

Aplicación de la norma técnica de interoperabilidad de reutilización de recursos de información en los catálogos de información pública de las comunidades autónomas

Yolanda MARTÍN-GONZÁLEZ

ymargon@usal.es

Paula SAINZ-IGLESIAS

pausainz@usal.es

Universidad de Salamanca. Facultad de Traducción y Documentación

Recibido: Julio 2015

Aceptado: Octubre 2015

Resumen: El propósito de este trabajo es elaborar un estudio sobre la aplicación que se realiza de la Norma Técnica de Interoperabilidad de Reutilización de Recursos de Información en lo relativo a la descripción de los datos abiertos y su correspondiente asignación de metadatos de las Comunidades Autónomas en España. La muestra del estudio la constituyen los doce catálogos de información pública puestos en marcha por dichas administraciones. Basándonos en la observación de las páginas de entrada al catálogo y de las fichas de datos durante, iniciada en octubre de 2014 y finalizada en abril de 2015, se efectúa un análisis de tipo descriptivo-comparativo de la asignación de los metadatos que describen los recursos. Los resultados alcanzados permiten concluir que los catálogos autonómicos cumplen de un modo dispar los requisitos exigidos por la citada norma especialmente en lo que se refiere a la inclusión de los metadatos obligatorios y que existe también heterogeneidad en la designación de los metadatos –obligatorios y adicionales– lo que dificulta en cierta medida la recuperación de la información requerida por el usuario.

Palabras clave: Catálogo; Comunidad Autónoma; Datos abiertos; España; Información pública; Norma Técnica de Interoperabilidad de Reutilización de Recursos de Información.

Application of Standard Technique of Interoperability of Reuse of Information Resources in the catalogues of public information of the autonomous communities

Abstract: The purpose of this paper is to draw up a study on the application that performs the standard technique of interoperability of reusability of information resources with respect to the description of open data of the autonomous communities in Spain. Constituted the sample of the study the twelve catalogues of public information implemented by these administrations. Based on the observation of the entry of catalog pages and data sheets, initiated in October 2014 and finished in April 2015, was a descriptive-comparative analysis of the allocation of the metadata describing resources. The results suggest that autonomic catalogs meet partially the

requirements demanded by the cited standard as regards the inclusion of the required metadata and heterogeneity in the designation of the metadata - additional and mandatory - there is also making it difficult to some extent to the recovery of the information required by the user.

Keywords: Autonomous Community; Catalog; Open data; Public information; Reuse of information resources Interoperability Standard NIF Spain; Spain.

1 INTRODUCCIÓN

El sector público, en el cumplimiento de sus funciones, genera, gestiona y conserva una gran cantidad de información de calidad. Dicha información presenta un importante potencial social y económico y su difusión contribuye a la transparencia en la gestión pública y a la generación de confianza por parte de la ciudadanía. Tal es así que los ciudadanos se convierten en demandantes de información y la Administración en una de las mejores proveedoras de este activo abundante y barato (Mendo y otros, 2013: 1).

Por otra parte, las políticas de datos abiertos surgen ligadas al concepto de administración electrónica –propiciada por el uso de las tecnologías de la información y la comunicación- que impulsa la generación de contenidos y documentos en formato digital que han de ser difundidos y estar accesibles para los ciudadanos. La finalidad de las iniciativas open data consiste en hacer públicos los datos en poder de las administraciones para que puedan ser posteriormente reutilizados con la finalidad de generar nuevos conocimientos, aplicaciones o servicios que en muchos casos, añadirán un valor al dato primario y contribuirán a potenciar la innovación.

Sin embargo, no toda la información generada por el sector público posee las mismas características y despierta el mismo interés de cara a su reutilización. La Organización para la Cooperación y el desarrollo Económico (OCDE), en su informe *Participative Web: User-Created content* (DSTI/ICCP/IE 2) elaborado en 2005, realiza una distinción entre dos grupos de información pública (cit. por Ramos, 2012: 3). Por un lado, la información del sector de las administraciones producida y recopilada por los organismos del sector público como parte de su actividad se considera información dinámica, continua y fácilmente accesible. Esta información, directamente generada por el sector público y asociada al funcionamiento de su actividad (p.e. información meteorológica, datos económicos...) se halla vinculada al concepto de conocimiento como recurso competitivo en la sociedad de la información. Por otra parte, existen otros contenidos públicos más estáticos, conservados, registrados y guardados -más que producidos- en los organismos públicos que no están asociados directamente a la función de gobierno, tampoco tienen un específico valor comercial, aunque se orientan a otros propósitos públicos, por ejemplo, culturales o educativos, es el caso de museos, archivos y bibliotecas.

Por tanto, el movimiento *open data* o de datos abiertos se puede entender como una práctica que persigue que determinados datos estén disponibles de forma libre a todo el mundo, en formatos que permitan que puedan ser utilizados, reutilizados y redistribuidos y sin ningún tipo de restricción legal, financiera o tecnológica y al mismo tiempo, como una forma de comunicación abierta, permanente y bidireccional entre la administración y los ciudadanos (Ferrer-Sapena; Peset; Aleixandre-Benavent, 2011: 261). Para conseguirlo, y tal como señalan Ramos Simón et alii (2012: 34), “desde 2009 hay una clara corriente internacional a favor de la apertura de datos públicos en Internet lo que ha inducido a la promoción, por parte de muchos gobiernos de todo el mundo, de portales para hacer accesibles los datos públicos”.

No obstante, no todos los datos se pueden liberar y los que se liberan no pueden hacerlo de cualquier forma. Los requisitos que debe reunir un grupo de datos para que pueda ser liberado son las siguientes (Open Linked Data: guía para la liberación de datos, 2014: 3):

- Deben ser procesables automáticamente por máquinas (machine readable)
- Deben estar publicados en formatos abiertos que, a poder ser, no dependan de empresas concretas para favorecer así su reutilización (p.e. es preferible publicar en formato CVS en lugar de formato Excel de Microsoft)
- Ha de facilitarse el acceso masivo a los mismos (normalmente “en bruto”)
- Deben tener una “granularidad” suficiente para que resulten útiles al usuario.
- Deben ser fáciles de encontrar en un único punto de acceso en línea, y ser puntualmente publicados (compromiso con los usuarios potenciales) y en estándares abiertos.
- Deben poder ser descargables desde diferentes tipos de terminales y navegadores.

En 2008, se puso en marcha en España el Proyecto Aporta que ha supuesto un importante dinamizador de las políticas de datos abiertos y de reutilización en nuestro país además de servir de modelo a otros países para la puesta en marcha de iniciativas similares (Marcos-Martín y Soriano-Maldonado, 2011: 294). El acceso a la información que custodia la Administración General del Estado español se realiza, desde marzo de 2010, a través del catálogo de información pública puesto a disposición por el gobierno a partir del citado proyecto. El RD 1494/2011, por el que se desarrolla la Ley 37/2007 de Reutilización de la Información del Sector público, otorga -en su artículo 4.1- apoyo normativo a dicho catálogo y establece obligaciones de actualización del mismo por Administraciones y organismos.

Las Comunidades Autónomas, por su parte, promueven la publicación de la información de sus administraciones y organismos mediante catálogos a través de los cuales permiten el acceso y la reutilización de sus datos a la ciudadanía y/o las empresas. Dichos catálogos permiten la consulta a los distintos documentos y

recursos de información en posesión de las instancias públicas autonómicas. La naturaleza de tales recursos puede ser diferente al permitirse el acceso a sitios web y/o a grupos de datos en formatos abiertos que se hallan estructurados y organizados siguiendo criterios basados en la temática del recurso, la institución que lo generó, su fecha o los formatos en los que se presenta, entre otros.

