

I. MERCADO, INDUSTRIA Y SECTOR INFORMACION

I.1. Reutilización de datos culturales

Antonia Ferrer-Sapena y Fernanda Peset

1 marzo 2012

Ferrer-Sapena, Antonia; Peset, Fernanda. "Reutilización de datos culturales".
Anuario ThinkEPI, 2012, v. 6, pp. 193-196.



Resumen: La reutilización de datos en la administración pública española está creciendo de manera notable gracias al *Plan Avanza 2*. No se da el mismo ritmo en el caso de los datos de las instituciones culturales. Se describen los tres aspectos clave que retrasan la liberación de los datos culturales: generar métricas que midan el negocio generado a terceros al abrir los datos; posibilidad de perder la atribución; y la pérdida de potenciales ingresos.

Palabras clave: Reutilización de datos, Instituciones culturales, Metadatos, *Europeana*, Datos culturales, Datos abiertos, *Open data*, *Open linked data*.

Title: Reuse of cultural data

Abstract: The reuse of data by public administration in Spain is growing significantly thanks to *Plan Avanza 2*. But the growth rate is not the same in the case of cultural institutions. The three key issues that delay the release of cultural data are described: the need for adequate metrics for measuring the business generated when data become open, the possibility of losing the attribution, and potential income loss. The most emblematic cultural data release projects are described, and the problems they have faced are identified.

Keywords: Data reuse, Culture institutions, Metadata, *Europeana*, Cultural data, Open data, Open linked data.

La reutilización de información del sector público es una apuesta de la *Unión Europea* desde 2003 (*Directiva 2003/98/CE*). Este campo es conocido internacionalmente como *open government data* (OGD) y en España como reutilización de información del sector público (RISP).

España aprobó hace ya 4 años, casi al mismo tiempo que la ley sobre administración electrónica, la normativa básica sobre reutilización (*Ley 37/2007*) y desde entonces se han ido desplegando paulatinamente estudios y proyectos, liderados por la *Administración General del Estado* y por algunas comunidades autónomas. Uno de ellos es el *Proyecto Aporta*, enmarcado en el *Plan Avanza 2* del *Ministerio de Industria, Energía y Turismo*, que está contribuyendo a hacer visibles en nuestro país las iniciativas sobre reutilización de la información en el sector público (no sólo de administraciones públicas, sino de muchas otras entidades de derecho público, como consorcios, fundaciones, etc.).

Entre las exclusiones especiales del marco jurídico nacional y europeo se encuentran las instituciones culturales (bibliotecas, archivos, museos).

No es la única exclusión en la directiva, pero es la que trataremos a continuación. La principal razón para esta excepción es la conveniencia de una regulación sectorial específica, por lo que hemos de partir de la idea de que el sector cultural no está excluido de la reutilización, sino tan sólo de la norma genérica.

Los tres aspectos clave que preocupan a las instituciones culturales para liberar sus datos son (sacado del informe *Problem of the yellow milkmaid*):

1. Potencial de ingresos. Hay que disponer de métricas que midan la cantidad del negocio facilitado a terceros (a industrias creativas) gracias a los datos liberados.

2. Pérdida de atribución. Las instituciones culturales son guardianas y conservadoras de la memoria colectiva, y desean que se mantenga una conexión indeleble entre ellas y sus objetos culturales. Tienen temor a que con la apertura de los datos se pierda su trazabilidad. Hace falta investigar sobre aspectos legales y técnicos para salvaguardar la integridad y la proveniencia de los datos.

3. Pérdida de ingresos. Miedo a la pérdida de la capacidad de vender los datos y de gestionar los ingresos cuando estén ampliamente disponibles con licencias abiertas y estén incluidos en otras ofertas creativas e innovadoras. Sin embargo puede establecerse una diferenciación por valor añadido: los datos pueden abrirse al público general de una forma y al mercado especializado de otra, con niveles de enriquecimiento diferentes.

