

luación ha abierto la puerta a nuevas iniciativas con una vocación fuertemente comercial (bases de datos bibliográficas, repositorios de pago, servicios de alojamiento masivo, etc.) que podrían tener un efecto no deseado sobre el objetivo final de mayor descubrimiento, apertura y reutilización en aras del progreso científico y socioeconómico. Otras cuestiones que quedan para trabajo futuro se refieren a la siempre delicada cuestión de la gestión y preservación de la integridad de los datos, la protección de privacidad y las dificultades de atribuir adecuadamente estas obras, de fuerte impronta colectiva en muchas disciplinas, a sus correspondientes autores.

1. http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf

Mossink, Wilma; Bijsterbosch, Magchiel; Nor-tier, Joeri (2013). *European landscape study of research data management*. Utrecht: SURF. <http://goo.gl/CJNITX>

Nota original completa en *Blok de BiD*:
<http://www.ub.edu/blokdebid/les/node/440>

* * *

El camino incierto de la gestión de los datos

Alicia García-García

Universidad Católica de Valencia
alicia.garcia@ucv.es

Fernanda Peset

Universitat Politècnica de València
mpesetm@upv.es

García-García, Alicia; Peset, Fernanda (2015). "El camino incierto de la gestión de los datos". *Anuario ThinkEPI*, v. 9, pp. 231-234.

<http://dx.doi.org/10.3145/thinkepi.2015.53>

Publicado en *Blok de BiD* el 26 de febrero de 2014

El informe *Research data management: Principles, practices, and prospects* del *Council on Library and Information Resources* examina cómo los organismos de investigación y los profesionales de las bibliotecas y de ciencias de la información pueden responder a los requerimientos de las agencias de financiación para la gestión de los datos de investigación (*National Science*

Foundation-NSF y *National Institutes of Health-NIH*). Recoge seis estudios que pivotan sobre el proyecto *DataRes*:

1. Halbert, Martin. "Prospects for research data management".

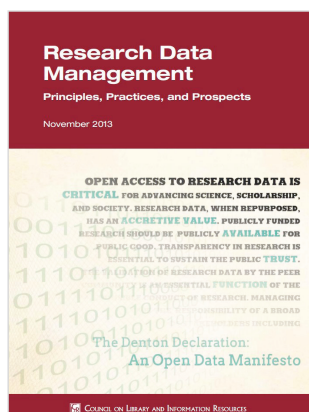
2. Keralis, Spencer D. C.; Stark, Shannon; Halbert, Martin; Moen, William E. "Research data management in policy and practice: The *DataRes* project"

DataRes utiliza encuestas (a bibliotecas y administradores), *textmining* sobre los mandatos de las agencias, y análisis de las políticas de algunos centros. Los resultados hacen visibles cinco factores críticos para alcanzar unas prácticas eficaces de gestión de datos de investigación (**González et al.**, 2013):

- Falta financiación para implantar las infraestructuras y los servicios que los investigadores necesitan. Se necesita más apoyo institucional para que las bibliotecas puedan proporcionar estos servicios.
- Falta de colaboración entre los órganos institucionales y departamentos. En este informe, aunque también en otros muchos foros, estamos escuchando lo necesaria que es la cooperación entre todos los actores del ciclo de comunicación académica.
- Los propios investigadores no dan especial prioridad a la gestión de sus datos, ni en los EUA ni en Europa (**Dallmeier**, 2012). En realidad los investigadores reclaman el reconocimiento de esta labor, y esto sólo será posible si las agencias de evaluación tienen en cuenta el esfuerzo por gestionarlos. Obviamente si los mandatos de los financiadores se extienden, el investigador se verá obligado. Reseñan un alto riesgo de pérdida de los datos, algo que los resultados preliminares de un cuestionario sobre gestión de datos a los investigadores de la *Universidad Politècnica de Valencia* (14 a 31 enero 2014, 224 respuestas) corrobora, pues casi la mitad lo almacena en ordenadores locales o en los instrumentos de medida.
- Falta de mandatos de las agencias financiadoras. Los autores observan que los mandatos van creciendo en número y son la vía más efectiva para que en las universidades se desarrollen las infraestructuras y servicios necesarios para gestionar los datos de los investigadores.



Agencias como *NSF*, *NIH*, *NEH-ODH*, etc., llevan años exigiendo planes de gestión (*data management plans*, *DMP*) en sus solicitudes. **Halbert** revisa la historia de los mandatos y destaca algo que puede llegar a ser un avance significativo: la



Office of Science and Technology Policy (OSTP) exigió en feb. 2013 a las agencias gubernamentales que hicieran públicas sus políticas de acceso a datos en seis meses, pero a mediados de septiembre aún no había dado sus frutos...

