

Eva Méndez Rodríguez: El mayor desafío que tienen las instituciones de la memoria es que el sector infomediario se fije en los datos culturales como una fuente de negocio

URL de la contribución <www.iaph.es/revistaph/index.php/revistaph/article/view/3983>



Eva Méndez Rodríguez, durante su intervención en el European Open Science Cloud Summit (Bruselas) | fuente Young European Research University Network (YERUN)

Eva Méndez es profesora titular e investigadora de la Universidad Carlos III de Madrid. Pertenece a la primera generación en España que es diplomada, licenciada y doctora en biblioteconomía y/o documentación. Su tesis obtuvo el premio extraordinario de doctorado de la Universidad Carlos III (2003). En el año 2015 recibió el Premio de Excelencia del Consejo Social UC3M-Banco de Santander en la modalidad de joven personal investigador.

Es autora y/o co-autora de más de 70 trabajos publicados, principalmente sobre metadatos, vocabularios y web semántica, gestión de información electrónica, sistemas y servicios de información digital, accesibilidad y normalización web y políticas de acceso a la información, *linked open data* y *open science*. Reviste particular interés la monografía *Metadatos y Recuperación de información: estándares, problemas y aplicabilidad en bibliotecas digitales* (Trea, 2002), así como el libro *Knitting the Semantic Web* (Haworth Press, 2007).

Reconocida experta o asesora de distintas organizaciones internacionales, fue una de las personas que participaron en la Declaración de la Haya *Knowledge Discovery in the Digital Age* (2015) y en el *workshop* de la Comisión Europea en la visión Open Science 2030. Miembro del Advisory Board de OpenAire2020 y del grupo de expertos de UNESCO de las Guidelines to Open Data and Open Data Repositories (2015). Pertenece a grupos de trabajo en España como el *thinktank* ThinkEPI o el Comité Técnico 50 de AENOR.

En relación a la actividad investigadora, merece especial atención su participación, desde 2006, como experto independiente de la Comisión Europea para evaluar/revisar proyectos de diversos programas, sobre todo en el ámbito de bibliotecas digitales, la mejora del aprendizaje a través de tecnologías y *open science*.

De 2009 a 2012 dirigió el Máster Universitario en Bibliotecas y Servicios de información digital. En la actualidad es Vicerrectora adjunta de Estrategia y Educación Digital de la Universidad Carlos III, miembro del Comité Ejecutivo de CRUE-REBIUN (Red Española de Bibliotecas Universitarias) y miembro del High Level Advisory Group de la Comisión Europea Open Science Policy Platform.

revista ph: *Open software, open data, open access, open science...* ¿lo *open* está de moda? ¿Qué hay detrás de estos conceptos y cómo se relacionan entre ellos? ¿constituye el conocimiento abierto una preocupación de las empresas, de los gobiernos, de los ciudadanos, de los investigadores... o es sólo una moda/tendencia pasajera?

Aunque pudiera parecerlo, *open* no es una moda pasajera, es una actitud provocada por las nuevas formas de comunicación en internet.

Es obviamente una tendencia que rodea a todas las manifestaciones de información y conocimiento a través de la red. Lo abierto ha convertido las publicaciones, los datos y el *software*, en abierto, pero también el gobierno, la cultura, la educación y la ciencia, llegando a convertir el conocimiento en abierto o la misma sociedad, en una sociedad abierta. Esta apertura, manifiesta en distintos ámbitos, dota a nuestra realidad de adjetivos no menos intensos y difíciles de explicar que el propio abierto: transparente, democrático, compartido, colectivo, libre, común e incluso procomún. No en vano el 2017 es The Year of OPEN, donde se pretende destacar que los procesos, herramientas y sistemas abiertos, creados desde un enfoque colaborativo, mejoran la educación, la ciencia, el gobierno y las organizaciones.

Pero, lo *open* es algo más que una tendencia, es un cambio de mentalidad provocado por el deseo de compartir no sólo recursos, datos o informaciones, sino ideas, procesos, resultados, pensamiento y formas de pensar o actuar. Esta tendencia a lo abierto genera un interés y una preocupación a gobiernos, investigadores, ciudadanos y empresas, y conlleva retos infinitos a todos los agentes implicados en la gestión de la información y el conocimiento, que incluye obviamente a las instituciones culturales, o instituciones de la memoria, que cada vez más denominamos GLAM (galerías, bibliotecas, archivos y museos).