Bibliotecas, archivos y museos (LAM)

Desde la publicación de la *Directiva* hasta ahora, la marea de los datos abiertos ha tenido eco en las bibliotecas y la cultura. En 2011 se concluyó el extenso informe del *Grupo Incubadora sobre Datos Abiertos Bibliotecarios* del W3C, dentro de las iniciativas para generalizar el uso de *linked open data*. Hemos asistido a impactantes liberaciones de datos de altas instituciones bibliotecarias (*Biblioteca Nacional de España, British Library, Deutsche Nationalbibliothek*) y a significativos pasos que indican que se acerca el momento de regular con mayor claridad las políticas de reutilización en el sector cultural.

Uno de los casos de mayor alcance es el nuevo acuerdo de intercambio de datos para participantes en *Europeana (Europeana Data Exchange Agreement)*, que sustituye al anterior *Data Provider and Data Aggregator Agreement*. En este acuerdo los metadatos difundidos desde *Europeana* utilizarán licencias de dominio público (CC0 1.0 *Universal public domain dedication*). Esta decisión se basa en la necesidad de reducir las barreras de uso, para estimular la innovación abierta de manera que aplicaciones y servicios incorporen metadatos de *Europeana*. Keller (2011) indica que los beneficios de abrir los contenidos para intercambiarlos superan los riesgos, así que permitirán incluso los usos comerciales.

La apuesta por la liberación total de datos parte de la conciencia de que en el contexto de *linked open data*, el papel de las instituciones culturales es promover que terceros usen datos fiables y de calidad (ver *Your metadata and Europeana*). Existe documentación detallada sobre la importancia de las licencias y de acuerdos flexibles con todos los proveedores de datos en *Europeana*. Podríamos resaltar una frase significativa sobre la competencia entre recursos de información en la Red:

“Si *Europeana* y sus proveedores no permiten estos usos, nos encontraremos superados por un medio que cambia con rapidez. *Wikipedia* y otros ocuparán nuestro lugar como la única fuente de confianza disponible” (Zeinstra; Keller).

El informe *New renaissance* diagnostica la necesidad de liberar datos y producir servicios

propios, pero también alerta de la sinrazón de buscar una explotación comercial directa o exclusiva de esos datos públicos: “Metadata related to the digitised objects produced by the cultural institutions should be widely and freely available for re-use”. Los datos (los metadatos) son un tipo de información específico, cuya regulación es de naturaleza diferente a los contenidos digitales de las industrias culturales. Un buen marco de referencia es el documento *Licensing open data: a practical guide* (Korn; Oppenheim, 2011) para el ámbito británico.

Podemos traer aquí el adagio con el que concluía Aguillo (2012) sus reflexiones sobre el horizonte tecnológico: “Los contenidos son el rey, pero con los metadatos se construyen imperios”.

Hay que declarar las licencias

Es posible que las administraciones culturales dediquen grandes esfuerzos a la calidad e interoperabilidad de los metadatos en la digitalización, pero descuiden la importancia de las licencias, como “tecnología legal” que permite el uso efectivo. Son bastante frecuentes las bibliotecas digitales y catálogos sin especificaciones adecuadas de derechos de reutilización. Sería útil algún tipo de etiqueta específica para los metadatos culturales. Durante la *LodLam summit* celebrada los días 2 y 3 de junio de 2011 en San Francisco, se propuso una nueva manera de evaluar la apertura y utilidad de los datos vinculados a la comunidad LAM mediante una clasificación de 4 estrellas.

El *Ministerio de Educación, Cultura y Deporte*, en la *Red digital de colecciones de museos de España* “se reserva el ejercicio exclusivo de los derechos de explotación sobre los contenidos de los catálogos de los museos de su titularidad y gestión, en cualquiera de sus formas (reproducción, distribución, comunicación pública y transformación), sin perjuicio de los derechos que puedan corresponder a los titulares de la propiedad intelectual de las obras o las imágenes que se reproduzcan en estos catálogos” (*Ceres*).

Se pronuncia sobre contenidos, pero es deseable que al diferenciar contenidos de metadatos, puedan esperarse mayores opciones de reutilización metadatos. La *Biblioteca Nacional de España* ha licenciado sus datos enlazados mediante *Creative commons zero* (CC0): confiamos en que su liderazgo influya en el resto de instituciones culturales bibliotecarias, de archivos y museos.

