

aspectos críticos, y el avance del OA y las estrategias varían de país a país.

### Estudio 3. *Open data policies and strategies in the European Research Area and beyond*<sup>6</sup>

Trata de políticas y estrategias para el OA a los datos de investigación. Los gobiernos nacionales han centrado la atención primordialmente en los datos producidos por ellos mismos más que en procurar que se pongan en abierto los datos de la investigación de los organismos.



Los entes que financian la investigación no han sabido gestionar el depósito de los datos. Si se está todavía atrás con la implementación de políticas institucionales sobre el acceso abierto a las publicaciones, aún más atrasada se presenta la situación sobre los datos de la investigación. Éstos están en constante crecimiento, y en cambio es mínima la colaboración entre gobiernos, instituciones universitarias y de investigación, la sociedad civil y el sector privado para desarrollar la infraestructura, estándares, prototipos, y modelos de negocio.

Los estudios confirman que el iniciado camino al OA no tiene ya retorno, pero hay todavía mucho trabajo por hacer.

### Notas

1. Science-Metrix Inc.

<http://www.science-metrix.com>

2. Estudio 1. **Archambault, Eric; Amyot, Didier; Deschamps, Philippe; Nicol, Aurore; Rebout, Lise; Roberge, Guillaume** (2013). *Proportion of open access peer reviewed papers at European level and worldwide – 2004-2011*. European Commission DG Research; Innovation Science-Metrix Inc.

[http://www.science-metrix.com/pdf/SM\\_EC\\_OA\\_Availability\\_2004-2011.pdf](http://www.science-metrix.com/pdf/SM_EC_OA_Availability_2004-2011.pdf)

3. **Gargouri, Yassine; Lariviere, Vincent; Gingras, Yves; Carr, Les; Harnad, Stevan** (2012). "Green and gold open access percentages and growth, by discipline". In: *17th Intl conf on science and technology indicators (STI)*, Montreal, CA, 05-08 Sep 2012. 11 pp.

4. **Björk, Bo-Christer; Welling, Patrik; Laakso, Mikael; Majlender, Peter; Hedlund, Turid; Guðnason, Guðni** (2010). "Open access to the scientific journal literature: Situation 2009". *PLoS one* <http://www.plosone.org/article/info:doi/10.1371/journal.pone.0011273>

5. Estudio 2. **Caruso, Julie; Nicol, Aurore; Archambault, Éric** (2013). *Open access strategies in the European Research Area*. European Commission DG Research & Innovation Science-Metrix Inc.

[http://www.science-metrix.com/pdf/SM\\_EC\\_OA\\_Policies.pdf](http://www.science-metrix.com/pdf/SM_EC_OA_Policies.pdf)

6. Estudio 3. **Nicol, Aurore; Caruso, Julie; Archambault, Éric** (2013). *Open data policies and strategies in the European Research Area and beyond*. European Commission DG Research; Innovation Science-Metrix Inc.

[http://www.science-metrix.com/pdf/SM\\_EC\\_OA\\_Data.pdf](http://www.science-metrix.com/pdf/SM_EC_OA_Data.pdf)

**Archambault, Eric; Amyot, Didier; Deschamps, Philippe; Nicol, Aurore; Rebout, Lise; Roberge, Guillaume** (2013). *Proportion of open access peer reviewed papers at European level and worldwide – 2004-2011*. European Commission DG Research; Innovation Science-Metrix Inc. [http://www.science-metrix.com/pdf/SM\\_EC\\_OA\\_Availability\\_2004-2011.pdf](http://www.science-metrix.com/pdf/SM_EC_OA_Availability_2004-2011.pdf)

**Caruso, Julie; Nicol, Aurore; Archambault, Éric** (2013). *Open access strategies in the European Research Area*. European Commission DG Research & Innovation Science-Metrix Inc.

[http://www.science-metrix.com/pdf/SM\\_EC\\_OA\\_Policies.pdf](http://www.science-metrix.com/pdf/SM_EC_OA_Policies.pdf)

**Nicol, Aurore; Caruso, Julie; Archambault, Éric** (2013). *Open data policies and strategies in the European Research Area and beyond*. European Commission DG Research; Innovation Science-Metrix Inc.

