

La reutilización de la información pública en la Comunidad Autónoma de Madrid

Silvia Cobo Serrano
Andrea Sala Jiménez,
L. Fernando Ramos Simón *

*Artículo recibido:
18 de agosto de 2011.*

*Artículo aceptado:
28 de febrero de 2012.*

RESUMEN

La reutilización de información fue introducida por la directiva europea 2003/98/CE y apoyada por la ley española 37/2007 para dar un enfoque uniforme a los aspectos relativos a la reutilización de datos del sector público. Se pretende caracterizar la información y analizar diversos aspectos de la reutilización de la información en la Comunidad Autónoma de Madrid. Se presenta el Portal de datos públicos como resultado de la investigación, y se recomienda la creación de productos de información, normalización, control, comunicación y difusión para potenciar la reutilización de la información madrileña.

* Los tres autores pertenecen a la Universidad Complutense de Madrid, España. (Silvia: s.cobo@estumail.ucm.es); (Andrea: andrea_ccdoc@hotmail.com); (Fernando: lframoss@pdi.ucm.es).
Con la colaboración del Grupo PUBLIDOC-UCM (Rosario Arquero Avilés, Iuliana Botezán, Rodrigo Sánchez Jiménez, Félix del Valle Gastaminza).

Palabras clave: Reutilización de información, Comunidad Autónoma de Madrid, Información del sector público, Activos de información, Portales, Open Data

ABSTRAC

The Re-use of Public Information in the Autonomous Community of Madrid

Silvia Cobo Serrano, Andrea Sala Jiménez and L. Fernando Ramos Simón

Re-use of information was introduced by European Directive 2003/98/EC and supported by the Spanish Law 37/2007 in order to provide a uniform approach on all aspects related to the re-use of public sector data. The study aims to characterize the information and analyze different aspects of the re-use of information in the Autonomous Community of Madrid. The article presents the public data Web portal as a result of our research and it recommends the creation of information products, standardization, control, communication and dissemination to promote the re-use of information in Madrid.

Keywords: Re-use of Information, Autonomous Community of Madrid, Public Sector Information, Databases, Web Portals, Open Data.

I. INTRODUCCIÓN

La información del sector público es considerada en muchos países de la Unión Europea un motor de desarrollo económico por contribuir a la competitividad y a la creación de empleo. De esta manera, los activos de información (buscadores, bases de datos, etc.) de los organismos públicos permiten el acceso, uso y reutilización de sus datos para generar un nuevo mercado de información.

En España, la reutilización de información fue introducida a través de la Directiva 2003/98/CE para fomentar la economía digital; recopilar, tratar y difundir la información de los organismos del sector público (PSB); establecer normas comunes para la creación de productos, y permitir la comunicación

entre las Administraciones, los ciudadanos y las empresas. Tal y como señala la Ley española 37/2007, se entiende por “reutilización de la información del sector público” el

uso de documentos que obran en poder de las Administraciones y organismos del sector público, por personas físicas o jurídicas, con fines comerciales o no comerciales, siempre que dicho uso no constituya una actividad administrativa pública (BOE núm. 276).

Aunque la Ley ha tenido escasa repercusión, destacan las iniciativas y actividades realizadas a nivel nacional por el Proyecto Aporta así como la apertura de datos públicos desarrollados por algunas Comunidades Autónomas: Cataluña, País Vasco o Asturias. El Ministerio de Industria Turismo y Comercio y el Ministerio de Política Territorial y Administración Pública han trabajado entre los años 2010-2011 en un borrador de Real Decreto para un mejor desarrollo de dicha Ley. Su objetivo es promover la reutilización, centrándose en las obligaciones de los organismos del sector público, fomentar el Catálogo de Información Pública (creado en el marco de las actuaciones del Proyecto Aporta) y, por último, regular ciertos aspectos de la reutilización, como pueden ser los datos personales y la propiedad intelectual, entre otros.¹

Uno de los mejores ejemplos de acceso, uso y reutilización de información pública y digital es la Web de la Comunidad Autónoma de Madrid,² que se considera uno de los recursos de referencia por la calidad y fiabilidad de sus datos digitales. No obstante y tras analizar con detenimiento su espacio virtual, se ha detectado cierta dispersión para buscar y localizar sus activos de información pública. Bajo este contexto, el grupo de investigación PUBLIDOC-UCM presentó el proyecto de investigación