Con el fin de normalizar el tratamiento y la gestión de los recursos de información en el ámbito público, el 19 de febrero de 2013 se publicó la Resolución de la Secretaría de Estado de Administraciones Públicas, por la que se aprobó la *Norma Técnica de Interoperabilidad de Reutilización de Recursos de Información* (en adelante NTIRRI). El objeto de dicha norma es “establecer objetivos 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, sobre empresas, educación, etc. cumpliendo con lo establecido en la Ley 37/2007, de 16 de noviembre, sobre reutilización de la información del sector público”.

En principio, los destinatarios del contenido de la norma serían los siguientes (Guía de aplicación de la NTIRRI, 2013: 12):

- Directivos y máximos responsables en la aprobación de políticas y planes de Reutilización de la Información del Sector Público en los organismos.
- Responsables de organismos que produzcan, gestionen o conserven recursos de información susceptibles de ser reutilizados y, por tanto, de garantizar los requisitos de interoperabilidad con el resto de organismos.
- Responsables de creación, producción, gestión, conservación y uso de documentos electrónicos que contenga información susceptible de ser reutilizada.
- Desarrolladores de aplicaciones de creación y gestión de los catálogos de información y recursos reutilizables, de forma que se asegure la interoperabilidad entre los catálogos.

En la citada guía se hace constar que dicha norma contiene únicamente “una serie de directrices generales y el contenido mínimo para el desarrollo de metodologías específicas por parte de los diferentes órganos de la Administración” por tanto, no constituye una metodología en sí misma para la publicación de datos reutilizables.

Según la NTIRRI, los catálogos de información pública reutilizable deberán equipar a éstos con:

- Una interfaz de publicación, que permita a los diferentes órganos y entidades públicas poner a disposición los metadatos de sus documentos y recursos de información reutilizables.
- Una interfaz de consulta, que permita que las aplicaciones de terceros puedan acceder a funcionalidades de búsqueda.

El contenido de los catálogos es información que corresponde a todos los ámbitos de actuación y temas en los que desarrolla su actividad cualquier administración y/o institución pública entre las que se encuentran los gobiernos nacional, autonómicos y locales.

La descripción del catálogo y de cada categoría de documentos o recursos de información se ha de realizar en fichas donde se recojan, al menos, los metadatos obligatorios, establecidos en el Anexo III de la NTIRRI. En dicho anexo, de cada uno de los metadatos exigidos, se ofrece la siguiente información: término recomendado para la representación del metadato, una breve descripción del mismo, las entidades o propiedades básicas del metadato, si es obligatorio o si admite más de un metadato y el tipo y esquema de valores propuesto.

Figura 1. Descripción de los metadatos en la NTIRRI

Catálogo (dominio dcat:Catalog)						
Metadato	Descripción	propiedad	R	M	Tipo y Esquema de valores	
Nombre	Breve título o nombre dado al catálogo de datos.	dct:title	✓	✓	Literal.	Cadena alfanumérica (se recomienda en varios idiomas).
Descripción	Resumen descriptivo del catálogo de datos.	dct:description	✓	✓	Literal.	Cadena alfanumérica (se recomienda en varios idiomas).
Órgano publicador	Entidad que publica el catálogo.	dct:publisher	✓	—	foaf:Agent.	Se especificará el URI correspondiente a un órgano público diferenciados por un código alfanumérico único para cada órgano/unidad/oficina, que será extraído del Directorio Común gestionado por el MINHAP según el esquema siguiente: http://datos.gob.es/recurso/sector-publico/org/Organismo/{ID- MINHAP}
Tamaño del catálogo	Número total de documentos y recursos de información inventariados en el catálogo.	dct:extent	—	—	dct:SizeOrDuration.	Se recomienda incluir el valor de un número entero y su representación textual equivalente.
Identificador	Referencia para identificar el catálogo.	dct:identifier	—	—	xsd:anyURI.	URI que identifica la descripción actual del catálogo.
Fecha	Fecha de publicación inicial del	dct:issued	✓	—	Literal.	Fecha/Hora con formato ISO-

La norma contempla además la posibilidad de asignar metadatos adicionales que contribuyan a mejorar la calidad de la información ofrecida en el catálogo.

Se confirma, por tanto, la necesidad de impulsar políticas e iniciativas que promuevan la reutilización de la información pública para lo cual se requieren mecanismos de apertura de los datos y los recursos informativos en poder de las administraciones central y autonómica. Asimismo se impone la necesidad de potenciar y definir medidas concretas para la creación y el desarrollo de este mercado como fuente de financiación y sostenimiento del sector público y como uno de los motores de la economía española (Arquero Avilés y otros, 2011: 440).

2 OBJETIVOS Y METODOLOGÍA

2.1 OBJETIVOS

El principal propósito de este trabajo es elaborar un estudio sobre la descripción de los datos abiertos de las Comunidades Autónomas en España y, en concreto, sobre la aplicación que, a este respecto, se realiza de la Norma Técnica de Interoperabilidad de Reutilización de Recursos de Información.

A partir de este objetivo general se derivan los siguientes objetivos específicos:

- Dibujar el mapa de catálogos de información pública existentes en España a nivel autonómico.
- Realizar un estudio comparativo de los catálogos de información pública existentes en España y, en concreto, sobre la forma de descripción de sus datasets o grupos de datos.
- Identificar las diferencias existentes en la asignación de los metadatos obligatorios y los metadatos adicionales.

2.2 METODOLOGÍA

2.2.1 REVISIÓN BIBLIOGRÁFICA

El proceso de elaboración de nuestro estudio comenzó con una primera fase de búsqueda de información especializada sobre el tema.

Para ello se recurrió a fuentes especializadas en el ámbito de la información pública y de los datos públicos abiertos con el fin de encontrar trabajos relacionados con el tema. Asimismo, se consultaron numerosas revistas, bases de datos y repositorios donde localizar trabajos científicos relacionados con las tres áreas que confluyen en este trabajo: la ciencia documental la gestión y la administración pública y la ciencia jurídica. Ello nos permitió localizar artículos como el publicado, en 2012, por Ramos Simón y otros autores, *De la reutilización de la información a los portales de datos abiertos en Europa*, donde se analizan,

por primera vez de forma pormenorizada, diferentes portales de datos abiertos en tanto que herramientas difusoras de información.

Se extrajo información especializada en las páginas web de instituciones y fundaciones interesadas en la apertura y la reutilización de datos públicos como es la *Open Knowledge Foundation* y la española *Fundación Centro Tecnológico de la Información y la Comunicación* (CTIC). Ambas organizaciones desarrollan importantes actividades en el mercado de los contenidos digitales y participan en destacados proyectos sobre implementación de políticas open data en el sector público.

Para conseguir una visión más completa del tema, se consultaron los portales públicos oficiales dispuestos tanto en España como en otros países en materia de datos abiertos y de reutilización de información.

En el ámbito europeo se estudiaron los portales open government data del Reino Unido –uno de los países pioneros en la puesta en marcha de este tipo de políticas y que cuenta, por tanto, con una dilatada trayectoria en el ámbito que nos ocupa-, Francia, Italia, Portugal, Bélgica y Países Bajos, entre otros. Se consideró relevante la consulta a catálogos del entorno de la Unión Europea porque los países que los avalan han asumido, al igual que España, en su ordenamiento jurídico nacional la Directiva europea 2003/98 relativa a la reutilización de la información del sector público.