- Falta de políticas institucionales que aseguren la eficacia en la gestión. Pocas instituciones de investigación cuentan con una política que exija la preservación y el intercambio de datos de la investigación. Las existentes están poco consolidadas y muchas no se encuentran accesibles al público (sólo 20 instituciones). La mayoría de los encuestados valoran positivamente que exista una política de gestión para toda la institución.
- Necesidad de formar profesionales para gestionar grandes conjuntos de datos. Los autores proponen extender programas integrales sobre *data curation* en los planes de estudio de las escuelas de biblioteconomía y ciencias de la información, que sólo se están desarrollando en 9 centros. *DataRes* valora el papel que están asumiendo las 32 bibliotecas que proporcionan servicios de apoyo a los investigadores. Ofrecen un amplio abanico del que podemos aprender: servicios de consultoría para los planes de gestión, acceso web a las políticas y buenas prácticas, apoyo práctico sobre almacenamiento... Son más que suficientes para orientar a los bibliotecarios de nuestro país, que en breve habrán de asumir estas funciones (**Nina-Alcocer; Blasco-Gil; Peset**, 2013).

3. Crabtree, Jonathan et al. "The Denton declaration: An open data manifesto".

Se trata de un breve manifiesto, la *Declaración de Denton*, resultado final de los estudios desarrollados en este informe. Entre otras cosas explicita que existe una responsabilidad en la rendición de cuentas a la sociedad que se extiende más allá de las fronteras de cada una de las instituciones o de las disciplinas. Invita a suscribirlo en: <http://openacces.unt.eduldenton-declaration>

4. Deards, Kiyomi. "Why, how, and where we're going next: A multi-institution look at data management services".

Recoge los resultados de investigar la existencia de servicios de apoyo a la gestión de datos en cuatro centros universitarios. Sus resultados pueden orientar a los servicios de apoyo a la gestión de datos, ya que como reflejan los resultados preliminares del mencionado cuestionario de la *Universidad Politécnica de Valencia*, más del 80% de los encuestados cree que su organización no le ayuda con la preservación de los datos (estableciendo procedimientos para nombrar ficheros, guías de servidores externos donde almacenar, personal para subir los ficheros...). El autor proporciona algunos ejemplos de buenas prácticas: trabajar con las autoridades académicas, organizar talleres para alumnos sobre buenas prácticas en la gestión de datos, revisar las políticas de los repositorios institucionales incluyendo ciertos metadatos referidos a datasets, etc. Los servicios evolucionan con tanta rapidez que recomienda monitorizar iniciativas como *DataCite* o *DataOne*.

"El trabajo con datos no debe limitarse a los recursos de cada grupo de investigación, ha de ser una responsabilidad a escala nacional en el ámbito de la investigación"

5. Jordan, Chris; Esteva, Maria; Walling, David; Urban, Tomilsav; Kulasekaran, Sivakumar. "Responses to data management requirements at the national scale".

Revisa las iniciativas a escala nacional que pueden ser de utilidad en la gestión de datos, pues no es posible hacer frente a ciertos aspectos de manera individualizada, institución a institución. Además, insiste en que el trabajo con datos no debe limitarse a los recursos de cada grupo de investigación individual, que es lo que están actualmente fomentando los *DMP* que exigen las agencias de financiación. Ha de pasar a ser una responsabilidad a escala nacional en el ámbito de la investigación, lo que denomina "research effort", dado que forman parte de su patrimonio y potencial. Esta idea tiene sus derivaciones, ya que muestra dos tipos de intereses: un grupo de investigación estará siempre más interesado en la etapa en que los datos tienen utilidad, en que están "activos"; mientras que la preservación a largo plazo, su etapa "inactiva", ha de ser una responsabilidad a escala nacional. Esta distinción es fundamental para que las instituciones tomen

conciencia de su doble papel: apoyo a los investigadores y preservación de su producción.

Entre las iniciativas revisa: las soluciones distribuidas para almacenaje y análisis de datos o ciberinfraestructuras para la computación: XSEDE y DPN; los proyectos que desarrollan buenas prácticas: *iDigBio* y *iPlant Collaborative*, y *DataNet*, del que sorprendentemente no ofrecen palabras muy halagadoras; y los servicios de software para la gestión de datos durante la investigación: *DuraCloud*, que usa *Amazon S3* para el almacenamiento.

Sus autores recalcan que estamos en un momento inicial, en el que no se sabe cuál es la manera adecuada de proceder durante la investigación ni cómo sería un plan de gestión de datos correcto. En definitiva, muestran mucho camino por recorrer. Un viaje que, como siempre ha recomendado Lluís Anglada, hay que hacer juntos. La gestión de datos es como un ecosistema orgánico, cuyas partes en ocasiones evolucionan a diferentes ritmos y maneras con solapamientos que son percibidos como beneficiosos, sorprendentemente.

