

Compartir datos de investigación: beneficios e incertidumbres



Por Sandra Sánchez García.
Directora Técnica de la Biblioteca General del Campus de Cuenca, UCLM
Coordinadora editorial de Ocnos

Publicado el 25 de noviembre de 2016.

Hace unas semanas se celebraba la Semana del Acceso Abierto, un evento internacional de carácter anual centrado en el movimiento *open access*, en el que durante esta edición ha adquirido gran protagonismo el debate en torno a la gestión y el acceso abierto a los datos de investigación. La investigación científica genera cada vez un volumen mayor de datos o *datasets*, algunos se usan en las publicaciones pero la mayoría permanecen ocultos (datos estadísticos, entrevistas, gráficos, vídeos...), y en muchas ocasiones su acceso depende de la voluntad y predisposición de los investigadores.

Por este motivo, la preservación y el intercambio de datos científicos se han convertido en temas de interés a escala internacional, una tendencia que se conoce en el ámbito anglosajón como *data sharing*, y que es mucho más compleja y poliédrica de lo que pudiese parecer en un principio (Torres-Salinas, Robinson-García, Cabezas-Cavijo, 2012).

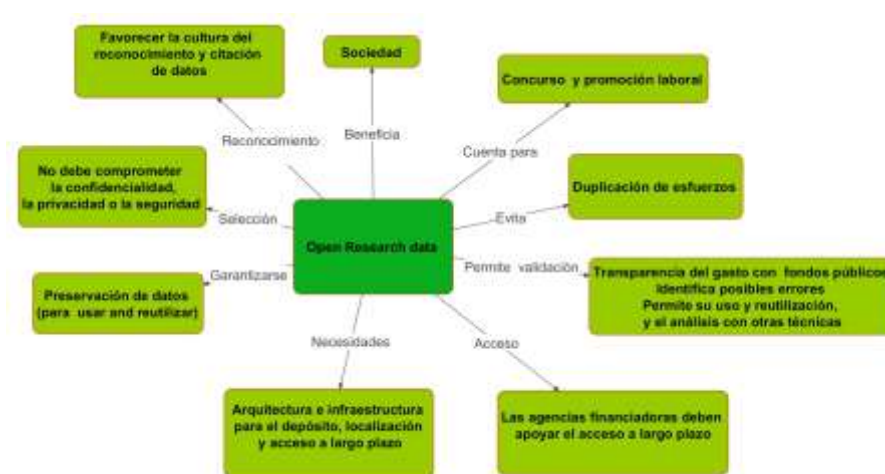
La investigación científica ha evolucionado en los últimos años gracias a los movimientos *open access* modificando en gran parte las prácticas de los investigadores y favoreciendo en cierta medida los flujos e intercambios de información. Además, el intercambio de información por parte de los investigadores no es una práctica novedosa, pero si es un hecho que los servicios de alojamiento de información en la nube (Dropbox, OneDrive o Google Drive) o redes sociales como Academia.edu o Researchgate lo han favorecido. Por ello, nos encontramos en un momento en el que los investigadores muestran una actitud más abierta a mostrar y difundir sus investigaciones, motivada en gran medida por las políticas open Access, la creación de repositorios y las herramientas de la ciencia 2.0. o ciencia abierta.

En este contexto marcado por el libre acceso a la información, los beneficios del *data sharing* son evidentes. En primer lugar supondrían un mayor rendimiento del dinero invertido en ciencia, contribuyendo a un mayor aprovechamiento de los recursos, especialmente en aquellos proyectos financiados con fondos públicos; pero además, facilitaría la labor de otros investigadores a la hora de replicar o proponer nuevos análisis, e incluso, ayudaría a la detección de errores, no necesariamente intencionados (Melero, Hernández-San-Miguel, 2014, Torres-Salinas, Robinson-García, Cabezas-Cavijo, 2012). Pero esta práctica aún causa recelo entre los investigadores, especialmente en determinadas

áreas o disciplinas. Como señalan González, Saorín, Ferrer-Sapena, Aleixandre-Benavent, Peset, (2013, 417), “existe una brecha entre el deseo de utilizar los datos ajenos (60%) y el de compartir los propios (40%).

