



DATA JOURNALS: ECLOSIÓN DE NUEVAS REVISTAS ESPECIALIZADAS EN DATOS

Data journals: emergence of new journals specializing in data



Alicia García-García, Alexandre López-Borrull y Fernanda Peset



Alicia García-García es doctora por la *Universidad Católica de Valencia* en 2015, licenciada en documentación por la *Universidad Politécnica de Valencia (UPV)*, diplomada en biblioteconomía por la *Universidad de Valencia (UV)* y *Máster en contenidos y aspectos legales en la sociedad de la información (UPV)*. Participa en proyectos como *Vestigium*, sobre estudios de científicos y humanistas valencianos del siglo XX; en el inventario *ODiSEA* y el proyecto de I+D *Datasea*. En 2013 colaboró con el grupo de *Gestión de Estándares de Información Agrícola (AIMS)* de *Food and Agriculture Organization (FAO)* de *Naciones Unidas*. Anteriormente participó en el grupo de investigación *Uisys* en el estudio de la actividad científica y de redes sociales de colaboración, en el *Instituto de Historia de la Medicina y de la Ciencia López Piñero (UV)*.

<http://www.vestigium.es>

<http://odisea.ciepi.org>

<http://orcid.org/0000-0003-3597-3964>

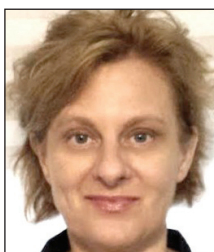
Universidad Politécnica de Valencia, Proyecto Datasea
Camino de Vera, s/n. 46022 Valencia, España
alimore17@gmail.com



Alexandre López-Borrull es profesor agregado de la *Universitat Oberta de Catalunya (UOC)*, donde es director de los *Estudios Información y Documentación*. Es doctor en química por la *Universitat Autònoma de Barcelona* y profesor ayudante en la misma universidad en el período 1998-2009. Licenciado en documentación por la *UOC*. Como investigador ha trabajado en aspectos legales de la información, fuentes de información electrónica en ciencia y tecnología, y ha participado en diversos proyectos. Sus intereses de investigación están relacionados con *open science*, *big science* y *e-research*. Forma parte del grupo de investigación *Knowledge and Information Management in Organizations (KIMO)*.

<http://orcid.org/0000-0003-1609-2088>

Universitat Oberta de Catalunya, Estudis d'Informació i Comunicació
Rambla del Poblenou, 156. 08018 Barcelona, España
alopezbo@uoc.edu



Fernanda Peset es titular de universidad de la *Universidad Politécnica de Valencia, Departamento de Comunicación Audiovisual, Documentación e Historia del Arte*. Viene del mundo de las unidades de información tradicionales –funcionario en el *Servicio de Información Bibliográfica* de la *Universidad de Valencia* hasta 1999-. Doctor por la *Universidad de Murcia* en 2002, es coordinadora del programa de doctorado *Industrias de la comunicación y culturales*. Su docencia y publicaciones se orientan a la comunicación científica, el acceso abierto y la implantación del protocolo OAI-PMH, normalización de la información, descripción de documentos, sistemas de documentación de museos y datos abiertos y enlazados. Participa en proyectos como *IraLIS*, en el gobierno de *E-LIS* y del *Grupo Ciepi*, el inventario *ODiSEA* y ahora dirige el proyecto de I+D *Datasea*.

<http://eprints.rclis.org/information.html>

<http://www.datasea.es>

<http://orcid.org/0000-0003-3706-6532>

Univ. Politécnica de Valencia, Depto. de Comunicación Audiovisual, Documentación e Historia del Arte
Camino de Vera, s/n. 46022 Valencia, España
mpesetm@upv.es

Resumen

Dos hechos principales, el ingente incremento de la cantidad de datos que los científicos producen, y la presión por la transparencia y la eficiencia económica de los presupuestos públicos, han ocasionado que la gestión de los datos haya adquirido mayor importancia. Se presenta una compilación y un análisis de algunas revistas especializadas en datos o *data journals*, de los que se describe su origen, evolución que están experimentando y características. Se reflexiona sobre el papel que pueden desempeñar en el ecosistema de la comunicación científica.

Palabras clave

Revistas de datos; Artículos de datos; Repositorios de datos; Publicación de datos; Publicación científica; Datos abiertos; Revisión por pares.

Abstract

Two major facts: the huge increase in the amount of data that scientists produce, and the pressure for transparency and economic efficiency of public budgets, have caused that a good data management has become much more important. A compilation and an analysis of some data journals –journals specializing in data- is presented. Their origins, evolution and characteristics are described. Finally, a reflection is done on the role that data journals can play in the scientific communication ecosystem.