El catálogo de información pública puesto en marcha por el gobierno estatal, datos.gob, constituyó una importante fuente de información para nuestro estudio pues además de ayudarnos a localizar las iniciativas autonómicas existentes sobre la materia también nos aportó información sobre el open data y la reutilización en general. Por último, se accedió a todos los portales autonómicos que desarrollan la iniciativa Open Government Data en España.

La Norma Técnica de Interoperabilidad de Reutilización de Recursos de Información, publicada en el BOE el 4 de marzo de 2013, constituye el documento de referencia en la realización de nuestro estudio. Su localización se llevó a cabo a través de la base de datos de legislación de la Agencia Estatal Boletín Oficial del Estado.

Todo este material recopilado nos sirvió para elaborar un repertorio bibliográfico adecuado al tema tratado en este estudio.

2.2 SELECCIÓN DE LA MUESTRA

El método elegido para realizar esta investigación ha sido el muestreo. El total de los catálogos puestos a disposición por las distintas Comunidades Autónomas conforma la muestra del trabajo. La elaboración de dicho trabajo transcurrió entre el mes de octubre de 2014 y el mes de abril de 2015.

Para la localización de los catálogos se utilizó como fuente principal de información el portal oficial del gobierno español sobre reutilización de la información pública *datos.gob* (<http://datos.gob.es/datos/>) donde se da cuenta de las distintas iniciativas surgidas a nivel español, europeo e internacional (no europeo).

Según la citada fuente, de las diecisiete comunidades autónomas existentes en España doce cuentan con un portal de datos abiertos. Aquellas que no han creado aún este tipo de herramienta son: Cantabria, Extremadura, Madrid, Murcia y Valencia.

Sin embargo, cuando se procedió a acceder a los distintos portales mediante los hiperenlaces facilitados en el portal estatal se constató la imposibilidad de acceder al portal del Principado de Asturias. En el caso del Gobierno de Canarias la dirección ofrecida en el portal datos.gob nos redirigió al portal del Instituto Canario de Estadística que ofrece sus datos en abierto. Sin embargo, decidimos no incluirlo en nuestro estudio puesto que, al contrario de los otros catálogos que constituyen la muestra, este presenta sólo un tipo de información (estadística) y está generada por una sola institución (el Instituto Estadístico).

Con el fin de confirmar la veracidad de los datos recogidos en el portal datos.gob se procedió a realizar búsquedas específicas en los portales oficiales de todas las comunidades autónomas de España. Así se comprobó que en el Portal de Transparencia y Colaboración Ciudadana del Gobierno de Extremadura se contempla el catálogo de datos públicos abiertos de la administración autonómica. Asimismo se localizó el portal Open Data Canarias que permite la consulta y el acceso a sus datos.

Pese a que algunas comunidades autónomas no poseen un portal de datos abiertos en *stricto sensu*, sí han puesto en marcha plataformas en las que explican cuáles son sus avances en la iniciativa Open Government Data tal como ocurre en la Comunidad Valenciana en el que se ha creado un blog para mostrar lo que está haciendo la Autonomía en este ámbito¹. En el caso de Cantabria, no cuenta con ningún portal de datos aunque en la página oficial de la Comunidad Autónoma aparece un apartado dedicado a las descargas de datos abiertos².

Las comunidades autónomas de Madrid y Murcia aún no han puesto en marcha portales de datos abiertos a nivel autonómico aunque sí existen iniciativas a nivel local como es el caso del Ayuntamiento de Madrid³.

Teniendo en cuenta todos estos datos, finalmente la muestra para nuestro estudio la conformaron los doce portales de datos abiertos dispuestos por las administraciones autonómicas para permitir el acceso a sus datos y a sus recursos de información (Tabla 1).

¹ <http://governobertc.wordpress.com/about-2/>

² <http://www.territoriodecantabria.es/Contenido/descargas-datos-abiertos-open-data/136/>

³ <http://datos.madrid.es/portal/site/egob/>

Tabla 1. Portales de datos abiertos de las Comunidades Autónomas en España

Comunidad Autónoma	Nombre del Portal	Dirección url
Andalucía	Datos Abiertos	http://www.juntadeandalucia.es/datosabiertos/portal.html
Aragón	Aragón Open Data	http://opendata.aragon.es/
Canarias	Open Data Canarias	http://opendata.gobiernodecanarias.org
Castilla-La Mancha	Portal de Datos Abiertos JCCM	http://opendata.jccm.es/
Castilla y León	Portal de Datos Abiertos	http://www.datosabiertos.jcyl.es/
Cataluña	Datos Abiertos GenCat	http://dadesobertes.gencat.cat/
Extremadura	Portal de Transparencia	http://gobiernoabierto.gobex.es/datos/
Galicia	Portal Open Data "Abert@s"	http://abertos.xunta.es/portada
Islas Baleares	Dades Obertes CAIB	http://www.caib.es/caibdatafront/index?lang=es
La Rioja	Dato Abierto Rioja	http://www.larioja.org/npRioja/default/index.jsp?idtab=758984
Navarra	Iniciativa Open Data Navarra	http://www.gobiernoabierto.navarra.es
País Vasco	Open Data Euskadi	http://opendata.euskadi.net

Todos estos portales poseen una estructura similar conformada por una serie de elementos básicos y comunes como son:

- El catálogo de datos donde se encuentran localizados todos los datasets y recursos de información con sus correspondientes metadatos y accesibles en distintos formatos.
- El catálogo o repertorio de aplicaciones desarrolladas a partir de la reutilización de la información y los datos públicos.
- El espacio de participación ciudadana que permite a los usuarios participar dando su opinión o realizando comentarios sobre cuáles son los datos que deberían abrirse.

- Información sobre la política o iniciativa de datos abiertos de la administración autonómica correspondiente acompañada de la base jurídica o normativa sobre la que se fundamenta.
- Recursos relacionados con la filosofía Open Data y la reutilización de la información tales como noticias, estudios, eventos, etc.

2.3 RECOGIDA Y ANÁLISIS DE LOS DATOS

Para la recogida de los datos se consideró en un primer momento la utilización de alguna metodología que permitiera la recuperación automatizada de los metadatos. Sin embargo, dada la especial naturaleza de la información del sector público y la todavía escasa normalización existente en lo referente a su apertura se optó finalmente por el método científico de la observación directa.

Dicho método se basó en la consulta sistemática de cada uno de los portales autonómicos que han desarrollado la iniciativa Open Government Data en España y, más concretamente, a sus catálogos de datos.

De cada uno de los catálogos se realizó un estudio general de la información contenida en ellos (p.e. relacionada con el diseño del catálogo, formas de organización de los recursos, etc.) y otro más específico de las etiquetas o metadatos asignados a los distintos recursos con el fin de describirlos.

La información sobre los metadatos que han de constar en la descripción de los datos en los catálogos españoles se extrajo de la Norma Técnica de Interoperabilidad de Reutilización de Recursos de Información.

Para el análisis de los metadatos se estableció una doble clasificación. En primer lugar, se distinguió entre los correspondientes a la página de entrada al catálogo y los relativos al registro del recurso. En segundo lugar, respetamos la distinción establecida por la propia norma entre metadatos obligatorios y metadatos adicionales.

Estos metadatos se han incluido en una tabla junto con cada uno de los catálogos, con el fin de poder comparar y obtener conclusiones acerca de qué catálogos cumplen o no con los requisitos establecidos en la norma. Se ha seguido el mismo procedimiento en el caso de los metadatos adicionales que contenían algunos de estos catálogos.

No obstante, debemos indicar que han existido ciertos problemas a la hora de la recogida de los datos, especialmente en lo que se refiere a la descripción de los datasets pues la asignación de las etiquetas, en un mismo catálogo, varía. Por esta razón, para el análisis de los metadatos presentes en la ficha del dataset decidimos realizar hasta diez búsquedas diferentes con el fin de extraer una conclusión final.