En este contexto es importante que los investigadores tomen conciencia del valor de los datos y de su potencial para el reconocimiento de su trabajo, tanto en sus círculos profesionales, como por el aumento en la citación que puede generar su consulta y reutilización. En este sentido existen trabajos en los que se ha puesto de manifiesto cómo el acceso a los datos de investigación hace que aumente su citación (Melero y Hernández-San-Miguel, 2014). Cada vez encontramos mayor bibliografía sobre los beneficios que se obtienen de compartir, reutilizar o transformar datos, un ejemplo de ello es el manifiesto [Denton Declaration: An Open Access Data Manifiesto](#) (2012).

Figura 1. Esquema basado en los principios de la Denton Declaration: An open access manifesto (<http://openaccess.unt.edu/denton-declaration>)



Fuente: Melero y Hernández-San-Miguel (2014)

Pero para que el *data sharing* sea una realidad aún quedan muchos aspectos más por perfilar, los más importante centrados sin duda en aspectos técnicos, de organización y procedimientos. Como señalan Nina-Alcocer, Blasco-Gil y Peset (2013) en este momento no están claras las vías para hacer los datos disponibles, y menos aún en España. Las que se barajan coinciden en cierto modo con las conocidas ruta verde y ruta dorada.

La ruta verde incluiría la opción de utilizar repositorios o depósitos de datos, cada vez más numerosos como podemos observar en el directorio [OpenDOAR](#), donde se recogen actualmente 174 repositorios que admiten *datasets*, de ellos 39 son temáticos y 135 institucionales, y tan solo 4 gestionados por instituciones españolas ([Digital.Csic](#), [RIO](#) de la Universidad Pablo de Olavide, [UPCommons](#) de la Universitat Politècnica de Catalunya y el [Repositorio de la Universitat Pompeu Fabra](#)). En este sentido cada vez son más numerosos los directorios de depósitos de datos que ayudan a los investigadores a conocer los sistemas disponibles para almacenar y difundir los datos de sus investigaciones. En nuestro país encontramos [ODiSEA: International Registry on research Data](#), un inventario internacional de los depósitos y revistas que admiten datos de investigación a escala mundial. Este inventario permite al investigador buscar bancos de datos por materiales, formato y disciplinas.

La ruta dorada, incluiría la opción de almacenar el material adicional en la plataforma online de la propia revista junto al artículo, a través de tecnología propietaria como es el caso de editoriales como Sciverse o Wiley-Blacwell, o sistemas OAI-PMH compatibles con OJS para el caso de las revistas open access.

El análisis de las políticas editoriales muestran una correlación entre los índices de impacto altos y el hecho de que las revistas acepten *datasets* (González, Saorín, Ferrer-Sapena, Aleixandre-Benavent, Peset, 2013). Por lo que desde las editoriales de las revistas científicas también deberíamos verlo como una oportunidad para hacer nuestras revistas más atractivas. Muchas revistas ya se han posicionado a favor de este intercambio de datos, ofreciendo a los autores las herramientas necesarias para difundir los datos que subyacen a sus investigaciones. Incluso, sociedades científicas como la American Psychological Association (APA), o la International Spinal Cord Society (ISCoS), proporcionan a autores y editores recomendaciones para el manejo de los datos (Nina-Alcocer, Blasco-Gil y Peset, 2013).

Finalmente, la gestión de los datos de investigación también requiere de unas políticas y mandatos a escala nacional e internacional que guíen las prácticas de todos los implicados en el proceso. La Ley 14/2011 de 1 de junio de la ciencia, la tecnología y la innovación, que obliga a depositar en abierto las publicaciones financiadas con fondos públicos, no incluye mención específica a los datos objeto de la investigación. Este tema apenas está tratado por las universidades y organismos públicos de investigación (OPI), que son quienes producen los datos (González, Saorín, Ferrer-Sapena, Aleixandre-Benavent, Peset, 2013, 417). En este sentido, debemos de ser conscientes que “los datos que provienen de proyectos financiados con dinero público también son resultado de la investigación y por tanto, al igual que los artículos, deben hacerse públicos” (Torres-Salinas, Robinson-García, Cabezas-Cavijo, 2012, 178).