La información del sector público de la Comunidad de Madrid: Descripción y análisis de activos de información para su reutilización (Directiva 2003/98/CE y Ley 37/2007)

- 1 Borrador de proyecto del Real Decreto por el que se desarrolla la Ley 37/2007, del 16 de noviembre, sobre reutilización de la información del sector público. Disponible en: http://www.mityc.es/dgdsi/es-ES/participacion_publica/Documents/BorradorRD_reutilizacion_consultaPublica_V1.pdf, [consulta: 26 abril 2011]. Debido a la disolución de las Cámaras por la convocatoria de elecciones generales (noviembre de 2011) es difícil saber si el R.D. será finalmente publicado.
- 2 La Comunidad Autónoma de Madrid alberga la capital del Estado español y es considerada una de las economías más importantes del país (Primer puesto en el PIB nacional). Se organiza, territorialmente, en 179 municipios y tiene una población estimada en más de 6.500.000 habitantes, según el Instituto Nacional de Estadística.

en el año 2009, se centró en los activos de información pública de la Comunidad Autónoma de Madrid para analizarlos desde la perspectiva de la gestión de la información y de la documentación, decisivas para aprovechar su potencial.

Con el presente trabajo se pretende identificar, analizar y centralizar las diferentes bases de datos en línea y enlaces Web, cuyo resultado se puede constatar en la herramienta de análisis *Portal de Información del Sector Público de la Comunidad Autónoma de Madrid*, que facilita a los reutilizadores (ciudadanos y empresas) el acceso a la información pública.

2. MATERIALES Y MÉTODOS

Para la realización del proyecto se disponía de un plazo temporal de doce meses (de noviembre de 2009 a noviembre de 2010), divididos en cuatro fases de desarrollo:

- *Primera fase* (tiempo de ejecución del trabajo: Noviembre 2009 – Enero 2010). Se recopiló toda la información necesaria para identificar las características generales de los datos públicos de la Comunidad Autónoma de Madrid. Para ello, se consultó el “Portal Web de la Comunidad Autónoma de Madrid” (www.madrid.org) y las páginas Web derivadas del mismo: Presidencia, nueve Consejerías, Entidades Institucionales Adscritas y Órganos Colegiados.
- *Segunda fase* (tiempo de ejecución del trabajo: Febrero 2010- Abril 2010). Con la herramienta de tratamiento y análisis de información Joomla,³ se diseñó, en fase piloto, el “Portal de Información del Sector Público de la Comunidad Autónoma de Madrid” (<http://multi-doc.rediris.es/isp-cam/>) para reunir publicaciones oficiales, bases de datos, capturas de páginas Web, etc. Como resultado, se identificaron más de 70 bases de datos y, aproximadamente, 300 recursos de información.
- *Tercera fase* (tiempo de ejecución del trabajo: Mayo 2010 – Septiembre 2010). Se clasificaron y ordenaron los datos relativos a bases de datos, publicaciones, normativas, noticias y otra información relevante en el “Portal de Información del Sector Público de la Comunidad Autónoma de Madrid”.

3 Joomla es un sistema que permite la gestión de contenidos Web y está programado en .php. Portal Joomla Spanish: <http://www.joomlaspanish.org/>, (consulta: 3 octubre 2011).

- *Cuarta fase* (tiempo de ejecución del trabajo: Octubre 2010 – Noviembre 2010). Se redactó la memoria final del proyecto con los resultados y recomendaciones, que presentan una visión global sobre la estructura de la información del sector público de la Comunidad Autónoma de Madrid. La memoria pretende dar a conocer y aplicar la reutilización en los organismos autónomos madrileños y tener un efecto positivo para el crecimiento económico de la Comunidad.

3. RESULTADOS

3.1. Reutilización de la información en la comunidad autónoma de Madrid

Para poder estudiar el grado de importancia de la reutilización de la información, se han analizado y comparado las formas y canales de distribución de la información, los sectores de la economía madrileña, las iniciativas nacionales de reutilización y la migración de datos.