3 RESULTADOS

Como ya hemos explicado anteriormente, la Norma Técnica de Interoperabilidad de Reutilización de Recursos de Información⁴ establece una serie de condiciones básicas para la reutilización de los documentos y recursos de información elaborados y custodiados por las instancias y organismos públicos. Aunque dichas pautas afectan a los distintos procesos de gestión de la información de carácter público (selección, identificación, descripción, formato, condiciones de uso y difusión), para nuestro estudio tomamos como referencia lo establecido para la descripción de los grupos de datos y que aparece estipulado en el apartado V de la norma, *Descripción de la información reutilizable*. Asimismo, y puesto que constituyen el objeto principal de nuestro trabajo se ha tenido en cuenta el apartado IX, *Catálogo de información pública reutilizable*.

En las siguientes líneas se exponen los resultados alcanzados de la realización del estudio, organizados en dos bloques: resultados relativos a la página de entrada del catálogo y resultados extraídos del análisis de la ficha del dataset o recurso de información.

En cada uno de los bloques se ofrece, además de la correspondiente tabla comparativa con los resultados extraídos del estudio concreto de cada apartado, una breve descripción del/os término/s utilizado/s para representar cada metadato –obligatorio y adicional- y un comentario sobre su presencia en cada uno de los catálogos analizados.

3.1 PÁGINA DE ENTRADA AL CATÁLOGO

3.1.1 METADATOS OBLIGATORIOS

Con el fin de obtener una idea general de lo que va a describirse a continuación, se muestra, en la siguiente tabla, la relación de metadatos considerados obligatorios en la norma y su presencia en la página de entrada de los catálogos analizados.

⁴ <http://datos.gob.es/saber-mas?q=node/2672>

Tabla 2. Metadatos obligatorios para la página de entrada del catálogo exigidos por la NTRIRI

<i>Metadato</i>	Andalucía	Aragón	Canarias	CLM	CyL	Cataluña	Extrem.	Galicia	Islas Baleares	La Rioja	Navarra	P. Vasco
<i>Nombre</i>	X	X	X	X	X	X	X	X	X	X	X	X
<i>Descripción</i>				X								
<i>Órgano publicador</i>	X	X	X	X	X	X	X	X	X	X	X	X
<i>Tamaño catálogo</i>	X	X	X				X	X				
<i>Identificador</i>	X	X	X	X	X	X	X	X	X	X	X	X
<i>Fecha de creación</i>												
<i>Fecha de actualización</i>												
<i>Idioma(s)</i>						X		X	X		X	X
<i>Cobertura geográfica</i>												
<i>Temática(s)</i>		X	X	X		X	X	X	X		X	X
<i>Página web</i>	X	X	X	X	X	X	X	X	X	X	X	X
<i>Términos de uso</i>				X	X	X					X	X
<i>Doc y recursos de información</i>	X	X	X				X	X	X		X	

El primer metadato asignado al catálogo es el metadato *Nombre* que aporta información sobre el título dado al catálogo de datos y lo identifica por lo que todos los catálogos consultados incluyen este dato.

La norma recomienda que, para facilitar el acceso a los recursos, este metadato se proporcione en varios idiomas tal como hace el catálogo del gobierno navarro que ofrece esta información en castellano, euskera, francés e inglés; los catálogos del País Vasco y de Cataluña lo hacen, además de en las lenguas oficiales de su Comunidad, en inglés; el catálogo de Galicia, en gallego y en español; el catálogo balear sólo permite la consulta en español y en balear.

El metadato *Descripción* incluye un resumen explicativo del contenido del catálogo de datos lo que nos permite tener una idea aproximada sobre la información que ofrece. Al igual que en el caso anterior, la norma también recomienda que este metadato se proporcione en varios idiomas. Sin embargo, esta etiqueta únicamente se contempla en el catálogo de Castilla-La Mancha que ofrece una breve descripción sobre cómo consultar la información y qué recursos o datos pueden encontrarse en él.

El metadato *Órgano publicador* ofrece información sobre la entidad que publica el catálogo de modo que el usuario pueda saber quién es el responsable de la información. Para ello la norma exige que se especifique el URI correspondiente

a un órgano/unidad/oficina público identificado éste mediante un código alfanumérico único.

El *órgano publicador* es uno de los pocos metadatos que aparecen en todos los catálogos que conforman la muestra de nuestro estudio y responde a los distintos órganos e instituciones de la administración autonómica aunque se incluyen también en algunos casos recursos procedentes de la administración pública estatal. La diferencia entre unos catálogos y otros radica en la terminología utilizada para designar este metadato. Así, el catálogo de Cataluña y el de Galicia lo denominan *Fuente de los datos* y el de Castilla la Mancha, *Órgano editor*.

El número total de documentos y de recursos de información inventariados en el catálogo se muestra bajo la denominación *Tamaño del catálogo*.

Este es un metadato que aporta información cuantitativa sobre el contenido del catálogo pero que se encuentra poco utilizada en los catálogos autonómicos ya que sólo Andalucía, Aragón, Canarias, Extremadura y Galicia proporcionan esta información.

En el caso concreto del catálogo de la Comunidad Autónoma de Cataluña esta información va asociada al metadato *Órgano Publicador* donde se ofrece un listado con el número de recursos accesibles en función del órgano del que provienen.

El metadato *identificador* ofrece la referencia para identificar el catálogo, es decir, el URI asociado a él. Por ejemplo, el identificador del catálogo de Extremadura es <http://gobiernoabierto.gobex.es/datos/>. Todos los catálogos cuentan como es lógico con este metadato.

La fecha de publicación inicial del catálogo constituye un elemento informativo que, según la Norma de Interoperabilidad ha de constar en el metadato *Fecha de creación* y que ha de redactarse utilizando el siguiente formato: Fecha/Hora con formato ISO-8601: YYYY-MM-DDThh:mm:ssTZD.

Curiosamente ninguno de los catálogos analizados contemple este metadato lo que impide al usuario conocer la fecha de su puesta en marcha.

Fecha de actualización es el término que se utiliza para mostrar la fecha en la que se modificó por última vez el catálogo. Para su redacción debe utilizarse un formato similar al del metadato *fecha de creación*: Fecha/Hora con formato ISO-8601: YYYY-MM-DDThh:mm:ssTZD.

Como ocurría en el caso anterior, la fecha de actualización no se facilita en ninguno de los catálogos analizados lo que a nuestro entender constituye un error puesto que, en el contexto de la reutilización de la información el grado de actualización de ésta y, por tanto, de las herramientas que la facilitan es muy importante. No podemos olvidar en este sentido que el sector infomediario está reclamando cada vez con mayor insistencia información en tiempo real para crear productos y servicios más competitivos.

El metadato *Idioma* muestra la/s lengua/s en la/s que se proporciona la información del catálogo. Este dato debe aparecer consignado mediante valores normalizados de etiquetas para identificar idiomas definidos en el RFC 4646 que es el código de referencia de la etiqueta de idioma que explica cuál es la estructura

y la construcción semántica que tienen que seguir: “es”: español; “ga”: gallego; “ca”: catalán; “eu”: euskera; “en”: inglés; “fr”: francés.

Muestran este metadato los catálogos de las Comunidades Autónomas que cuentan con varias lenguas oficiales como son Cataluña, Islas Baleares, Galicia, Navarra y País Vasco. Sólo en este último caso se utiliza el código de referencia propuesto por la Norma para ofrecer esta información.