Por el momento las iniciativas emprendidas en este sentido vienen de organismos específicos y de áreas muy concretas, destacando las directrices y recomendaciones realizadas por organismos como el NIH, National Institutes of Health, la OECD (Organisation for Economic Co-operation and Development), e, incluso en España, el Centro Nacional de Datos Polares, en las que se insta a los investigadores a declarar cual será su plan de gestión de datos y su política de acceso a los datos, siendo en algunos casos requisito imprescindible la existencia de un compromiso de acceso abierto a los datos. A pesar de estos intentos puntuales, en gran parte de las agencias nacionales de investigación de la UE no existen políticas bien documentadas que sirvan como referente a los investigadores, a las universidades y a las agencias que financian la investigación.

En este sentido queremos hacer mención especial al programa marco Horizonte 2020, y a su cláusula 29.3, en la que se incluye un mandato específico para los proyectos financiados por la CE en siete áreas concretas, en las cuales tanto las publicaciones derivadas de estos proyectos como los datos de investigación deberán depositarse en repositorios y estar disponibles en abierto, y contar con un plan de gestión de datos. Además, pueden adherirse al mandato de forma voluntaria proyectos de otras áreas. En el documento [Guidelines on FAIR Data Management in Horizon2020 \(version 3.0, july 2016\)](#), se recogen las principales directrices que guían este mandato; y en la siguiente infografía, elaborada por CRUE-REBIUN con motivo de la Semana del Acceso Abierto se resumen algunos de sus aspectos fundamentales.

Figura 2. Infografía sobre el cumplimiento de la art. 29.3 del Programa Horizonte 2020 (<http://www.accesoabierto.info/como-cumplir-con-los-mandatos-sobre-gestion-y-publicacion-de-datos-en-2052328033.html>)



Fuente: CRUE-REBIUN

A pesar de ser una realidad cada vez más cercana, lo que es evidente que todavía hay muchos interrogantes pendientes en este sentido. Además en España no existen todavía infraestructuras técnicas preparadas para su almacenamiento, a excepción de los casos citados anteriormente, y lo que es más importante no existen políticas y mandatos a escala nacional que ayuden a materializar esta realidad. Nos encontramos en un momento incierto y lleno de interrogantes tanto para investigadores como para editores e instituciones de investigación, pero lo que es evidente es que la gestión de los datos de investigación se va a convertir en una actividad crucial en los años venideros.

Referencias bibliográficas:

- Melero, R.; Hernández-San-Miguel, J. (2014). Acceso abierto a los datos de investigación, una vía hacia la colaboración científica. *Revista Española de Documentación Científica*, 37 (4): e066. doi: <http://dx.doi.org/10.3989/redc.2014.4.1154>
- Millán-González, L.; Saorín, T.; Ferrer-Sapena, A.; Alexandre-Benavent, R.; & Peset, F. (2013). Gestión de datos de investigación: infraestructuras para su difusión. *El profesional de la información*, 22 (5), 415-423. doi: <http://dx.doi.org/10.3145/epi.2013.sep.06>
- Nina-Alcocer, V.; Blasco-Gil, & Peset, F. (2013). Datasharing: guía práctica para compartir datos de investigación. *El Profesional de la Información*, 22 (6), 562-568. doi: <http://dx.doi.org/10.3145/epi.2013.nov.09>
- Torres-Salinas, D.; Robinson-García, N.; & Cabezas-Clavijo, Á. (2012) Compartir los datos de investigación en ciencia: introducción al data sharing. *El Profesional de la Información*, 21 (2), 173-184. doi: <http://dx.doi.org/10.3145/epi.2012.mar.08>

Cómo citar esta entrada:

Sánchez García, S. (2016). Compartir datos de investigación: beneficios e incertidumbres. Debates en torno a las métricas fiables de reputación académica. *Aula Magna 2.0*. [Blog]. Recuperado de: <https://cuedespyd.hypotheses.org/2245>