sup> Se refiere a cada uno de los elementos que componen el catálogo.

<sup>13</sup> Si la propiedad es una URI o es un literal interpretable directamente por humanos.

<sup>14</sup> Número de instancias (copias) que se pueden incluir de la propiedad en la clase.

<sup>15</sup> Aunque tenga el mismo identificador, este título se refiere al dataset y no al anterior referido al catálogo.

<sup>16</sup> Aunque tenga el mismo identificador, no confundir con el indicador del catálogo.

- Debe tener una descripción.
- Proporcionar información en todos los aspectos definidos como **obligatorios**<sup>17</sup>, para los registros de datasets.
- Una descripción de los datasets que componen el catálogo, incluyendo un título y una descripción de cada uno.
- Una URL para cada distribución de un dataset (en caso de existir).
- Un nombre de los agentes que publican el catálogo o los dataset.
- El título de un esquema de categorías que contenga todas las categorías asignadas a los datasets.
- El título de las categorías asignadas a los datasets.
- Seguir los vocabularios controlados listados en el perfil de aplicación.
- **Para utilizadores:**
  - Deben ser capaces de procesar (aceptar los datos e incorporarlos) la información de las clases obligatorias, recomendadas y opcionales, así como de las propiedades asignables a dichas clases.
  - Debe ser capaz de utilizar de los vocabularios controlados definidos.

### 2.3.3 OTROS ASPECTOS

Hay otros aspectos que deben ser tenidos en cuenta en el catálogo:

- **Accesibilidad:** Dependerá de las herramientas que visualicen los datos dependiendo de los campos tipo de medio (media-type) y formato (format) asociadas a cada distribución de un dataset.
- **Multi-idioma:** En muchos casos los literales pueden ser repetidos en los idiomas necesarios siempre que se incluya la etiqueta idioma conforme a norma **BCP47** Fuente: <https://tools.ietf.org/html/bcp47> y sus transformaciones por la norma **RFC6497** Fuente: <https://tools.ietf.org/html/rfc6497> de IETF.

### ¿CÓMO PODEMOS SABER SI UN PORTAL DE DATOS UTILIZA DCAT-AP?

La primera prueba para comprobar si un portal de datos utiliza DCAT-AP es comprobar el número de **categorías** Fuente: <http://publications.europa.eu/mdr/resource/authority/data-theme/html/data-theme-eng.html> en las que clasifica la información, y si cumplen el anexo II de la norma.

---

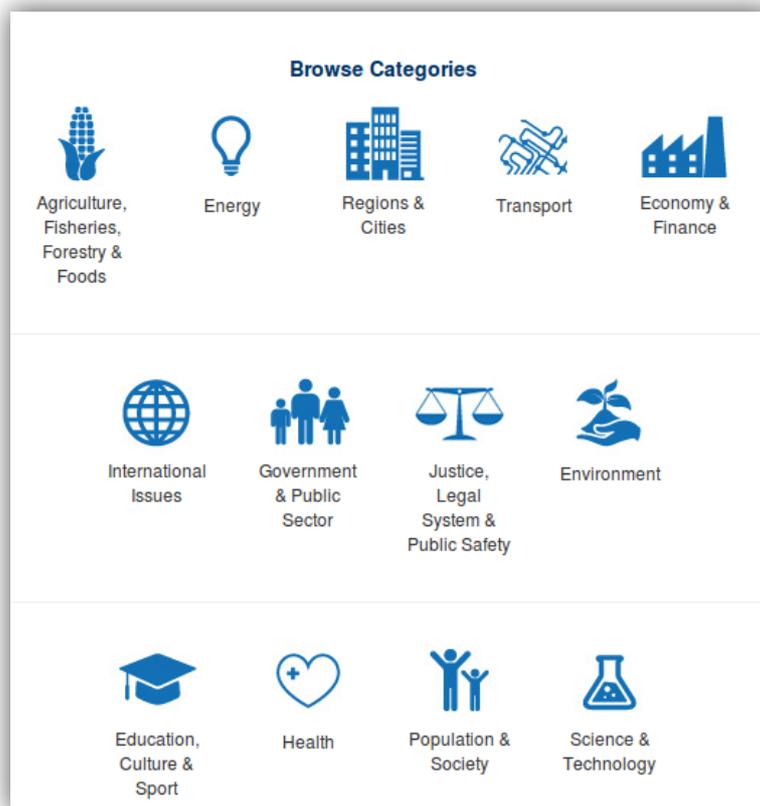
<sup>17</sup> Ver sección 4.2.1 de la norma.