La *open science* no es sólo una manifestación más de este cambio en la mentalidad, donde se entiende que la investigación financiada con fondos públicos, debe generar resultados públicos. La ciencia en abierto es

devolver la ciencia a quien la paga, y entender que la investigación pública es parte de la información de sector público (PSI, public sector information), y genera datos, información y conocimiento susceptible (de manera deseable) de reutilización.

revista ph: ¿Podría llamar nuestra atención sobre ejemplos de buenas prácticas en el uso de datos abiertos en países y sectores específicos, especialmente en el sector cultural? ¿Alguna iniciativa que merezca la pena seguir con especial atención?

Cada vez más, existen proyectos e iniciativas de datos abiertos, sobre todo de datos públicos para potenciar la transparencia de las administraciones, y también en otros ámbitos vinculados como las *smart cities*. El sector cultural no es una excepción en el ámbito de los datos abiertos, al menos, desde dos perspectivas.

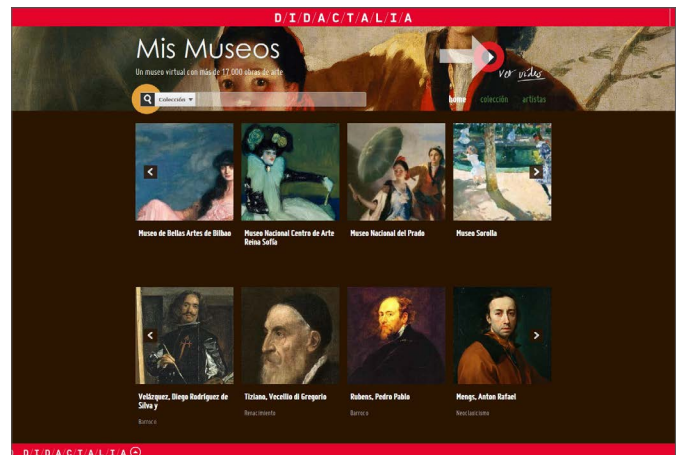
Por un lado, desde el punto de vista de que las instituciones GLAM están incluidas en la nueva Directiva sobre la Información del Sector Público (2013/37/EU), ya que bibliotecas, archivos y museos son instituciones custodias de información pública y por tanto susceptible de ser reutilizada. La directiva anterior (2003) se adoptó mucho antes que apareciera el movimiento de los datos abiertos, y desde una perspectiva tradicional que restringía la información pública a la información gubernamental.

Sin embargo la nueva directiva promueve la adopción firme y clara de los datos abiertos, en el sentido de la actitud de lo abierto, y se reclama el uso de estándares abiertos tanto para documentos como para metadatos. No en vano, muchas de estas instituciones están poniendo en el diagrama de Linked Open Data cloud (LOD) sus conjuntos de datos (catálogos, vocabularios, etc.). Así, por ejemplo, grandes actores en el sector de la biblioteca, como son la Library of Congress u OCLC han desarrollado sus modelos de datos abiertos utilizando tecnologías LOD.

El caso más emblemático es quizás Europea, que se ha convertido en algo mucho más allá que la biblioteca digital europea que comenzó siendo, en una infraestructura de servicios digitales (DSI) para la cultura y el patri-



Proyecto de publicación de datos como *linked open data* de la Biblioteca Nacional de España (<http://datos.bne.es/inicio.html>)



MisMuseos visibiliza *datasets* de museos para su potencial reutilización (<http://mismuseos.net/>)