Las formas y canales de distribución de la información se dividen, principalmente, por medios electrónicos, mediante la página Web de la Comunidad Autónoma de Madrid, el CanalCAMTV y la participación en las principales redes sociales como Facebook, Twitter y YouTube. En segundo lugar, el medio de distribución más utilizado es la publicación impresa: edición de obras de la Administración de la Comunidad Autónoma de Madrid y sus organismos dependientes, el Boletín Oficial de la Comunidad y el catálogo de publicaciones impresas.

La consulta de los informes nacionales europeos (Bélgica, Austria, República Checa, Polonia, Italia y Francia),⁴ relativos a la reutilización de información del sector público, permite contrastar los dominios económicos abordados en cada país e identificar los posibles sectores económicos madrileños. El denominador común de los subdominios de información de estas investigaciones son los registros oficiales de negocios, las patentes, la información de marcas registradas, las bases de datos públicas de licitación y la información de la Cámara de Comercio. En España, hay que añadir licitaciones, agricultura, ganadería e industria pesquera, energía, telecomunicaciones, información financiera y estadísticas económicas e industriales.

En nuestro caso, para determinar los sectores de la economía madrileña con mayor demanda de información pública se estableció un criterio temático de acuerdo con la información, previamente analizada, del Portal de la Comunidad

4 Fuente: Comisión Europea. Sociedad de la Información. Información del Sector Público, disponible en: http://ec.europa.eu/information_society/policy/psi/facilitating_reuse/exclusive_agreements/index_en.htm, [consulta: 26 abril 2011].

Autónoma de Madrid. El resultado obtenido fueron los siguientes sectores: comercio, consumo, economía, estadística, industria, presupuestos, tesorería, tributos y turismo.

Actualmente algunos poderes públicos españoles están desarrollando actividades, a nivel nacional y autonómico, con objeto de promover y poner a disposición de los ciudadanos y empresas la información pública para su reutilización por medios electrónicos y favoreciendo así el desarrollo de la sociedad de la información. En este sentido, debemos señalar los siguientes portales y proyectos:

Tabla I. Iniciativas nacionales y autonómicas de reutilización de la información del sector público en España

Portal	Organismo público responsable	Características
APORTA. Catálogo de Información Pública ⁵	Ministerio de Industria, Turismo y Comercio. Secretaría de Estado de Telecomunicaciones y para la Sociedad de la Información.	<ul style="list-style-type: none"> • Nivel de actuación: Administración General del Estado (AGE). • Promover la publicación, mejorar el acceso y favorecer la reutilización de la información de la AGE. • Conseguir una Administración más transparente, eficiente y cercana. • Crear nuevas oportunidades (sociales y económicas) a partir de la industria de contenidos digitales.
Dades Obertes de la Generalitat de Catalunya	Generalitat de Catalunya	<ul style="list-style-type: none"> • Nivel de actuación: autonómico. • Fomentar el uso y la reutilización de la información procedente de la Administración. • Disponibilidad de datos en diferentes formatos para su fácil reutilización. • Facilitar la ordenación interna de los sistemas de información dentro de la Administración. • Fomentar la interoperabilidad entre los servicios del sector público.
Open Data Euskadi	Gobierno Vasco	<ul style="list-style-type: none"> • Nivel de actuación: autonómico. • Obtener productos derivados de datos por parte de empresas, intermediarios y ciudadanos. • Permitir la reutilización de los datos expuestos para analizar y evaluar la gestión pública. • Facilitar la creación de servicios para la sociedad. • Promover la eficiencia en la documentación y clasificación de datos.
Datos de Asturias	Principado de Asturias	<ul style="list-style-type: none"> • Nivel de actuación: autonómico. • Fomentar el uso y la reutilización de información de la Administración pública para hacer ésta más útil y adecuada a la ciudadanía. • Favorecer la transparencia informativa de la Administración. • tener disponibilidad de datos en diferentes formatos para su fácil reutilización.

5 Creado a partir de la publicación del directorio “Bases de datos de libre acceso difundidas por la Administración General del Estado”, realizado por el grupo PUBLIDOC-UCM.

Por último, la Web de la Comunidad Autónoma de Madrid ha sido sometida a un análisis interno mediante un proceso de migración de datos con el programa mBOT,⁶ diseñado para trabajar y completar la totalidad de enlaces de la URL de origen (www.madrid.org), así como los diferentes niveles de análisis, previamente configurados. mBOT consta de un sistema para la detección de enlaces duplicados y sitios Web fuera de funcionamiento.