## 2.4 Caso de uso: Portal paneuropeo

El Portal paneuropeo [Fuente: http://www.europeandataportal.eu/](http://www.europeandataportal.eu/) de datos agrega y hace disponibles más de un cuarto de millón de conjuntos de datos provenientes de 33 países europeos (UE + Suiza, Noruega, Moldavia, Serbia, Liechtenstein) y en 21 idiomas diferentes.

Actualmente incluye más de **60 catálogos de datos** provenientes de otras tantas instituciones europeas (de las cuales 3 son españolas). Debe indicarse que uno de estos últimos, datos.gob.es, a su vez está recopilando datos de otros portales españoles.

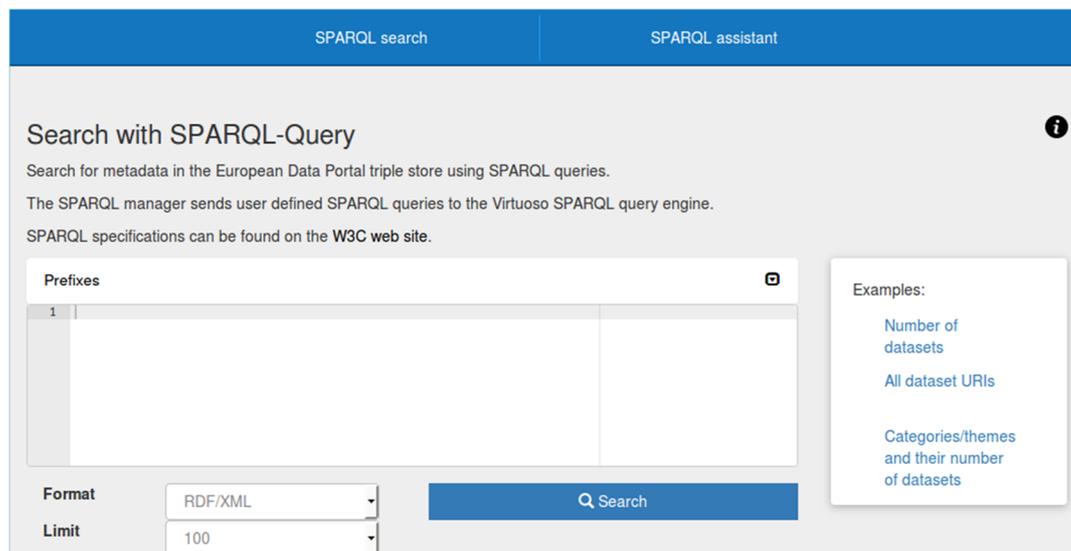
Es la mayor implementación mundial de DCAT-AP y atesora la experiencia de realizar conversiones a DCAT-AP desde multitud de fuentes dispares. En su portada puede verse esta imagen, con la clasificación en 13 categorías.



## + MÁS INFORMACIÓN

El código fuente de la plataforma está disponible para su [descarga Fuente: https://gitlab.com/groups/european-data-portal](#) y utilización bajo la [licencia Apache 2.0 Fuente: http://www.apache.org/licenses/LICENSE-2.0](#), una licencia de código abierto. Cuenta, además, con un punto de consulta [SPARQL Fuente: http://eurostat.linked-statistics.org/sparql](#).

Puedes ampliar información sobre SPARQL en la Unidad relativa a diseño de APIs y Linked Data.