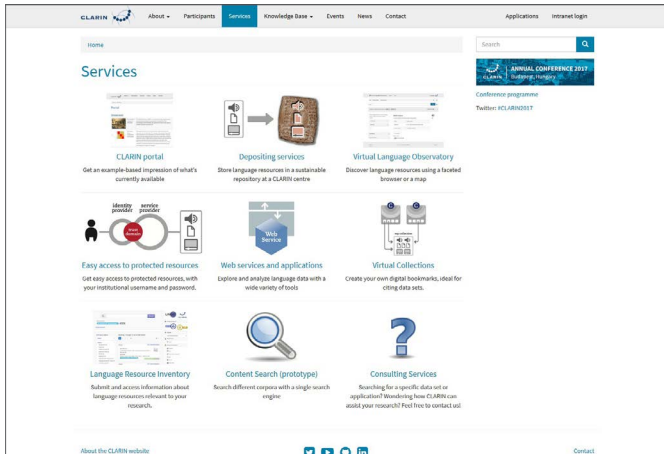
monio europeo, que incluye servicios de publicación de datos y metadatos de forma abierta, para hacerlos compartibles y reutilizables para la educación, la investigación y las industrias creativas. Europea da acceso hoy en día a más de 50 millones de registros, de los cuales, más de un millón de *datasets* están disponibles como *linked open data*. En nuestro país podemos destacar datos.bne.es y MisMuseos, que abren, enlazan y visibilizan *datasets* culturales para su potencial reutilización.

Por otro lado, porque las instituciones de la memoria en el entorno digital se generan en torno a la creación de activos digitales que, por su naturaleza, son datos para la investigación en humanidades, y también se generan repositorios e infraestructuras de datos de investigación en el ámbito cultural, dando cabida a todo un mundo de reflexiones e interacciones en las humanidades digitales. Las dos iniciativas más significativas en Europa en el ámbito de las infraestructuras para la investigación en humanidades son DARIAH (Digital Research Infrastructure for the Arts and Humanities), y CLARIN (European Research Infrastructure for Language Resources and Technology). CLARIN y DARIAH son en ambos casos ERIC, es decir, consorcios europeos de infraestructuras de investigación, que conectan, por un lado, comunidades de investigación en humanidades digitales (arte y lingüística, también) y, por otro, los datos que éstas producen, facilitando la investigación basada en datos, en el ámbito de las humanidades.

revista ph: La ley española 37/2007, de 16 de noviembre, sobre reutilización de la información del sector público, regula la reutilización de la información y los datos generados o custodiados por las administraciones públicas ¿Cuánto de avanzada está en España y América Latina la liberación de dicha información pública? ¿Cuáles son los principales inconvenientes que percibe en el proceso?

Se ha avanzado mucho, claro... han pasado 10 años desde esta primera ley de la reutilización en España. Sin embargo la Ley 37/2007, siempre me pareció una ley tímida, que podría haber ido mucho más allá. Es la transposición de la directiva europea de 2003 (Directiva 2003/98/CE), que quizás, también es tímida...

No obstante, en la última edición del Open Data Barometer (abril 2016) que mide cómo los gobiernos de los distintos países publican y utilizan los datos abiertos, España está en el puesto número 11, siendo el quinto país europeo en publicación y reutilización de datos abiertos. México (ya que me preguntas también por América Latina), comparte la misma posición que España, el 11 en el *ranking* mundial de *open data*, y es el país más destacado, junto a Uruguay, en este ámbito. Este barómetro, como todos los *rankings*, no es perfecto, pero nos da un indicador de los esfuerzos que están haciendo los gobiernos de los distintos países por compartir y reutilizar datos públicos, y el caso de España, es notable.



Clarín, consorcio europeo de infraestructuras de investigación (<https://www.clarin.eu/content/services>)



Open Data Barometer (<http://opendatabarometer.org>)

Abrir los datos públicos, y fomentar su reutilización, no es algo fácil, porque, como ocurre en otros ámbitos de la *openness* implica un gran cambio cultural. No obstante, los inconvenientes en este proceso son de distinta naturaleza... pero yo destacaría los aspectos legales y los técnicos. La parte legal tiene siempre muchas aristas sobre todo porque hablar de “abierto” debe de ser sinónimo de “libremente reutilizable” no de disponible como muchas veces se entiende. Y ello implica que los datos y los metadatos lleven aparejadas licencias abiertas. Por supuesto, el tema de la protección de datos que también es un cuello de botella para la apertura de datos.

Y en lo que respecta a la parte técnica, es fundamental utilizar estándares internacionales que permitan la reutilización, por ejemplo, la adopción del OCDS (*open contracting data standard*) para los datos sensibles de contratación, que aún no se contempla en España. Ni la ley de transparencia (19/2013), ni la de reutilización (37/2007) son determinantes en cuanto a la adopción de estándares.

revista ph: La modificación legal posterior, la Ley 18/2015, amplió la autorización de reutilización de los documentos a las bibliotecas, los museos y los archivos ¿Cuáles son los desafíos a los que se enfrentan las instituciones culturales en términos de derechos de autor, formación de su personal técnico, presupuestos, consulta con la ciudadanía y sector infomediario?