Con respecto a los enlaces de carácter interno y con base en cinco niveles de profundidad, podemos manifestar que hay un total de 4.626 páginas Web analizadas bajo el dominio de la Comunidad Autónoma de Madrid. Atendiendo a parámetros de estudio más concretos, se han extraído 35.199 enlaces, 2.143 metadatos, 822 canales de sindicación, 15.319 imágenes, 657 documentos y 68 enlaces de la Web social.

3.2. Características de la información pública de la comunidad autónoma de Madrid

El “Portal Web de la Comunidad Autónoma de Madrid” presenta cinco tipos de formatos para la presentación de la información, y todos ellos alcanzan el mayor porcentaje en el quinto nivel de profundidad según el análisis realizado por el programa mBOT:

Tabla II. Formatos de presentación de la información en la Web de la Comunidad Autónoma de Madrid.

Formato de presentación	Núm. Total de documentos
.doc	132
.pdf	365
.html	5.706
.php	5.706
.asp	265
Total	12.174

Otro punto de interés son los activos con mayor capacidad de información, es decir, bases de datos y recursos con información rica y elaborada. En

6 Desarrollado por el Prof. Dr. Manuel Blázquez Ochando de la Facultad de Ciencias de la Documentación de la Universidad Complutense de Madrid.

Su plataforma/soporte es Apache/PHP/MySQL, y alcanza hasta 10 niveles configurables. Con respecto a la extracción de datos, mBOT tiene una extracción configurable y organizada por campos en bases de datos, y puede analizar enlaces a páginas web, metadatos, canales de sindicación, imágenes, documentos, videos, audios, código fuente y/o texto indexado. Su interfaz es completamente web, y permiten la exportación en formatos .dot, .csv, .rdf...; la creación de archivo estadístico con el total de enlaces analizados, sitios, dominios y páginas por nivel.

general, la información recuperada es textual o de imagen/texto, mientras que los activos que proporcionan datos en hojas de cálculo, imágenes y otros formatos son todavía escasos.

Con respecto a la tipología de información reutilizable, los trabajos especializados en materia de reutilización de información –como son las investigaciones realizadas por Arquero Avilés (2008, p. 33-40), los estudios relativos a las publicaciones periódicas oficiales (Arquero Avilés, Mendo Carmona y Ramos Simón, 2006) y las aportaciones sobre los recursos públicos disponibles en Internet (Ramos Simón, Mendo Carmona y Arquero Avilés, 2006, 2009)– sirven de base conceptual y teórico-metodológica para abordar el tema que nos ocupa en el presente artículo de investigación. En consecuencia y basándonos en la siguiente clasificación, podemos señalar que los datos proporcionados en la Web son, fundamentalmente, informes, datos estadísticos y actos y resoluciones.

Tabla III. Tipología de la información pública.
Grupo de investigación PUBLIDOC-UCM.

Tipología de la información
Actos y Resoluciones
Catálogos
Censos y Directorios
Datos estadísticos
Imágenes
Informes
Mapas y Planos
Normas y Jurisprudencia
Noticias y Actualidad
Otros

Con respecto a los sectores y sub-sectores del mercado de la información pública, los reutilizadores (ciudadanos y empresas) manifiestan interés por las cuestiones sociales, que junto con el sector legal y económico constituyen los mercados más demandados en la Web. Por otro lado, los sub-sectores con mayor potencial de reutilización son la legislación, las licitaciones y las estadísticas, (en este último caso, las estadísticas de ciencia y tecnología, las correspondientes a la Administración Pública y las económicas e industriales).