The screenshot shows a web interface for SPARQL search. At the top, there are two tabs: "SPARQL search" (selected) and "SPARQL assistant". Below the tabs, the main heading is "Search with SPARQL-Query" with an information icon. The text below explains that the search is for metadata in the European Data Portal triple store using SPARQL queries, and that the manager sends user-defined queries to the Virtuoso SPARQL query engine. It also mentions that SPARQL specifications can be found on the W3C website.

There is a "Prefixes" table with one row containing the number "1". Below the table, there are two dropdown menus: "Format" set to "RDF/XML" and "Limit" set to "100". A blue "Search" button is located to the right of these dropdowns.

On the right side, there is a box titled "Examples:" containing three links: "Number of datasets", "All dataset URIs", and "Categories/themes and their number of datasets".

## + SABER MÁS

Si quieres ampliar información sobre el valor que pueden crear los portales de datos abiertos (poniendo datos a disposición del público), puedes consultar un informe creado por el proyecto titulado [“Creating value through Open data” Fuente: 19\\_edp\\_creating\\_value\\_through\\_open\\_data\\_0.pdf](#) 

### 3. NTI-RISP

La **Norma Técnica de Interoperabilidad (NTI-RISP<sup>18</sup>)** tiene por objeto establecer las condiciones comunes sobre **selección, identificación, descripción, formato, condiciones de uso y puesta a disposición** de los documentos y recursos de información elaborados o custodiados por el sector público, relativos a numerosos ámbitos de interés como la información social, económica, jurídica, turística, empresarial, educación, etc. Esta norma está enmarcada dentro del **Esquema Nacional de Interoperabilidad (ENI)** [Fuente: 20\\_1\\_NTI%20Norma%20Técnica\\_BOE-A-2013-2380.pdf](#)  que se describe a continuación.

#### 3.1 Introducción al Esquema Nacional de Interoperabilidad

El **Esquema Nacional de Interoperabilidad (ENI)** [Fuente: 20\\_2\\_RD%204\\_BOE-A-2010-1331.pdf](#)  establece la serie de **Normas Técnicas de Interoperabilidad** que son de obligado cumplimiento por las administraciones públicas de España y que desarrollan **aspectos concretos de la interoperabilidad entre las administración públicas y con los ciudadanos**.

Todas las normas y documentos están accesibles en el **Portal de Administración Electrónica (PAe)**:



[Fuente: 20\\_3\\_PAE%20Normas\\_técnicas\\_de\\_interoperabilidad.pdf](#) 

Puedes realizar una primera aproximación al ENI consultando su **guía de adecuación** [Fuente: 20\\_4\\_Guia\\_adequacion\\_al\\_ENI\\_PDF\\_NIPO\\_630-14-238-6.pdf](#) .

#### 3.2 Objetivos de la norma y ámbito de aplicación

##### OBJETIVO

El objetivo es **facilitar y garantizar la reutilización de información pública**, asegurando:

- ✓ La persistencia de la información.
- ✓ El uso de formatos.
- ✓ Los términos y condiciones de uso adecuados.

Si bien el concepto “reutilización” es ampliamente aludido, hay limitados estudios en cuanto a:

---

<sup>18</sup> Norma publicada en Febrero de 2013 por la Secretaría de Estado de Administraciones Públicas de España.

- ✓ Las condiciones que facilitan su desarrollo
- ✓ Los mecanismos para su promoción
- ✓ Las métricas de su impacto.

## ¿CÓMO FOMENTAMOS LA REUTILIZACIÓN?

Gracias a:

1. Aspectos semánticos: relacionados con el modelo de datos.
2. Aspectos técnicos: referidos a los formatos o a los mecanismos de acceso.
3. Otros aspectos relevantes:
  - ✓ Aspectos legales.
  - ✓ Georeferenciación de la información.
  - ✓ Características de tiempo real.

El análisis de la reutilización de la información que publiquemos en nuestro portal de datos, nos ayudará a determinar las políticas de promoción y a seleccionar la información para su publicación.