Revisión de la Directiva

Por otro lado, en la *Comisión Europea* se ha completado este año la propuesta de revisión de la *Directiva*, y es interesante contemplar cómo

trata la cuestión de los datos culturales. En mayo de 2011 se elaboró un informe sectorial sobre reutilización (*PSI re-use in the cultural sector*) desde el punto de vista de su compatibilidad con los principios de la *Directiva*, más que desde la obligatoriedad de hacer disponibles los datos. Indica que la falta de flexibilidad a la hora de permitir la reutilización puede bloquear la capacidad de las instituciones culturales para adoptar las licencias abiertas.

El estudio señala también carencias en la disponibilidad de catálogos normalizados online, lo que actúa de rémora en las adaptaciones tecnológicas que necesitarían en caso de una fuerte demanda de datos por parte de agentes sociales y empresas. Hay que añadir que los servidores de acceso a los datos exigen rendimiento y escalabilidad.

Se modificarán los derechos y las obligaciones establecidos en la *Directiva*, y en concreto se ampliará su ámbito de aplicación a los sectores actualmente excluidos (instituciones culturales, educativas y de investigación, además de organismos de difusión de los servicios públicos).

Un motor económico

Además de la argumentación tradicional del acceso a la cultura, se incluye su valor como motor económico: “la digitalización convierte estos recursos en un activo duradero para la economía digital y crea numerosas oportunidades para la innovación”; aunque considera que aún falta mucho recorrido. Por ello, el informe refuerza la importancia de “garantizar la disponibilidad generalizada de la información del sector público (*Directiva ISP*) y poner los bienes culturales digitalizados a disposición de empresas creativas e innovadoras (política de digitalización)”.

“No se trata sólo del acceso generalizado a la cultura. Además debe obtenerse rendimiento económico”

No es nuevo que la UE sitúe los aspectos económicos en primer plano. La digitalización y los datos culturales son la materia prima de la economía de la información, y no sólo un derecho educativo y de ocio. Otro asunto serían los “contenidos culturales”, por lo general las imágenes o los textos completos digitalizados. No es un asunto menor, pero hemos preferido centrarnos en los metadatos que permiten organizarlos, describirlos y conectarlos al espacio digital de información.

La lechera amarilla



La lechera, una de las más famosas piezas de Johannes Vermeer, representa la escena de una mujer que plácidamente vierte leche en un tazón. En una revisión del Rijksmuseum se descubrió que había más de 10.000 copias de la imagen en internet, en su mayoría pobres y amarillentas. —“Con tantas reproducciones deformadas la gente no creía que las tarjetas postales de nuestra tienda del museo mostrarán la pintura original. Eso nos movió a poner en la web una reproducción de la obra original en alta resolución, con los correspondientes metadatos. La apertura de datos es nuestra mejor defensa contra la ‘lechera amarilla’”, dicen desde el museo.

Referencias bibliográficas

Aguillo, Isidro F. “Tecnologías, investigación y futuro de la profesión”. *El profesional de la información*, 2012, enero-febrero, v. 21, n. 1, pp. 5-7.
<http://dx.doi.org/10.3145/epi.2012.ene.01>

Ceres. *Condiciones de uso*.
<http://ceres.mcu.es/pages/issued.jsp>

Clapton, Genevieve; Hammond, Max; Poole, Nick. *PSI re-use in the cultural sector: final report*. Curtis+Cartwright Consulting Ltd, 2011.
http://ec.europa.eu/information_society/policy/psii/docs/pdfs/report/cc462d011_1_1_final_report.pdf

Datos abiertos. Un motor para la innovación, el crecimiento y la gobernanza transparente.
http://ec.europa.eu/information_society/policy/psii/docs/pdfs/lopendata2012/open_data_communication/es.pdf

Directiva 2003/98/CE del Parlamento Europeo y del Consejo de 17 de nov. de 2003 relativa a la Reutilización de la información del sector público. Diario Oficial de la Unión Europea, 2003, 31 dic., pp. L 345/90-L 345/96.
<http://eu.vlex.com/source/doue-23/issue/2003/12/31/1>

Europe's Information Society. *Thematic portal. Public sector information - Raw data for new services and products*.
http://ec.europa.eu/information_society/policy/psii/index_en.htm

Europeana agreement.
<http://version1.europeana.eu/web/europeana-project/newagreement>

Europeana new agreement.
<http://version1.europeana.eu/web/europeana-project/newagreement>