6. Jahnke, Lori M.; Asher, Andrew. "Dilemmas of digital stewardship: research ethics and the problems of data sharing".

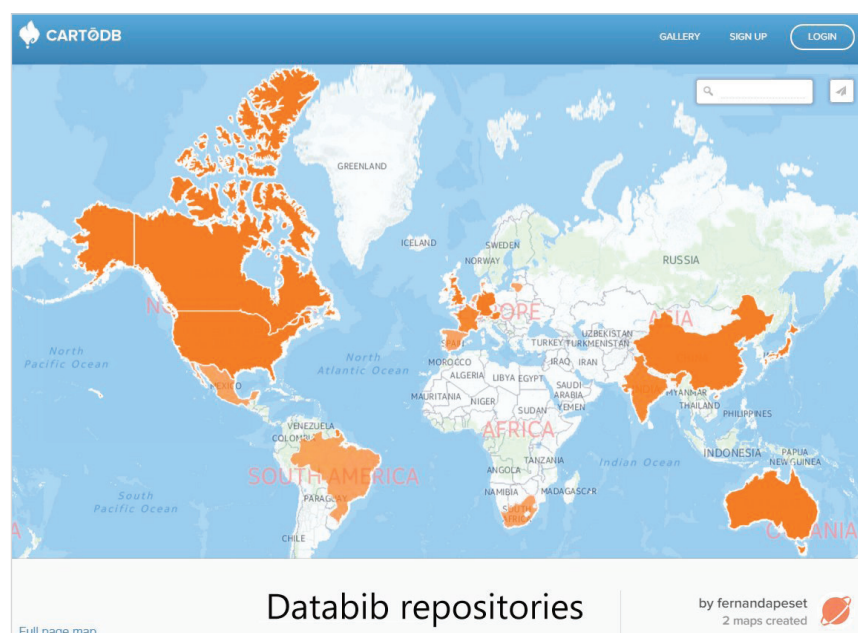
El último trabajo reflexiona sobre los problemas de almacenar los datos en la nube, especialmente cuando se utilizan servicios comerciales. Los autores llegan al extremo de pronosticar que si no se logra garantizar la privacidad de los datos en el futuro, se podría derivar a un sistema de ciencia distinto, con unas nuevas normas éticas o al abandono de la noción de confidencialidad actual. De hecho, hemos de reconocer que en ocasiones cualquier persona infringe lo que sería razonable en cuanto a garantía de privacidad. El autor pone de manifiesto que, como sucedió en el campo de la edición científica, de nuevo se está dejando entrar al sector comercial en el sistema de la ciencia. Y nos hace conscientes de que los objetivos de uno y otro sector pueden ser no sólo diferentes sino perjudiciales para que la ciencia siga siendo un bien común.

Un abaratamiento inicial de la externalización esconde unos costes que son insoslayables si se quiere garantizar la permanencia de los servicios. Ejemplo de gastos son la nece-

sidad de estudiar y negociar con los proveedores privados de servicios las condiciones de seguridad, licencias de uso, en estos momentos todavía no muy transparentes tanto en *Google*, como *Dropbox* o *WeTransfer*; o la gestión de litigios por infracciones de la política de privacidad.

Los autores detectan graves preocupaciones dada la escala a la que un dato puede ser compartido en la nube. Las preocupaciones de los investigadores con respecto al almacenamiento son: garantizar la privacidad de los datos sensibles; desconocimiento y ambigüedad sobre la propiedad de los datos cuando se trabaja a escala internacional; pérdida de control del uso de los datos. En el primer caso, los autores documentan la inseguridad de la nube con unos casos prácticos que producen escalofríos. Hacen tomar conciencia de que aunque los equipos de investigación necesitan compartir datos a escala multi-institucional, están asumiendo un riesgo en la toma de decisiones. Riesgo muy incierto, por demás. Para el segundo caso, la propiedad, el conocimiento legal necesario para tomar una decisión acertada no es tarea habitual en los equipos de investigación. Ni siquiera está al alcance de los abogados más cualificados, ya que no están suficientemente estudiados los problemas que pueden surgir en la nube. Respecto a la pérdida de control sobre los datos, los problemas no sólo se producen por accesos indebidos o el vacío legal respecto al uso de los datos, sino también por la pérdida de control sobre la integridad de los ficheros (datos corruptos).