### 3.3 Preceptos de la norma

#### 3.3.1 SELECCIÓN DE LA INFORMACIÓN

Para seleccionar la **información reutilizable** debemos cumplir los siguientes preceptos:

- ✓ Debemos **priorizar la información**<sup>19</sup> [Fuente: 22 ISA%20Programme%20-%202014%20-%20Report%20on%20high-value%20datasets%20from%20EU%20institutions.pdf](#) de mayor potencial.
- ✓ Incluir la información con el **mayor detalle posible**, evitando agregaciones.
- ✓ Utilizar **recursos primarios**, evitando alteraciones que puedan provocar errores en la información.
- ✓ En la medida de lo posible, **enriquecer semánticamente** la información.
- ✓ **Actualización** a las últimas versiones, indicando la fecha de la última actualización.

---

<sup>19</sup> La norma no proporciona guías sobre cómo determinarlo. Este informe del programa ISA esclarece este punto.

Existen, además, otras serie de requerimientos que la norma establece para las informaciones:

- **Identificación de la información:** Debe realizarse por medio de URI http únicas según el patrón descrito en el Anexo II de la norma siendo en la medida de lo posible neutro desde el punto de vista técnico.
- **Descripción de la información:** En la medida de lo posible se deben **enriquecer semánticamente**. Para consultar las ontologías y vocabularios disponibles se puede visitar el **CISE** Fuente: [http://administracionelectronica.gob.es/pae/Home/pae/Estrategias/pae/Interoperabilidad/Inicio/pae/Centro/Interoperabilidad\\_semantica.html](http://administracionelectronica.gob.es/pae/Home/pae/Estrategias/pae/Interoperabilidad/Inicio/pae/Centro/Interoperabilidad_semantica.html) o Centro de interoperabilidad semántica del Ministerio de Hacienda y Administraciones Públicas. Deben utilizarse los metadatos descritos en el anexo III.
- **Formatos:** En general deben utilizarse **estándares abiertos**, salvo las excepciones recogidas en el Esquema Nacional de Interoperabilidad. Además, se recomienda que sean en **varios formatos/distribuciones**, para facilitar su reutilización.
- **Condiciones de uso:** Las condiciones de accesibilidad a la información están basadas en la ley 37/2007 sobre reutilización de la información del sector público y el real decreto 1495/2011 que la desarrolla.
- **Puesta a disposición:** En principio cada organismo debe ubicar en la dirección **/datosabiertos**<sup>20</sup> de su sede electrónica un **catálogo** con los recursos que dispone susceptibles de reutilización. Dichas condiciones están reguladas por la ley 11/2007 de acceso electrónico de los ciudadanos a los Servicios Públicos y artículo 8 del decreto 1495 (2011)

### 3.3.2 CATÁLOGO DE INFORMACIÓN PÚBLICA

La NTI-RISP **obliga** a los organismos a:

1. Una interfaz de publicación donde poner a disposición los **metadatos**<sup>21</sup> de sus documentos y los **recursos de información**<sup>22</sup> reutilizables.
2. Tener un interfaz de consulta utilizable por aplicaciones de terceros.
3. La inclusión de al menos los **metadatos obligatorios**<sup>23</sup>.
4. Acceso al catálogo en HTML para personas, y al catálogo en **formato automatizable**<sup>24</sup>.

<sup>20</sup> Por ejemplo, en una dirección del tipo [our.domain.com/datosabiertos](http://our.domain.com/datosabiertos), donde “our.domain.com” debe ser sustituido por el dominio real del organismo.

<sup>21</sup> La normativa no evita que algunas informaciones sean de pago (aunque recomienda el coste marginal), pero los metadatos deben ser accesibles.

<sup>22</sup> Es decir, los contenidos apuntados por los metadatos en terminología semántica.

<sup>23</sup> Ver anexo III del Data Catalog Vocabulary (DCAT) W3C Recommendation 16 January 2014.

<sup>24</sup> Normalmente como RDF de acuerdo a las especificaciones de DCAT.