El ámbito geográfico cubierto por el catálogo se especifica en el metadato *Cobertura geográfica* sin embargo ninguno de los catálogos examinados lo incluye. Aunque se dé por sobreentendido que, en el caso de los catálogos de las Comunidades Autónomas, esta herramienta cubre este ámbito territorial sería aconsejable que se indicara puesto que el catálogo puede contener documentos producto de las relaciones institucionales que mantenga una Comunidad con otra, con el gobierno central o con otras administraciones.

Cada uno de los catálogos debe mostrar, bajo la etiqueta *temática/s*, las distintas materias que cubre. En algunos casos, la clasificación temática es elaborada por los propios responsables del catálogo tal como ocurre con los catálogos de Aragón, Canarias, Extremadura, País Vasco, Galicia, Islas Baleares, Navarra y Cataluña apareciendo en este último caso el metadato bajo la denominación *Categoría*. En otros casos, se toma como referencia la clasificación propuesta en el Anexo IV de la NTIRRI para la categorización de los recursos de información. Esta clasificación solo es respetada en su totalidad en el catálogo de datos de Aragón. El catálogo de Castilla-La Mancha contempla la mayor parte de las categorías temáticas propuestas por la norma excepto “Seguridad” y “Vivienda”. Los demás catálogos contemplan sólo algunas temáticas. En ocasiones ocurre que se incluyen dos categorías bajo un mismo nombre o dividen una temática en varias como por ejemplo, ocurre en el País Vasco donde la categoría “Medio ambiente” se sub-divide en “Medio ambiente” y, en “Meteorología”.

El metadato *Página web* ha de especificarse mediante la dirección web o la url que permita el acceso al catálogo de datos. Se proporciona esta información en todos los catálogos examinados para el estudio.

Los *Términos de uso* aparecen, bajo el metadato del mismo nombre, en los catálogos de Cataluña, Castilla-La Mancha, Castilla y León, Navarra y País Vasco. Los catálogos de Castilla y León, Canarias y Navarra redirigen además la información de este metadato a otra pantalla en el que se explican cuáles son esos requisitos que han de cumplirse.

Bajo el término *Documento(s) y recurso(s) de información* se facilita acceso a un listado de cada uno de los documentos y recursos de información del catálogo. Todos los catálogos consultados ofrecen este metadato a excepción de los catálogos de País Vasco, Castilla La Mancha, Cataluña y La Rioja. El catálogo de Castilla y León redirige a otra página donde se encuentra esta información.

3.1.2 METADATOS ADICIONALES

Además de los metadatos considerados obligatorios, la Norma Técnica de Interoperabilidad de Reutilización de Recursos de Información contempla la posibilidad de asignar metadatos adicionales con el fin de mejorar la calidad de la información ofrecida.

Por lo que se refiere a la página de entrada al catálogo la norma recomienda la asignación de dos metadatos: número de accesos al catálogo y formato.

En la tabla incluida a continuación se muestra la asignación de metadatos adicionales en la página de entrada al catálogo.

Tabla 3. Metadatos adicionales para la página de entrada al catálogo

Metadato	Andalucía	Aragón	Canarias	C La M	CyL	Cataluña	Extremadura	Galicia	Baleares	La Rioja	Navarra	P. Vasco
<i>N. acceso</i>		x					x					
<i>Formato</i>			x		x	x	x	x	x	x	x	x

El metadato *Número de accesos* muestra el número de veces que los usuarios han accedido a un recurso de información específico. Este metadato se encuentra presente en el catálogo de Aragón donde los recursos a los que se ha accedido más de diez veces en los últimos catorce días aparecen señalados con un asterisco. De modo que, cuando se realiza una búsqueda, el catálogo devuelve en primer lugar los registros o los datos señalados con este símbolo tipográfico. Este metadato también se encuentra en el catálogo de Extremadura aunque, en este caso, aparece bajo el término “visitas”.

Exceptuando los casos de Andalucía, Aragón y Castilla La Mancha, el resto de catálogos de información públicos autonómicos ofrecen el metadato *Formatos* que proporciona información sobre los tipos de formatos en que se encuentran disponibles los datasets y los recursos y, además, permite su búsqueda en función de este criterio.

3.2 REGISTRO DEL DOCUMENTO Y DEL RECURSO DE INFORMACIÓN

3.2.1 METADATOS OBLIGATORIOS

En la Norma Técnica de Interoperabilidad de Reutilización de Recursos de Información, también en su anexo III, se contemplan los metadatos mínimos asociados

a los documentos y recursos de información y que pasamos a describir a continuación indicando si se contemplan o no en los catálogos de datos públicos autonómicos.

Se muestran en la siguiente tabla dichos metadatos y su presencia en los registros de los datos y de los recursos de información –también denominadas *Fichas Técnicas*- de los catálogos analizados.

Tabla 4. Metadatos obligatorios para la ficha del dataset exigidos por la NTIRRI

<i>Metadato</i>	Andalucía	Aragón	Canarias	CLM	CyL	Cataluña	Extrem.	Galicia	Islas Baleares	La Rioja	Navarra	P. Vasco
<i>Nombre</i>	X	X	X	X	X	X	X	X	X	X	X	X
<i>Descripción</i>	X	X	X	X	X	X	X	X	X	X	X	X
<i>Temática</i>		X	X	X		X				X	X	X
<i>Etiquetas</i>		X	X		X		X			X	X	X
<i>Identificador</i>	X	X	X	X	X	X	X	X	X	X	X	X
<i>Fecha de creación</i>	X		X	X	X	X	X	X	X		X	X
<i>Fecha de últ. actualización</i>	X		X	X			X	X	X	X	X	X
<i>Frecuencia de actualiz.</i>		X	X		X	X	X	X	X	X	X	X
<i>Idioma(s)</i>		X	X			X						
<i>Organismo</i>	X	X	X	X	X	X	X	X	X		X	X
<i>Condiciones de uso</i>	X	X	X	X	X	X	X	X	X	X	X	X
<i>Cobertura geográfica</i>	X	X	X	X	X	X				X	X	
<i>Cobertura temporal</i>	X	X	X									X
<i>Vigencia del recurso</i>												
<i>Recursos relacionados</i>					X	X		X			X	X
<i>Normativa</i>												
<i>Distribución</i>	X	X	X	X	X	X	X	X	X	X	X	X

El metadato *Nombre* indica el título o designación del dataset o recurso de información correspondiente y se encuentra presente en todos los catálogos que conforman la muestra aunque en algunos casos aparece bajo la denominación *Título*.

La descripción del dataset y de su contenido ha de realizarse, según la norma, de una manera detallada bajo la etiqueta *Descripción*.

En el caso de ambos metadatos, imprescindibles para una óptima recuperación del dataset, la NTIRRI recomienda además que la consiguiente información se proporcione en varios idiomas lo que únicamente ocurre en el caso de las

comunidades autónomas con lenguas cooficiales y siempre y cuando el usuario modifique el idioma de consulta al catálogo.