[http://www.science-metrix.com/pdf/SM\\_EC\\_OA\\_Data.pdf](http://www.science-metrix.com/pdf/SM_EC_OA_Data.pdf)

Nota original completa en Blok de BiD

<http://www.ub.edu/blokdebid/es/node/449>

\* \* \*

## Alud de datos

### Isabel Bernal

*Unidad de Recursos de Información Científica para la Investigación, Consejo Superior de Investigaciones Científicas (CSIC)*  
[isabel.bernal@bib.csic.es](mailto:isabel.bernal@bib.csic.es)

**Bernal, Isabel** (2015). "Alud de datos". *Anuario ThinkEPI*, v. 9, pp. 228-231.

<http://dx.doi.org/10.3145/thinkepi.2015.52>

Publicado en Blok de BiD el 8 de enero de 2014



*Sim4rdm* (*Support infrastructure models for research data management*) es un proyecto europeo iniciado en octubre de 2011 y financiado por el 7 Programa Marco de la Comisión Europea.  
<http://www.sim4rdm.eu>

Se encuadra en el creciente número de iniciativas que, ante el alud de datos científicos, promueven buenas prácticas y estándares en su gestión para garantizar el progreso científico y económico mediante su acceso, reutilización y preservación digital. En concreto, *Sim4rdm* se marca como objetivo facilitar a la comunidad científica el uso efectivo de infraestructuras emergentes para datos de investigación y parte de la consideración previa del relativo desconocimiento por los investigadores de metodologías, instrumentos y capacidades necesarias. Para ello, el consorcio *Sim4rdm* se beneficia de experiencias acumuladas por las organizaciones participantes en el proyecto, entre las que destacan JISC (Reino Unido) y SURF (Holanda).

La piedra angular del informe *European landscape study of research data management*, publicado en mayo de 2013, gira alrededor de los resultados –desiguales en su cobertura geográfica y representatividad de las respuestas– de los cuestionarios enviados a 4 grandes grupos involucrados en la gestión de datos científicos en Europa (agencias de financiación, organismos nacionales, instituciones de investigación y editores) para definir sus acciones para que la comunidad científica mejore el manejo de sus datos una vez finalizados sus proyectos de investigación. El informe se completa con una introducción bastante exhaustiva que recorre, a partir de buenas prácticas existentes, los elementos que deben formar parte de las políticas y los planes de gestión de datos, y con un par de secciones anejas al análisis de las respuestas a los cuestionarios: por un lado, entrevistas a un grupo de investigadores sobre sus necesidades y un paquete de recomendaciones generales y específicas para cada grupo encuestado.

La sección introductoria ofrece una muy buena panorámica de las claves para diseñar políticas y planes de gestión de datos a partir del estudio de varios documentos de referencia, como el informe *Riding the wave* (2010), y su continuación *SURFboard for riding the wave* (2011), los *Principios y guía de la OCDE*, y varios ejemplos concretos de políticas y planes de gestión, principalmente del mundo anglosajón, como son los de la *National Science Foundation* (2011), la *US National Academy of Sciences* (1995) en los Esta-

dos Unidos, las recomendaciones del informe de *JISC Dealing with data* (2007), los principios de políticas de datos de los *Research Councils* en el Reino Unido (2012) y el plan de gestión de la *Australian National University* (2010).

Entre las muchas consideraciones, podemos destacar las relativas a los incentivos, a la formación y al apoyo para los investigadores, la inclusión de planes de gestión de datos en la preparación de propuestas, el desarrollo de infraestructuras y de códigos de conducta, la creación de perfiles profesionales específicos, la definición de responsabilidades entre las partes implicadas (autores de datos, agencias financiadoras e instituciones científicas) y las políticas de preservación digital y de utilización de los datos. Una parte especialmente interesante en la introducción se refiere a estudios sobre prácticas en la gestión de datos en disciplinas científicas concretas (p. e., *Eudat*) y sobre las posibles reticencias y carencias de la comunidad científica para compartir sus datos (p. e., el proyecto *Parse*).

El grueso del informe está compuesto por el análisis de las respuestas de las encuestas sobre políticas de gestión de datos y de las entrevistas a investigadores preparadas por el equipo de *Sim4rdm*. A pesar de su vocación europea, los resultados de los cuestionarios son relativamente desiguales y en algunos casos –especialmente, la encuesta a editores– han impedido ofrecer conclusiones suficientemente argumentadas. Sin embargo, el ejercicio en su conjunto refleja tendencias generales en materia de planificación de gestión y reutilización de datos a distintos niveles.

Las encuestas con el mayor número y variedad de respuestas son las dirigidas a agencias financiadoras y a organismos de investigación. Con respecto a las primeras, cabe destacar el creciente interés y la puesta en marcha de políticas para la financiación y la gestión de datos (casi un 50% de los encuestados), aunque la cifra cae con respecto a la obligatoriedad y la evaluación de planes de gestión de datos en las propuestas de proyectos.

En general, se evidencia que la garantía de preservación digital es otra área relativamente poco desarrollada por las agencias financiadoras. Por otro lado, el análisis de los resultados de la encuesta sobre políticas institucionales es la más rica en cuestiones y en



cobertura geográfica (participación de 15 países) y refleja que más de un tercio de las instituciones encuestadas disponen de una política, si bien suele ser de muy reciente creación.