### 3.4 Esquema de URI

A continuación se describe el esquema de URI a implementar bajo NTI-RISP, donde los contenidos entre corchetes pueden ser incluidos o no

**http:// {Base} / {carácter} [/ {Sector} ]/[ {Dominio} ]/[ {concepto} ][. {ext} ]**

- **{Base}**  
our.domain.com el dominio básico
- **{carácter}**  
Puede tomar los valores Catálogo, Def, Kos o Recurso
  - ✓ Catálogo: es la relación de recursos del mismo dominio,
  - ✓ Def: se refiere a vocabulario u ontología, codificado en RDFs o en OWL.
  - ✓ Kos: taxonomía o tesoro, normalmente codificado en SKOS.
  - ✓ Recurso: exactamente eso.
 Tienes más información en el Anexo II de la NTI.
- **{Sector}**  
Para determinar el área de conocimiento que representa conforme a la taxonomía de **sectores primarios**<sup>25</sup> definida en la Anexo VI de la propia norma.
- **{Dominio}**  
Identificar los **elementos específicos**<sup>26</sup> dentro de un sector.
- **{concepto}**  
Representa al recurso o término específico dependiendo del carácter del recurso.
- **.{ext}**  
**Extensión o formato**<sup>27</sup>. Para recursos concretos se omitirá la extensión, pero para el resto la extensión informa del formato técnico del elemento apuntado (p.e n3, .rdf, .html).

Adicionalmente se puede apuntar a un elemento de una colección indicada por un recurso con un identificador tras el carácter #.

<sup>25</sup> Cada uno de los 22 sectores definidos en la norma. Son parcialmente 'mapeables' con las 13 categorías definidas en DCAT-AP (como se ha visto en el punto 2.3.3).

<sup>26</sup> La codificación de estos elementos es abierta.

<sup>27</sup> Este aspecto es solo parcialmente equivalente con DAT-AP.

Por ejemplo serían equivalentes:

- ✓ <http://our.domain.com/Recurso/nombre-recurso/NNN> o
- ✓ <http://our.domain.com/Recurso/nombre-recurso#NNN>

donde:

*nombre-recurso* debe sustituirse por el nombre específico y *NNN* debe sustituirse por un identificador específico del elemento (p.e. un número).

### 3.5 Metadatos

La NTI-RISP utiliza como base los metadatos definidos en DCAT, y por tanto comparte los metadatos en buena medida con DCAT-AP, como se aprecia en la siguiente tabla. El detalle de qué metadatos están incluidos puede encontrarse en el anexo III de la Norma. A continuación se incluyen únicamente los obligatorios:

#### + NOTA EXPLICATIVA

Se han sombreado aquellas filas describiendo propiedades que son obligatorias en NTI-RISP, y que o bien no son obligatorias en DCAT-AP (pueden existir pero ser opcionales), o no se utilizan.

- **Clase<sup>28</sup>**: Nombre – Literal - 1..n - Uno por idioma
- **Catálogo (formato: dcat:catalog)**
  - Título – Literal - 1..n - Uno por idioma
  - Descripción – Literal - 1..n - Uno por idioma
  - Organismo publicador/Editor - URI (foaf:Agent) 1 - La organización que publica el catálogo
  - Fecha de creación – Literal – 1 - Formato ISO 8601
  - Fecha de actualización – Literal – 1 - Formato ISO 8601
  - Idioma(s) – Literal - 1..n - Formato RFC4646, uno por idioma
  - Temáticas - URI (dcat:themeTaxonomy) - **1..22<sup>29</sup>** - **Basado en SKOS: conceptScheme<sup>30</sup>** Fuente: <https://www.w3.org/TR/skos-reference/#schemes>
  - Página web - URI (foaf:homepage) – 1 - Una URL

<sup>28</sup> Los nombres de las propiedades deben coincidir con las que usa la NTI-RISP.

<sup>29</sup> Es el número de temáticas definidas.

<sup>30</sup> SKOS significa Simple Knowledge Organization System, y es un lenguaje de descripción de recursos (simplificado para descripción de ontologías y vocabularios) conceptScheme es un agregador de conceptos dentro de SKOS.