Tabla IV. Sectores y sub-sectores del mercado de la información. Grupo de investigación PUBLIDOC-UCM.⁷

Sector del Mercado de la información pública	Sub-sector del Mercado de la información pública
Negocios / Información Económica	<ul style="list-style-type: none"> • Información de las Cámaras de Comercio • Registros mercantiles • Patentes y marcas registradas • Licitaciones • Agricultura, ganadería e industria • pesquera • Energía • Telecomunicaciones • Información financiera • Estadísticas económicas e industriales
Información Geográfica	<ul style="list-style-type: none"> • Información de direcciones • Fotos aéreas • Edificaciones • Información catastral • Redes geodésicas • Geología • Datos hidrográficos • Información topográfica
Información Legal	<ul style="list-style-type: none"> • Decisiones de tribunales internacionales y extranjeros • Decisiones de tribunales nacionales y regionales • Legislación nacional (boletines oficiales) • Tratados
Información Meteorológica	<ul style="list-style-type: none"> • Datos climatológicos (incluye modelos) • Previsión meteorológica
Datos Sociales	<ul style="list-style-type: none"> • Estadísticas de empleo • Información y estadísticas sobre salud • Estadísticas de población • Estadísticas de la Administración Pública • Estadísticas sociales • Cultura y estadísticas de producción cultural • Estadísticas de Ciencia y Tecnología • Estadísticas de consumo • Discapacitados y personas dependientes • Estadísticas de educación
Información de Transportes	<ul style="list-style-type: none"> • Información sobre el tráfico de vehículos • Información sobre obras viales • Información sobre transporte público • Estadísticas de registros de vehículos

Por último, cabe destacar que la búsqueda simple predomina en la mayoría de los activos (se caracteriza por presentar una única casilla de búsqueda),

7 PUBLIDOC-UCM. PSI: Identification of potential Exclusive Agreements – Spain, disponible en: <http://multidoc.rediris.es/ppo/acuerdos-exclusivos-isp-publidoc.pdf>, [consulta: 26 abril 2011].

mientras que la consulta avanzada suele ser frecuente en los catálogos (por ejemplo, en las bibliotecas) y en los buscadores de publicaciones.

3.3. Portal de información pública de la comunidad autónoma de Madrid

Dado el potencial de reutilización de información pública de los recursos de la Comunidad Autónoma de Madrid, difundidos mediante el dominio www.madrid.org, el grupo de investigación PUBLIDOC-UCM utilizó una herramienta de tratamiento para análisis y difusión de dicha información. El resultado obtenido es el “Portal de Información del Sector Público de la Comunidad Autónoma de Madrid”, accesible en la dirección <http://multidoc.rediris.es/isp-cam/> y en fase experimental. Este Portal es un claro ejemplo de reutilización —con fines públicos y de investigación— por reunir de forma clara, organizada y sistemática las bases de datos y recursos de información más importantes de la página Web de la Comunidad Autónoma de Madrid.

El objetivo primordial de este instrumento de ensayo es centralizar los activos, facilitar el acceso, fomentar su utilización y, consecuentemente, impulsar la reutilización de la información pública madrileña potenciando una cultura de la misma.

Este Portal, innovador y pionero, se puede considerar el primer mecanismo encargado de fomentar y recopilar la información reutilizable de la Comunidad Autónoma de Madrid, y sirve como base y punto de partida de un posible catálogo que sistematice y describa, de forma individual y normalizada, los productos y servicios informativos de los organismos del sector público de la autonomía objeto de estudio.

Actualmente se han identificado más de 70 bases de datos y, aproximadamente, 300 recursos de información —de los que se facilita su organismo de dependencia y enlace Web— que están organizados en las siguientes categorías:

Tabla V. Categorías del “Portal de Información del Sector Público de la Comunidad Autónoma de Madrid”.

Categorías del Portal de Información del Sector Público
Bases de datos de la CAM
Noticias de la CAM
Normativa de la CAM
Publicaciones de la CAM
Entidades Institucionales Adscritas
Órganos Colegiados

Exceptuando el primer bloque, en el que todas las bases de datos son listadas directamente, las categorías restantes presentan sus recursos de información agrupados en las Consejerías de la Comunidad Autónoma de Madrid.

4. DISCUSIÓN Y CONCLUSIONES

- a) A pesar de la Directiva europea 2003/98/CE, la Ley 37/2007 y de las diferentes iniciativas nacionales y autonómicas relativas a Open Data, la reutilización de la información sigue siendo un tema desconocido, incluso en el ámbito público. Por ello, se recomienda una mayor difusión de la misma mediante actividades (reuniones o eventos sectoriales) y políticas activas de comunicación.
- b) El estudio realizado es extrapolable a otras Comunidades Autónomas españolas, lo que nos permitiría conocer el estado de la cuestión de las mismas para la reutilización de su información, y el mejoramiento de la visibilidad y disponibilidad de los datos.
- c) El “Portal de Información del Sector Público de la Comunidad Autónoma de Madrid” puede considerarse el primer instrumento organizado que permite el acceso a los activos de información para su posterior reutilización. En consecuencia y como continuidad del trabajo iniciado con el Portal, sería aconsejable la creación de una herramienta o producto informativo, por ejemplo un catálogo, en el que se puedan recoger todos los recursos de información de la Comunidad Autónoma de Madrid difundidos en Internet, lo que potenciaría una cultura madrileña de reutilización.