Otro aspecto a resaltar es la precariedad de la mayoría de las muchas iniciativas en marcha para producir herramientas, recursos, guías y apoyo para mejorar la gestión de los datos, y en este sentido quizá pueda establecerse una relación con el bajo número de respuestas recogidas (fundamentalmente de Finlandia, Holanda y Reino Unido) en la encuesta a organismos nacionales encargados de crear códigos de conducta y de coordinar actividades -para el acceso, la formación, la preservación y creación de infraestructuras de datos y el apoyo a agencias financiadoras-.

Entre las principales motivaciones de las instituciones de investigación para desarrollar planes de acción destacan posibles requerimientos legales de agencias financiadoras o institucionales (sobre todo en lo relativo al depósito, gestión y preservación de los datos).

### **“El apoyo de las instituciones a la reutilización de los datos suele canalizarse mediante la creación de repositorios”**

La exhaustividad de los planes de gestión varía considerablemente entre las instituciones encuestadas aunque en mayor o menor medida en la mitad de los casos suelen incluir líneas de acción para la captura, mantenimiento y retención de datos, la aplicación de estándares técnicos y bibliográficos, y cuestiones de confidencialidad y de seguridad y de licencias de uso. Por otro lado, el apoyo de las instituciones a la reutilización de los datos suele canalizarse mediante la creación de repositorios, pero los sistemas de incentivos para que los investigadores depositen sus datos siguen siendo minoritarios, con sólo un cuarto de las instituciones encuestadas ofreciendo servicios de apoyo como la inclusión de los datos en las listas de publicaciones institucionales, la gestión de datos en la evaluación y auditorías de proyectos, la asignación de DOIs, la traducción de los metadatos al inglés para su mayor visibilidad, la financiación parcial para aplicar formatos estándares, etc.

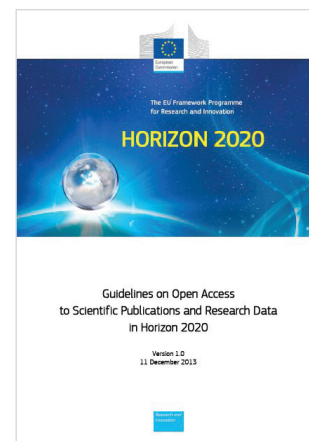
Las partes menos detalladas del informe corresponden a la encuesta a los editores y las entrevistas a investigadores aunque se dejan entrever cuestiones de relevancia, como la relativa carencia de políticas editoriales para el depósito y el acceso a los datos y la falta de consenso en la citación de los datos y su integración con

los artículos correspondientes. Por su parte, las 12 entrevistas a investigadores dejan en evidencia las cuestiones que ellos consideran prioritarias a la hora gestionar sus datos así como las necesidades más acuciantes: existencia de infraestructuras de depósito fáciles de usar, reconocimiento de la autoría de los datos, apoyo institucional, incentivos para su depósito, y formación.

Por último, el informe cierra con una serie de conclusiones generales y recomendaciones para que los organismos nacionales, agencias financiadoras, instituciones de investigación y editores promuevan una mejor gestión de datos por parte de la comunidad científica. El trabajo siguiente de *Sim4rdm* (hasta la primavera de 2014) incluirá más stakeholders en el ámbito de los datos científicos, como por ejemplo más órganos nacionales y organizaciones internacionales, centros de datos y proveedores de infraestructuras (bibliotecas), comités editoriales de revistas científicas y sociedades académicas.

Informes como éste son una prueba del interés que despierta este tema, y están alentados en gran medida por el auge del acceso abierto, la multiplicación de mandatos sobre la difusión de los resultados de investigación y por los desarrollos tecnológicos de última hora. El impulso de las agencias financiadoras es clave para cambiar de modo fundamental las prácticas en la gestión y la comunicación de estos resultados de investigación y las *Guidelines on open access to scientific publications and research data in Horizon 2020*<sup>1</sup>, publicadas en diciembre de 2013, tendrán sin duda un efecto positivo en el panorama europeo. Sin embargo, existen varios interrogantes en torno a la gestión de este trozo muy considerable de ciencia, empezando por la necesidad de abordar desde un consenso internacional una definición estándar de lo que entendemos por “datos de investigación” y no repetir los errores derivados de la relativa falta de armonización entre los múltiples modelos de mandatos de acceso abierto a las publicaciones científicas –en este sentido, la aplicación de estándares internacionales y la identificación de excepciones son importantes.

Por otra parte, la creciente facilidad tecnológica para recuperar información sobre su impacto y el tráfico web que genera, para analizarlos mediante técnicas de minería y las perspectivas de integrarlos en ejercicios de eva-





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](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/es/node/440>

\* \* \*

## El camino incierto de la gestión de los datos

**Alicia García-García**

Universidad Católica de Valencia  
[alicia.garcia@ucv.es](mailto:alicia.garcia@ucv.es)

**Fernanda Peset**

Universitat Politècnica de València  
[mpesetm@upv.es](mailto: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 Politécnica 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.