- Términos de uso (licencia) - URI (dct:license) 1
- Dataset (Documento / recurso información) URI (dcat:dataset) - 1..n - Uno por cada dataset que componga el catálogo
- **Documento/ recurso de información equivalente al dataset (formato dcat:dataset)**
  - Nombre/Título - Literal - 1..n - Uno por idioma
  - Descripción – Literal - 1..n - Uno por idioma
  - Temática - URI (skos:Concept) - 1..n - Según clasificación de 22 temáticas
  - **Organismo que expone y publica los datos**<sup>31</sup> (Editor) - URI (dct:publisher) – 1 - Bajo esquema (foaf:Agent)
  - Distribución - URI (dcat:distribution) - 1..n - Uno por cada distribución (versión por formato, rango, etc, que tenga el dataset)
- **Distribución de documento o recurso de información(dominio: dcat:distribution)**
  - URL de acceso – Literal – 1 - Bajo formato (dcat:accessURL)
  - Formato - URI (dcat:mediaType) – 1 - Describe los mime types bajo dct:MediaTypeOrExtent)

La NTI-RISP establece más propiedades obligatorias que DCAT-AP.

### 3.6 Vocabularios

Los vocabularios base, **necesarios** para la identificación de algunos de los elementos identificados en la norma, son los siguientes:

- ✓ XML Schema
- ✓ Simple Knowledge Organization System (SKOS)
- ✓ **Dataset Catalog (dcat)**<sup>32</sup>
- ✓ **Dublin Core Terms**<sup>33</sup>
- ✓ Dublin Core Elements



<sup>31</sup> Esto posibilita que en un catálogo se expongan datasets de varias organizaciones.

<sup>32</sup> Es la base junto con DCT.

<sup>33</sup> Dublin Core Terms y elementos son denominaciones de codificaciones de metadatos diseñadas por el Dublin Core Metadata Initiative, y que se utilizan bien como campos obligatorios o como voluntarios en algún momento de la norma.

- ✓ **W3C Time Ontology**<sup>34</sup>
- ✓ **Friend Of A Friend (FOAF)**<sup>35</sup>

En este caso, los vocabularios base sí son totalmente comunes con DCAT-AP (con diferencias mínimas en la codificación de los idiomas), si bien su utilización como hemos visto en el punto anterior es más prescriptiva.

---

<sup>34</sup> Definición de marcos temporales (actualización).

<sup>35</sup> Describe a las personas y organizaciones.

---

## RESUMEN

Hemos llegado al final de la unidad, y ahora vamos a repasar los **puntos principales** tratados:

- ✓ La interoperabilidad es la capacidad de dos organizaciones de poder intercambiar información entre ellas. En la UE se ha definido para las administraciones públicas el **marco europeo de interoperabilidad** en cuatro niveles: legal, organizativo, semántico y técnico.
- ✓ La **interoperabilidad semántica** es la posibilidad de interacción a nivel de datos entre distintos sistemas, garantizando que el intercambio de información respeta la coherencia y el significado de los contenidos intercambiados.
- ✓ El **DCAT** es una especificación (vocabulario) para la **descripción de catálogos de datos** en la red, que ha sido elaborada por el W3C.
- ✓ El **perfil de aplicación DCAT** o **DCAT-AP** para portales de datos es una especificación basada en DCAT para **describir catálogos de conjuntos de datos** del sector público en Europa.
- ✓ La primera versión de DCAT-AP es de 2013, y la versión actual 1.1 fue publicada en octubre de 2015. En noviembre de 2015 fue lanzado el **portal paneuropeo de datos**, cumpliendo este nuevo estándar, y contiene más de un cuarto de millón de conjuntos de datos.
- ✓ Los **catálogos** son repositorios que almacenan las descripciones y metadatos de los juegos de datos; en algunas ocasiones, los catálogos pueden ser publicados por organizaciones editoras que los agreguen o bien por ellas mismas.
- ✓ Los **vocabularios** permiten describir algunos de los elementos incluidos en la especificación DCAT-AP. Para que éstos sean aceptables como vocabulario en la especificación deben cumplir una serie de características. La mayoría de los elementos incluidos en la especificación DCAT-AP pertenecen a vocabularios controlados, es decir que tienen que cumplir las características de éstos para ser elementos válidos.