El tema principal del documento o recurso de información se consigna bajo el término *temática/s*. Es recomendable, según la Norma, que se haga referencia a un tema asociado del sector público según la taxonomía definida en el anexo IV del texto jurídico. En dicho anexo se establecen las siguientes categorías temáticas:

Tabla 5. Taxonomías propuestas por la NTIRRI

Ciencia y tecnología <i>Incluye: Innovación, Investigación, I+D+i, Telecomunicaciones, Internet y Sociedad de la Información</i>	Energía <i>Incluye: Fuentes renovables</i>
Comercio <i>Incluye: Consumo</i>	Hacienda <i>Incluye: Impuestos</i>
Cultura y ocio <i>Incluye: Tiempo libre</i>	Industria <i>Incluye: Minería</i>
Demografía <i>Incluye: Inmigración y Emigración, Familia, Mujeres, Infancia, Mayores, Padrón</i>	Legislación y justicia <i>Incluye: Registros</i>
Deporte <i>Incluye: Instalaciones deportivas, Federaciones, Competiciones</i>	Medio ambiente <i>Incluye: Meteorología, Geografía, Conservación fauna y flora</i>
Economía <i>Incluye: Deuda, Moneda y Banca y finanzas</i>	Medio Rural <i>Incluye: Agricultura, Ganadería, Pesca y Silvicultura</i>
Educación <i>Incluye: Formación</i>	Salud <i>Incluye: Sanidad</i>
Empleo <i>Incluye: Trabajo, Mercado laboral</i>	

Las clasificaciones temáticas utilizadas en los catálogos suelen ser de carácter general y contemplan categorías tales como economía, salud, estadísticas, administraciones públicas y urbanismo, entre otras. Además, en la mayoría de las ocasiones, se asocian como máximo tres categorías a un solo recurso o dataset.

No todos los catálogos consultados incluyen este metadato lo que puede dificultar el proceso de recuperación del dataset dado que uno de los criterios de búsqueda documental más habituales es precisamente la materia. Aquellos catálogos que sí contemplan este metadato lo hacen utilizando distintas denominaciones, por ejemplo, el catálogo de Aragón utiliza el término *Categoría* y el de Navarra, la denominación *Tema*.

Se utiliza el metadato *Etiqueta(s)* para añadir las etiquetas textuales que permiten clasificar libremente el documento o recurso de información. Se contempla este metadato en siete de los doce catálogos autonómicos. Puede

aparecer también bajo otras denominaciones como *Palabras clave* tal como ocurre en el caso del catálogo de Aragón, o utilizando el término inglés *Tag* como hace el catálogo de La Rioja.

En el metadato *identificador* se añade el URI del documento o recurso de información. Esta información aparece en todos los registros de los catálogos analizados puesto que es el elemento que permite identificar y localizar de forma unívoca en la web el dataset o el recurso de que se trate.

Se utiliza la etiqueta *Fecha de creación* para indicar la fecha de elaboración del documento o del recurso de información. La Norma Técnica insta a que se siga el formato siguiente: Fecha/Hora con formato ISO-8601: YYYY-MM-DDThh:mm:ssTZD.

A excepción de los catálogos de Aragón y de La Rioja, todos los demás catálogos examinados ofrecen esta información aunque sólo siguen el formato establecido por la Norma el catálogo de las Islas Baleares y el de Castilla-La Mancha. El resto de catálogos señalan la fecha pero no la hora.

En el caso del catálogo de Castilla-La Mancha establece además dos tipos de fecha de creación, una para la definición y otra para los datos.

El metadato *Fecha de última actualización* indica la última fecha conocida en la que se modificó o se actualizó el recurso. Para redactar tal información la Norma Técnica establece la siguiente forma: Fecha/Hora con formato ISO-8601: YYYY-MM-DDThh:mm:ssTZD.

Tal como ocurría en el caso de la página de entrada al catálogo, este metadato resulta relevante especialmente para el sector infomediario que se sirve de los datos alojados en estos catálogos para crear nuevos servicios y aplicaciones y que solicita sean lo más actuales posibles. Conscientes de ello, todas las comunidades autónomas proporcionan esta información a excepción de los catálogos de Cataluña, Aragón y Castilla y León que no lo hacen. Por otra parte, y tal como ocurría con el metadato anterior, solo el catálogo de las Islas Baleares cumple totalmente con el esquema que establece la NTIRRI. Además, este catálogo utiliza el término *Última ejecución* para referirse a este aspecto.

Bajo la denominación *Frecuencia de actualización* se muestra, de un modo aproximado, el periodo de tiempo transcurrido entre las distintas actualizaciones del recurso. La NTIRRI recomienda especificar periodos normalizados con formato ISO-8601 (P<date>T<time>).

Este metadato aparece en todos los catálogos analizados excepto en los catálogos de Andalucía y de Castilla La Mancha. En los catálogos balear y gallego se contempla en el registro del dataset este metadato aunque vacío de contenido, “sin especificar”.

La frecuencia de actualización, aspecto fundamental en el contexto de los datos abiertos, varía notablemente de unos catálogos a otros tal como se observa en la siguiente tabla:

Tabla 6. Frecuencia de actualización de los catálogos que incluyen el metadato
Frecuencia de actualización

Comunidad Autónoma	Catálogo	Periodicidad
Canarias	Open Data Canarias	Diaria
Castilla La-Mancha	Transparencia-Buen Gobierno	Anual
Cataluña	Datos abiertos gencat	Anual
Extremadura	Gobex	Sin especificar
Castilla y León	Portal de datos abiertos	Variable
Galicia	Abert@s	Sin periodicidad
Islas Baleares	Dades obertes caib	Sin especificar
La Rioja	Dato Abierto Rioja	Mensual
Navarra	Iniciativa Open Data Navarra	Diaria
País Vasco	Open Data Euskadi	Semanal

El idioma o idiomas en el que se encuentra el documento o recurso debe especificarse con su etiqueta correspondiente y según lo dispuesto en el RFC 4646 (“es”, “ga”, “ca”, “eu”, “en”, “fr”). Únicamente los catálogos de Aragón, Canarias y Cataluña facilitan esta información.

La etiqueta *Organismo que expone y publica los datos* proporciona información sobre la entidad que difunde el recurso lo que constituye, sin duda, un dato necesario para que el usuario conozca la procedencia de cada recurso. Se especifica el URI correspondiente al organismo público diferenciado por un código alfanumérico único para cada órgano/unidad/oficina que se extrae del Directorio Común gestionado por el Ministerio de Hacienda y Administraciones Públicas.

El catálogo de La Rioja es el único de los catálogos examinados que no ofrece este metadato. Los demás catálogos emplean numerosos términos para definirlo así, por ejemplo, el catálogo de Aragón utiliza el término *Publicador* y el catálogo de Navarra utiliza la denominación *Editor*.

Las condiciones de uso o licencia específica que se aplica al documento o recurso de información se indican también bajo el metadato correspondiente en este apartado mediante el URI del recurso.

Todos los catálogos asocian este metadato a sus recursos utilizando simultáneamente las denominaciones *Condiciones de uso* o *Licencia*. En este caso, el conjunto de catálogos utiliza la licencia Creative Commons (CC) que es el modelo más habitual y comúnmente aceptado en el ámbito de los datos abiertos con la excepción de los catálogos de Cataluña y las Islas Baleares que no cuentan con ninguna licencia definida.

Excepto el catálogo de Andalucía que utiliza la versión Creative Commons 4.0, el resto utilizan la misma versión de licencia, Creative Commons 3.0.

El metadato *Cobertura geográfica*, que hace referencia al territorio cubierto por el recurso, no se contempla en los catálogos de Extremadura, Galicia, Islas Baleares y País Vasco. El catálogo de Castilla-La Mancha denomina este metadato *Ámbito geográfico* y el catálogo de La Rioja, *Ámbito territorial*.

La etiqueta *Cobertura temporal* describe la fecha de inicio y de fin así como la duración del periodo cubierto por el documento o recurso de información.

Solo los catálogos de Andalucía, Aragón, Canarias y País Vasco utilizan este metadato. En el caso del catálogo del País Vasco distingue además entre la “fecha de inicio de datos” y la “fecha de fin de datos”.