5. REFERENCIAS BIBLIOGRÁFICAS

- Arquero Avilés, R., “Metodología para la localización, identificación y análisis de las bases de datos de la Administración General del Estado”, en PUBLIDOC-UCM. *Bases de datos de libre acceso difundidas por la Administración General del Estado*, Madrid: Dirección General del Libro, Archivos y Bibliotecas, 2008, pp. 33-40, ISBN 978-84-7491-940-0.
- Arquero Avilés, R.; Mendo Carmona, C.; Ramos Simón, L. F., “Publicaciones periódicas oficiales en España: evaluación y características de la producción”, en *Seminario Hispano-Mexicano de investigación en bibliotecología y documentación*, 2006, pp. 445-469.

- Borrador de proyecto de Real Decreto por el que se desarrolla la Ley 37/2007, de 16 de noviembre, sobre reutilización de la información del sector público, http://www.mityc.es/dgdsi/es-ES/participacion_publica/Documents/BorradorRD_reutilizacion_consultaPublica_V1.pdf, [consulta: 26 abril 2011].
- Comisión Europea. Sociedad de la Información. Información del Sector Público, http://ec.europa.eu/information_society/policy/psi/facilitating_reuse/exclusive_agreements/index_en.htm, [consulta: 26 abril 2011].
- Comunidad Autónoma de Madrid. Portal de la Comunidad Autónoma de Madrid, www.madrid.org. [consulta: 26 abril 2011].
- Directiva 2003/98/CE del Parlamento Europeo y del Consejo de 17 de noviembre de 2003 relativa a la reutilización de la información del sector público, http://www.aporta.es/c/document_library/get_file?uuid=824b77a0-b033-4c9d-b8a6-52807ea30648&groupId=10128, [consulta: 26 abril 2011].
- Generalitat de Catalunya. Dades obertes gencat, <http://dadesobertes.gencat.cat/index.html>, [consulta: 4 octubre 2011].
- Gobierno del Principado de Asturias. Datos de Asturias, <http://risp.asturias.es/catalogo/index.html>, [consulta: 4 octubre 2011].
- Gobierno Vasco. Open Data Euskadi, <http://opendata.euskadi.net/w79-home/es/>, [consulta: 4 octubre 2011].
- Ley 37/2007, de 16 de noviembre, sobre reutilización de la información del sector público, http://www.aporta.es/c/document_library/get_file?uuid=28c02272-f4a6-4090-b375-d10b87299e92&groupId=10128, [consulta: 26 abril 2011].
- PUBLIDOC-UCM, *Bases de datos de libre acceso difundidas por la Administración General del Estado*, Madrid: Dirección General del Libro, Archivos y Bibliotecas, 2008, ISBN 978-84-7491-940-0.
- PUBLIDOC-UCM, PSI: Identification of potential Exclusive Agreements – Spain, <http://multidoc.rediris.es/ppo/acuerdos-exclusivos-isp-publidoc.pdf>, [consulta: 26 abril 2011].
- Ramos Simón, L. F.; Mendo Carmona, C.; Arquero Avilés, R., "La producción informativa y documental del Estado: hacia un inventario de los recursos públicos", en *Revista Española de Documentación Científica*, 2009, enero-marzo, v. 32, n. 1, pp. 40-59.
- _____, "Producción editorial de los servicios de publicaciones oficiales: hacia un nuevo entorno", en *Seminario Hispano-Mexicano de investigación en bibliotecología y documentación*, 2006, pp. 431-444.
- Secretaría de Estado de Telecomunicaciones y para la Sociedad de la Información (SETSI), Proyecto APORTA, Catálogo de Información Pública, http://www.aporta.es/web/guest/buscador_de_catalogo, [consulta: 4 octubre 2011].