Keller, Paul. *Open data in the cultural sector: cultural heritage data in Europeana*, 20 oct. 2011, Bruselas.
<http://www.slideshare.net/paulkeller/open-data-in-the-cultural-sector-cultural-heritage-data-in-europeana>

Korn, Naomi; Oppenheim, Charles. *Licensing open data: a practical guide (version 2.0)*. Hefce: JISC, junio 2011.
http://discovery.ac.uk/files/pdf/Licensing_Open_Data_A_Practical_Guide.pdf

Ley 37/2007, de 16 de nov., sobre Reutilización de la información del sector público. BOE n. 276, pp. 47160-47165.
<http://www.boe.es/boe/dias/2007/11/17/pdfs/A47160-47165.pdf>

LodLam. International linked data in libraries, archives and museums. *Proposed: a 4-star classification-scheme for linked open cultural metadata*.
<http://lod-lam.net/submit/2011/06/06/proposed-a-4-star-classification-scheme-for-linked-open-cultural-metadata>

New renaissance report.
http://ec.europa.eu/information_society/activities/digital_libraries/doc/refgroup/final_report_cds.pdf

Propuesta de directiva del Parlamento y del Consejo Europeo por la que se modifica la Directiva 2003/98/CE relativa a la reutilización de la información del sector público.
http://ec.europa.eu/information_society/policy/psil/docs/pdfs/directive_proposal/2012/es.pdf

Verwayen, Harry; Arnoldus, Martijn; Kaufman, Peter B. *The problem of the yellow milkmaid: a business model perspective on open metadata*, White paper n. 2, Nov. 2011.
<http://version1.europeana.eu/web/europeana-project/whitepapers>

Your metadata and Europeana.
http://version1.europeana.eu/c/document_library/get_file?uuid=c11f09d9-6579-4fc8-8a84-a754e1e23475&groupid=10602

Zeinstra, Maarten; Keller, Paul. *Open linked data and Europeana*.
http://version1.europeana.eu/c/document_library/get_file?uuid=374c381f-a48b-4cf0-bbde-172cf03672a2&groupid=10602

* * *

No hay problema en seguir avanzando Xavier Agenjo

Hispana es el cuarto proveedor de datos de *Europeana* y en torno al 50% de los datos de *Europeana linked open data* proceden así mismo de varios data sets españoles.



<http://www.europeana.eu/portal/europeana-providers.html>

<http://pro.europeana.eu/linked-open-data>

<http://thedatahub.org/data-set/europeana-lod>

En cuanto a la LOD-LAM está disponible la traducción oficial del W3C LLD.

http://www.larramendi.es/i18n/cms/elemento.cmd?grupo=estaticos/paginas&elemento=documentacion_tecnica_lam.html

Resumiendo: no hay problema para seguir avanzado en la reutilización de los datos culturales en España en el entorno Europeo o mundial.

Me permito así mismo recomendar dar un vistazo a:

http://www.larramendi.es/i18n/cms/elemento.cmd?id=estaticos/paginas/Biblioteca_Virtual_Escuela_de_Salam.html

En ella se ha llevado a cabo una reutilización de datos culturales de *Europeana*, los *Cultural heritage objects*. Los registros reutilizan otros registros europeos, que se agregan con autoridades enriquecidas –*enhaced*– en RDA.

xavier.agenjo@larramendi.es

Agregar y combinar datos no es suficiente

Tomás Saorín



Me gustaría precisar algo sobre *Europeana* y la reutilización (lo que sigue sirve igualmente para *Hispana* o cualquier biblioteca digital).

Europeana es preciosa pero no es suficiente. Agregar y combinar datos no es suficiente. En su estrategia *Europeana* entiende

la reutilización como un giro en su modelo de negocio, desde la oferta a la demanda. Menos acumular y más repartir.

La medida de la reutilización para *Europeana* es conseguir ser usada en productos digitales que la gente disfrute.

No basta acumular millones de metadatos, hay que atraer y la reutilización es un medio necesario para ello. Hasta que no se perviertan sus registros, y lleguemos a ellos sin darnos cuenta a través de *apps* y productos innovadores, no podremos hablar de reutilización con impacto.

tsp@um.es