La fecha de validez del documento o recurso de información se consigna bajo el metadato *Vigencia del recurso*. La Norma de Interoperabilidad establece la siguiente forma: Fecha/Hora con formato ISO-8601: YYYY-MM-DDThh:mm:ssTZD.

Curiosamente esta etiqueta no aparece en ninguno de los catálogos de información pública autonómicos lo que impide al usuario conocer la validez real del dataset o del recurso de cara a una posible reutilización.

El metadato *Recurso(s) relacionado(s)* está previsto para añadir los enlaces a recursos relacionados con el dataset y completar así la información proporcionada por el documento inicial.

El catálogo de Cataluña, el de Castilla y León, el de Galicia, el de Navarra y el del País Vasco proporcionan este metadato aunque lo hacen utilizando distintos términos. En el caso del País Vasco aparece bajo el nombre *Documentación asociada*; Navarra utiliza la denominación *Fichas relacionadas*; Galicia lo hace bajo el título *También te puede interesar* y Cataluña lo denomina *Enlaces relacionados*.

La Norma contempla la posibilidad de vincular el dataset o el recurso a uno o a varios texto/s legal/es para lo que se sirve de la etiqueta *Normativa*. Únicamente en el catálogo Abert@s (Galicia) -en el apartado *Enlaces*- se consignan las referencias a los textos normativos relacionados con el tema del que trata ese recurso. Por ejemplo, los grupos de datos relativos a las elecciones al Parlamento de Galicia cuentan con enlaces vinculados a los textos jurídicos que fundamentan tal proceso democrático.

Bajo la etiqueta *Distribución(es)* se identifica la referencia al recurso que describa una distribución del documento o recurso de información. Dado que la utilización de este término propuesto por la NTIRRI puede conducir a confusión, la mayor parte de los catálogos utilizan la etiqueta *Formato* para referirse a la forma en la que se muestra el dato o el recurso y que habitualmente puede ser en xml, csv, ods, html, pdf, json o rdf, entre otras.

Esta información resulta a todas luces relevante puesto que del formato en el que se presente el dataset dependerá el grado de dificultad a la hora de recuperarlo. Por esta razón todos los catálogos autonómicos consultados incluyen este metadato en el registro del dataset o del recurso.

3.2.2 METADATOS ADICIONALES

Además de los metadatos exigidos por la Norma, algunas comunidades autónomas incluyen en sus catálogos metadatos adicionales en el registro del documento o del recurso de información.

Tabla 7. Metadatos adicionales para la ficha del documento y/o recurso de información

<i>Metadato</i>	Andalucía	Aragón	Canarias	C L M	CyL	Cataluña	Extrem.	Galicia	Islas Baleares	La Rioja	Navarra	P. Vasco
<i>Diccionario</i>		X		X								
<i>Nivel de detalle</i>		X										
<i>Fecha de creac. defin.</i>				X								
<i>Sistema de actualización</i>				X								
<i>Más información</i>	X				X							
<i>RDF</i>		X				X		X		X		X
<i>Documentación asociados</i>		X	X					X				X
<i>Otros datasets</i>					X	X		X			X	
<i>Valoración del acceso</i>			X				X		X			
<i>Tipo de dato</i>		X								X		X
<i>Dctc. técnica</i>				X								

Diccionario de datos es un metadato observado en los catálogos de Aragón y de Castilla-La Mancha. En el primer caso, este metadato ofrece una explicación de los elementos informativos que se ofrecen en el recurso o dataset. En el segundo caso, se ofrece este metadato adicional dentro de un apartado denominado “Descripción de los datos”. Después de realizar varias búsquedas comprobamos que este metadato se haya vacío de contenido ya que nos redirige nuevamente a la ficha técnica o registro del dataset elegido en cada ocasión.

El catálogo de Aragón ofrece también la etiqueta *Nivel de detalle* que muestra una serie de términos contemplados en el texto del recurso a modo de descriptores.

El catálogo de Castilla-La Mancha incluye el metadato *Fecha de creación de la definición*. Como ya se ha mencionado al hablar del metadato obligatorio “fecha de creación”, este catálogo distingue entre la fecha de creación de los datos y la fecha de creación de la definición sin embargo, después de realizar varias búsquedas confirmamos que ambas fechas son siempre coincidentes. La única diferencia es que en la fecha de creación de la definición se especifica la hora.

El metadato *Sistema de actualización de datos* está incluido en el catálogo de Castilla-La Mancha. Éste explica la forma en que se actualizan los datos que puede ser de dos tipos: por anexión (añadiendo nuevos datos) o por sustitución (reemplazando los datos). Esta información resulta relevante si queremos saber cómo se renuevan los dataset.

La Junta de Castilla y León y la de Andalucía incluyen además la etiqueta *Más información* donde ofrecen información adicional a la proporcionada en la ficha del dataset como por ejemplo, la publicación en el boletín oficial autonómico correspondiente de la información contenida en el recurso lo que para algunos usuarios puede constituir una información bastante útil.

Los catálogos de Cataluña, Aragón, País Vasco, Galicia y La Rioja ofrecen la posibilidad de acceder el dataset en formato *RDF*, esto es, en un modelo flexible que describe los recursos web utilizando expresiones con la forma sujeto-predicado-objeto.

Los catálogos de Aragón, Canarias, Galicia y País Vasco ofrecen información complementaria al dataset a través de la etiqueta *Documentación asociada o complementaria al dataset* que viene a ampliar la información ofrecida por éste. Los documentos complementarios que añade cada uno de estos catálogos suele contener información sobre el diseño o la metodología seguida en la elaboración de los datos o, si se trata de un dataset con información estadística, suministra el cuestionario realizado.

Algunos catálogos ofrecen sugerencias sobre *otros datasets* relacionados con el recurso en cuestión como son el catálogo de Galicia que lo hace bajo la denominación *También te podía interesar* y el catálogo de Navarra bajo la expresión *Fichas relacionadas*. Ambos catálogos ofrecen un apartado en el que muestran resultados con características relacionadas. También proporcionan esta información los catálogos de Castilla y León, Cataluña y Extremadura. En este último caso se informa sobre el número de visitas o consultas que ha recibido el recurso.

Los catálogos baleares, canario y extremeño ofrecen una *valoración del nivel de acceso* mediante el esquema de cinco estrellas de Tim Berners-Lee.

Los catálogos de Aragón, Euskadi y La Rioja incluyen el metadato *Tipo de dato* en el que se muestra la clase de información (estadística, económica, estudios, informes, etc.) que contiene cada ficha del dataset.

El catálogo de Castilla La Mancha ofrece además el metadato *Documentación técnica* que permite acceder a documentos normativos y/o técnicos relacionados con el recurso o dataset.

4 CONCLUSIONES

Una vez finalizado nuestro estudio sobre la descripción de los datos abiertos en los catálogos de las Comunidades Autónomas en España, se pretende dar respuesta a los objetivos planteados inicialmente al tiempo que ofrecer propuestas de mejora.

En términos generales, el estudio nos permite concluir que existe disparidad en cuanto al cumplimiento de los requisitos exigidos por la Norma, en especial, en lo referente a la asunción de los denominados *metadatos obligatorios* y a la terminología utilizada para designarlos.

Por lo que respecta a la página de entrada al catálogo, y a la asignación de metadatos obligatorios se observa lo siguiente:

- El metadato *Nombre* únicamente se ofrece traducido a otro idioma en los casos de las comunidades autónomas con varias lenguas oficiales cuando la norma considera este requisito como una recomendación de aplicación a todos los catálogos con independencia de su origen geográfico.
- El metadato *Descripción* se obvia en prácticamente todos los casos lo que impide a priori conocer al usuario cuál es el contenido del catálogo y el tipo de información que ofrece.
- La denominación de algunos metadatos difiere en casos tales como la temática y el órgano responsable del documento, entre otros.
- En ninguno de los casos se ofrece información sobre la fecha de creación y de actualización del catálogo lo que, en nuestra opinión resulta erróneo ya que es un dato esencial para los reutilizadores responsables de crear nuevos productos a partir de los datos abiertos.
- La cobertura geográfica creemos que los responsables la dan por sobreentendida ya que curiosamente no se contempla en ninguno de los catálogos estudiados.