Efectivamente la ley de 2015 sobre Reutilización de la información del sector público (RISP), contempla una modificación, necesaria por otra parte, de la primera ley de RISP en nuestro país, no sólo por la ampliación que contempla la Directiva Europea de 2013, del ámbito de aplicación al patrimonio cultural digital que han generado bibliotecas, archivos y museos (instituciones GLAM, de las que hablábamos antes), sino también por todos los avances tecnológicos surgidos en el ámbito de los estándares de *open data* de los últimos 10 años.

La modificación de la ley, urge la necesidad de “acelerar el acceso de los ciudadanos al patrimonio cultural” (como dice el texto de la ley), lo que no siempre implica que exista una reutilización real de la información cultural. El mayor desafío que tienen las instituciones de la memoria en este sentido es justamente potenciar esa reutilización, que el sector infomediario se fije en los datos culturales como una fuente de negocio, además de difusión.

Para ello, las bibliotecas, archivos y museos, tienen que unirse a ese cambio de actitud y liberar, en lo que es su responsabilidad, sus metadatos, con licencias abiertas, además de fomentar esas mismas licencias para los materiales y activos digitales que describen a través de ellos. El resto de retos son justamente los que se han citado en la pregunta: la cualificación del personal que sea capaz de implementar técnica y legalmente todas

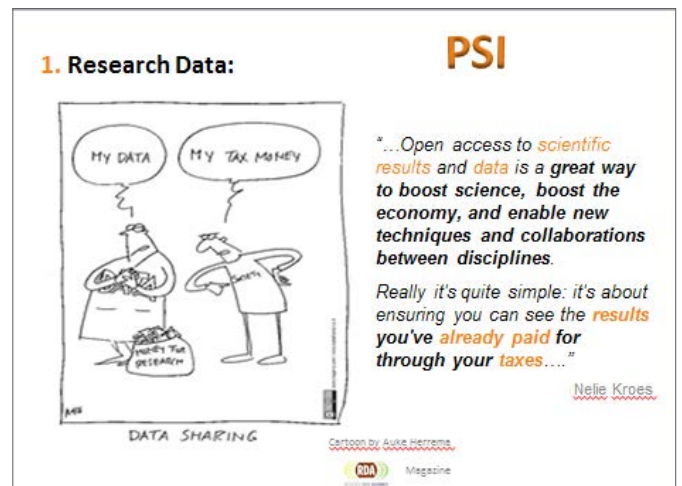
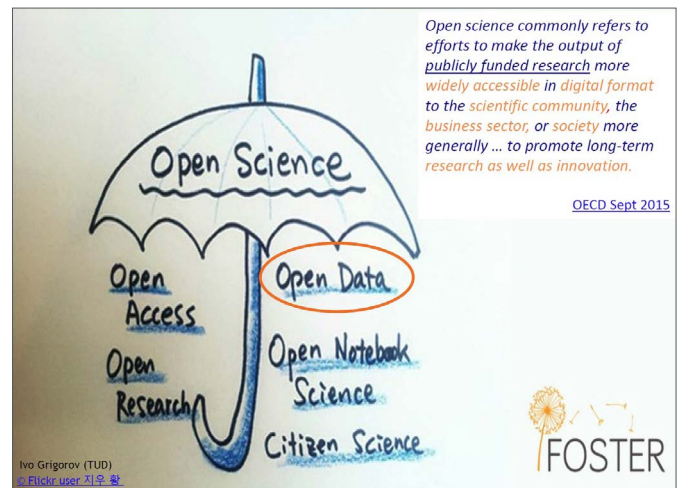
estrategias de apertura, amén del coste-oportunidad de abrir los datos culturales, que pasa por lo que decíamos antes, por incentivar el interés por su reutilización.

revista ph: Al parecer quedan al margen de la citada legislación los documentos resultantes de trabajos de investigación. Sin embargo, la Unión Europea está trabajando en el acceso abierto a todas las publicaciones científicas y en la reutilización de los datos de investigación. ¿Qué cambios pueden producirse en los próximos años?