Para concluir sólo cabe decir que este interesante texto da cuenta del grado de inmadurez de los aspectos legales y sociales, en un momento en que se han disparado las posibilidades técnicas que cualquier investigador tiene al alcance para



trabajar con sus datos. Advierte de que los equipos han de lidiar situaciones altamente complejas y multidimensionales sin suficiente conocimiento experto ni ayuda por parte de sus instituciones. Se encuentran en el centro de intereses contrapuestos que provienen tanto de las políticas de las comisiones de evaluación de proyectos o artículos, de los códigos éticos de sus sociedades científicas, de los mandatos de las agencias de financiación para compartir datos, de las políticas y normativas sobre propiedad y privacidad de los países con los que trabajan... Y en definitiva, no hay solución perfecta. Reproducimos la cita inicial del informe, muy ilustrativa:

"Nos encontramos, pues, que no hay una sola regla, por muy buena y por muy firmemente basada en la epistemología que esté, que no se viole en un momento u otro. Se hace evidente que tales violaciones no son accidentales, que no son el resultado de un conocimiento insuficiente o de una falta de atención que podrían haberse evitado. Sin embargo vemos que son necesarias para el progreso" (Feyerabend, 2010).

Las construcciones sociales progresan a base de incertidumbre y conflicto. Caminos tan inciertos como los de la propia investigación.

Referencias

Dallmeier-Tiessen, Sunje; Darby, Robert; Gitmans, Kathrin; Lambert, Simon; Suhonen, Jari; Wilson, Michael (2012). *Compilation of results on drivers and barriers and new opportunities*. <http://goo.gl/ChTK6>

Feyerabend, Paul (2010). *Against method: Outline of an Anarchist theory of knowledge* (4ª ed.). Brooklyn, NY: Verso Books. ISBN: 9781844674428

González, Luis-Millán; Saorín, Tomás; Ferrer-Sapeña, Antonia; Aleixandre-Benavent, Rafael; Peset, Fernanda (2013). "Gestión de datos de investigación: infraestructuras para su difusión". *El profesional de la información*, v. 22, n. 5, pp. 415-423. <http://dx.doi.org/10.3145/epi.2013.sep.06>

Nina-Alcocer, Víctor; Blasco-Gil, Yolanda; Peset, Fernanda (2013). "Los inicios del datasharing: guía práctica para compartir datos de investigación". *El profesional de la información*, noviembre-diciembre, v. 22, n. 6, pp. 562-568. <http://eprints.rclis.org/20907/1/datasharing.pdf> <http://dx.doi.org/10.3145/epi.2013.nov.09>

Council on Library and Information Resources (2013). *Research data management: Principles, practices, and prospects*. Washington, DC: CLIR. ISBN: 978 1 932326 47 5 <http://www.clir.org/pubs/reports/pub160/pub160.pdf>

Nota original completa en *Blok de BiD* <http://www.ub.edu/blokdebidles/node/459>

* * *

Establecimiento de un mercado justo de pagos por publicar en acceso abierto

Tomàs Baiget

El profesional de la información (EPI)

baiget@gmail.com

Baiget, Tomàs (2015). "Establecimiento de un mercado justo de pagos por publicar en acceso abierto". *Anuario ThinkEPI*, v. 9, pp. 234-237.

<http://dx.doi.org/10.3145/thinkepi.2015.54>

Publicado en *Blok de BiD* el 23 de abril de 2014

Objetivo y alcance de este informe

Como se sabe, los autores que quieren publicar artículos en revistas de acceso abierto (OA) comerciales, deben pagar la llamada "tarifa de proceso de artículos" (*article processing charges*), en adelante APC. Esto constituye la *vía dorada* al



acceso abierto.

Varias agencias europeas financiadoras de investigaciones comisionaron este informe para conocer el mercado de las tarifas que cobran las revistas con el objetivo de tener criterios cuando concedan las subvenciones.

Dentro de la "vía dorada" hay revistas que están subvencionadas al 100% por sus instituciones o por la publicidad, y que no cobran APC (algunos llaman a esto "vía platino"). Éstas también se consideran en este estudio.

Finalmente también se tienen en cuenta las revistas híbridas, que siguen financiándose por las suscripciones, pero ofrecen la llamada opción abierta (*open choice*), de manera que los autores pueden liberar sus artículos individuales pagando una APC.

Situación actual del mercado de APC

El mercado de APC está creciendo de un 30% (Laakso; Björk, 2012) a un 34% anual (Outsell, 2013), aunque se espera que la tasa de crecimiento disminuya a aproximadamente al 20% en los próximos 5 años, a medida que va aumentando su volumen (Outsell, 2013), que se estimó en aproximadamente 182 millones de US\$ en 2012.

Entre los editores de revistas OA indexadas en *Scopus*, el promedio de APC creció un 5% anual durante los últimos dos años, pero este aumento se debió a un reducido número de revistas: el