Los catálogos correspondientes a Navarra y Galicia son los que cumplen en mayor medida las exigencias de la norma a la hora de asignar metadatos obligatorios en su página de entrada (13/17 metadatos lo que supone el 61,5%). En el lado contrario se sitúan los catálogos de La Rioja y Castilla y León (30,7% y 38,4% respectivamente).

Por lo que se refiere a los metadatos adicionales, el número incluido en las páginas de entrada de los catálogos autonómicos es muy reducido (número de acceso y formato). A excepción del catálogo de Extremadura que contempla ambos, el resto solo ofrece información sobre los formatos en los que se puede acceder a los recursos ya que se trata de información que los usuarios de los datos abiertos y, en especial, los reutilizadores necesitan conocer.

En cuanto a la asignación de metadatos en la ficha o registro de los recursos se observan ciertas discrepancias entre los catálogos de información pública autonómicos.

Catálogos como el de Canarias incluyen en sus registros hasta catorce de los diecisiete metadatos obligatorios requeridos por la Norma. En esta categoría, le siguen con trece los catálogos de Navarra y el País Vasco. Curiosamente, estos tres catálogos excluyen del registro de sus datasets el metadato *vigencia del recurso* lo que nos llama la atención puesto que para la reutilización de los datos esta

información se considera relevante. De hecho, cada vez más las empresas infomediarias solicitan que se faciliten los datos en tiempo real con el fin de mejorar los productos o servicios que crean a partir de la apertura de los datos. Otro metadato que no se encuentra contemplado en ninguno de los catálogos señalados es el que hace referencia a la normativa a la que responde o que afecta al recurso de información.

El catálogo balear es el que presenta mayores deficiencias en la aplicación de metadatos obligatorios en la ficha del recurso o documento al asignar únicamente nueve de los diecisiete metadatos recomendados y obviando información esencial como por ejemplo, la categoría temática en la que se ubica el recurso.

En lo referente a los metadatos adicionales, a diferencia de lo que ocurría en el caso de la página de entrada al catálogo, estos son más numerosos y hacen referencia a aspectos complementarios del dataset. No obstante, en algunos casos, resulta complicado saber qué información se debe consignar bajo algunas etiquetas como por ejemplo en el metadato *diccionario* o en el relativo a *nivel de detalle* pues la norma nos es muy clara a este respecto.

El catálogo que ofrece una información adicional más completa es el de Aragón pese a que considera únicamente la mitad de los metadatos de este tipo propuestos por la Norma. Andalucía, Extremadura, Baleares y Navarra se encuentran en peor situación ya que sólo proporcionan un metadato adicional en su ficha del recurso por lo que queda un tanto incompleta la información que ofrecen.

Teniendo en cuenta estos datos y con el fin de mejorar las prestaciones que ofrecen los catálogos de información pública autonómicos, creemos que sería recomendable que los responsables de los mismos cumplieran de un modo más estricto lo estipulado en la Norma Técnica de Interoperabilidad de Reutilización de los Recursos de Información. Es imprescindible que, en pos de una óptima recuperación de la información, se contemplen todos y cada uno de los metadatos obligatorios establecidos por la norma. Sería también aconsejable que se llegara a un acuerdo sobre la denominación utilizada para designar los metadatos puesto que la utilización de diferentes términos puede conducir, en algunos casos, a confusión tal como ocurre en el caso de “tema” y “categoría”. Asimismo, deberían normalizarse los términos utilizados para clasificar temáticamente los recursos. Puesto que la Norma propone ya su propia taxonomía para la ordenación sistemática de los sectores primarios, los catálogos autonómicos deberían ajustarse a lo estipulado a este respecto en el texto normativo.

También sería recomendable que se considerara la inclusión, en el grupo de metadatos obligatorios, del metadato adicional “tipo de dato” puesto que dada la amplia gama de datos existentes (económicos, estadísticos, jurídicos, etc.) podría facilitar la recuperación de los mismos. Por otra parte, sería necesario reagrupar algunos metadatos adicionales correspondiente a la ficha del registro ya que no son demasiado explícitos y aportan información similar. Tal es el caso de los metadatos siguientes: “Documentación asociada o complementaria”, “Otros

datasets” y “Más información” en los que puede consignarse la misma información relativa al recurso o documento.

Por último, sería aconsejable evitar la inclusión de metadatos en las fichas de los recursos que, sin embargo, se encuentran vacíos de contenido lo que provoca cierta frustración en el usuario del catálogo.

Todo ello contribuiría, sin duda, a una verdadera normalización de los recursos alojados en los catálogos de información pública y por ende a una recuperación más efectiva de los datos propiedad de las administraciones autonómicas españolas.

5 REFERENCIAS BIBLIOGRÁFICAS

- ARQUERO AVILÉS, Rosario et al. (2011) “Reutilización de la información generada en el sector público español”. *Revista Española de Documentación Científica*, nº 3, vol. 34, pp. 427-446.
- ESPAÑA. “Norma Técnica de Interoperabilidad de Reutilización de Recursos de Información”. Boletín Oficial del Estado, 4 de marzo de 2013, núm. 54.
- FERRER SAPENA, Antonia; PESET, Fernanda; ALEIXANDRE-BENAVET, Rafael (2011). “Acceso a los datos públicos y su reutilización: Open data y Open Government”. *El profesional de la información*, nº. 3, vol. 20, pp. 260-269.
- MARCOS MARTÍN, Carlos; SORIANO MALDONADO, Salvador Luis (2011). “Reutilización de la información del sector público y open data en el contexto español y europeo. Proyecto Aporta”. *El profesional de la información*, nº 3, vol. 20, pp. 291-297.
- MENDO, C. et al. (2013). “Del acceso a la reutilización, del dato al documento: una visión conceptual de la información pública”. *Revista Española de Documentación Científica*, nº 3, vol. 36, pp. 1-12.
- MINISTERIO DE HACIENDA Y ADMINISTRACIONES (2013). “Guía de aplicación del Real Decreto 1495/2011 por el que se desarrolla la Ley 37/2007 sobre Reutilización de la Información del Sector Público”. <<http://administracionelectronica.gob.es/>> [Consulta: 16/12/2014].
- RAMOS SIMÓN, Fernando; ARQUERO AVILÉS, Rosario (2012). “Las bibliotecas, archivos y museos ante la nueva directiva de reutilización de información del sector público”. *VI Congreso Nacional de Bibliotecas Públicas/ Europeaana Conference*. <<http://www.mcu.es/bibliotecas/MC/2012/CongresoBP/comunicaciones.html>> [Consulta: 20/01/2015]
- RAMOS SIMÓN, Luis Fernando et alii (2012). “De la reutilización de información del sector público a los portales de datos abiertos en Europa”. *BiD: textos universitaris de biblioteconomia i documentació*, nº. 29. <<http://bid.ub.edu/29/ramos2.htm>> [Consulta: 09/03/2015].
- XIMDEX. “Linked Open Data. Guía para la publicación de datos abiertos y enlazados” (2014). <http://www.ximdex.com/Guides/OXE_OpenData_Guide_ES.pdf> [Consulta: 6/12/2014]