Eso es algo muy curioso (bueno, no es curioso, es extraño y “mosqueante”) y siempre me ha sorprendido de la propia Directiva 2013/98/CE que traspone nuestra ley, claro, en los mismos términos. Ambas normas pues, excluyen “los documentos producidos o conservados por instituciones educativas y de investigación (incluidas las organizaciones para la transferencia de los resultados de la investigación, centros escolares y universidades, exceptuando las bibliotecas universitarias) así como los museos y archivos estatales como agentes de ejecución del Sistema Español de Ciencia, Tecnología e Innovación siempre que sean resultado de una investigación” (Ley 18/2015 art. 3, 3. g).

Justamente esa frase de la directiva y de nuestra Ley de 18/2015 de reutilización de la información del sector público dieron lugar a una ponencia que hice en la IODC2016 (International Open Data Conference) que se celebró el año pasado en Madrid. El título de mi charla era: *Are Scientific and Research Data PSI... Aren't they?* donde incidía en que los resultados de investigación financiada con fondos públicos dan lugar a información del sector público y, por tanto, su publicación debe ser abierta (*open access*), así como los datos resultantes de dichas investigaciones también deben de ser abiertos (*open data*), legitimando además, la reproducibilidad de la ciencia, y la investigación en sí misma.

Como decía esa excepción es “mosqueante” e invita a pensar que sea fruto del *lobby* de las grandes editoriales científicas. Sin embargo, la Comisión ya ha anunciado la revisión de la Directiva Europea. Para ello ha lanzado una consulta pública (Public Consultation on the review



Algunos conceptos desarrollados por Eva Méndez durante la ponencia en la IODC2016, con el título *Are Scientific and Research Data PSI... Aren't they?*

of the directive on the re-use of Public Sector information) donde anuncia además que está preparando una iniciativa sobre accesibilidad y reutilización de los datos públicos o financiados públicamente, donde parece que claramente se va a enfrentar el tema de los datos de investigación financiados con fondos públicos. .

En el caso de la ley española de RISP (ley 18/2015) existe una pequeña luz en este tema en la disposición adicional cuarta, se contempla la obligatoriedad de publicar los datos (en realidad, los microdatos) de las encuestas de investigaciones cuantitativas en sociología. ¡Algo es algo! Pero es insuficiente aún.

Además de la revisión de la directiva que promete hacer la Comisión Europea en el ámbito de RISP, en el ámbito de la *open science* y del dominio de acción de la Comisión Europea en investigación e innovación, la ciencia abierta es un plan. Es algo que va a pasar. Y hay una agenda de apertura de publicaciones y de datos científicos muy clara, concreta y ambiciosa, tanto desde el punto de vista de las políticas como de las infraestructuras. La *open science* provocará un cambio estructural en la ciencia donde Europa quiere llevar el liderazgo. Se pretende el *open by default*, y que todas las publicaciones y los datos de investigación estén disponibles de forma FAIR (*findable, accessible, interoperable, and reusable*)... Así pues los cambios que esto provocará en los próximos años, deben de ser también estructurales: el sistema de evaluación científica, la formación de investigadores y gestores de información en *open science* y gestión de datos, el impacto de la investigación en la sociedad, la ciencia ciudadana, nuevos códigos de integridad académica, y un sinfín de cambios más, que no han hecho más que empezar.

revista ph: ¿Cree que la investigación sobre temas culturales financiada con fondos públicos debería protegerse mediante derechos de autor y afines, patentes, secreto industrial u otras fórmulas para facilitar su comercialización por las empresas o, por el contrario, debería priorizarse su difusión en abierto, en la línea de lo señalado?

La investigación en temas culturales, en humanidades, en ciencias sociales y en todas las ciencias... si está financiada con fondos públicos, debe de difundirse en abierto como el resto de las disciplinas. Para mí no es, en absoluto, una excepción. Además de la ciencia abierta han de abrirse los procesos de innovación: la *open innovation*, que es otra de las tres grandes oes del Comisario Europeo de Investigación (*open science, open innovation, open to the world*). Justo una de las premisas básicas de la innovación abierta es abrir no sólo datos y conocimiento, sino el propio proceso de innovación para que todos puedan participar y el conocimiento se pueda transformar libremente en productos y servicios que creen nuevos mercados... ¿por qué no un nuevo mercado para las industrias culturales, basado en el emprendimiento en vez de en la propiedad intelectual *per se*?