Keywords

Data journals; Data papers; Data repositories; Data publication; Scientific publishing; Open data; Peer review.

García-García, Alicia; López-Borrull, Alexandre; Peset, Fernanda (2015). "Data journals: eclosión de nuevas revistas especializadas en datos". *El profesional de la información*, v. 24, n. 6, pp. 845-854.

<http://dx.doi.org/10.3145/epi.2015.nov.17>

1. Introducción

Las revistas científicas y la comunicación científica en general están en pleno proceso transformación. A finales del siglo XX la fuerza motriz del cambio se dirigió a adaptarlas a las versiones electrónicas (Hildyard; Whitaker, 1996). Ahora el exceso de revistas hace que compitan entre ellas para mejorar en los rankings, y algunos las emplazan a fusionarse (Rodríguez-Yunta; Giménez-Toledo, 2013), a crear nuevos modelos de negocio (López-Borrull, 2012; Villarroya et al., 2012) e incluso a renovar el modelo de artículo (Cronin, 2013).

La evolución de las tecnologías de la información y la comunicación (TIC) ha provocado muchos de los cambios ocurridos en la comunicación científica. Destacan el movimiento *open access* (Wellen, 2013), el *open data* (Peset; Ferrer-Sapena; Subirats-Coll, 2011) y el aumento de la colaboración científica para poder llevar a cabo la investigación (Ortoll et al., 2014). Entre los principales retos de las revistas están la competencia que sufren por parte de las redes sociales generales y académicas como canales de información paralelos; la mejora de su calidad tanto formal como de contenidos; y la realización de marketing académico para lograr atraer buenos autores y lectores (López-Borrull, 2014).

Otro reto de las revistas es la adopción de una estrategia para ofrecer los datos utilizados por los autores de sus artículos, para que en el futuro cualquier usuario pueda reutilizarlos, libremente, pagando o mediante la firma de un acuerdo. Como paso previo la revista debe disponerlos con los correspondientes metadatos para que sean citables y encontrables (Halbert, 2013; Kratz; Strasser, 2014). Los datos podrán ofrecerse en la forma y condiciones que acuerden las revistas y los autores. En caso de que sean éstos los en-

cargados de publicar sus datos de investigación tienen tres posibilidades (Lawrence et al., 2011; Pampel; Dallmaier-Tiessen, 2014; Peset; González, en prensa):

1) Publicación como objeto de información independiente en un repositorio de datos de investigación.

El punto principal a favor de esta posibilidad es que asegura la preservación y el correcto tratamiento de datos si los repositorios cuentan con una calidad reconocida. Por el contrario, puede tomarse como debilidad la separación entre los datos y las conclusiones que se derivan de ellos, ya que se publican en artículos que se depositan en otros repositorios. Incluso aunque exista un enlace entre ellos, tal y como sucede desde hace años en muchas disciplinas, dificulta la serendipia. Este sería el caso de *Cambridge structural database* o de *Protein data bank* (López-Borrull, 2003), que ofrece acceso a los datos asignándoles un código que se añade en los artículos.

2) Publicación de datos de investigación en forma de *data paper* en un *data journal*.

Las revistas de datos son diferentes de las revistas tradicionales que conocemos. Amplían la forma clásica de concebir los artículos de investigación y preconizan un futuro acercamiento, casi un solapamiento, entre los repositorios de datos y las revistas científicas. Por ello, parece oportuna una reflexión mayor sobre tales revistas de datos, que según nuestro punto de vista constituye uno de los últimos intentos de hacer avanzar la comunicación científica desde concepciones más clásicas.

3) Publicación de datos de investigación junto al artículo, en la forma de *enriched / enhanced publication*. Presentar en un mismo sitio artículos y datos, puede considerarse óptimo

para difundirlos. Esta es la estrategia por la que ha optado por ejemplo *PLoS* (Silva, 2014). No obstante, revistas como las publicadas por la *International Union of Crystallography* llevan años publicando los datos junto al artículo, antes incluso de que se hablase de datos abiertos. Quizá debido a que es una publicación que nació ya online aceptaba desde un principio este tipo de datos. No es necesario que artículo y datos estén físicamente en el mismo sitio: pueden estar distantes enlazados entre sí (Vernooy-Gerritsen, 2009).

Dos hechos principales, el ingente incremento de la cantidad de datos que los científicos producen, y la presión por la transparencia y la eficiencia económica de los presupuestos públicos, han ocasionado que la correcta gestión de los datos haya adquirido mayor importancia.

Cuando se trata de investigación con cargo a fondos públicos, las instituciones financiadoras van obligando a que los autores hagan públicos sus datos al igual que las publicaciones (OCDE, 2006).

En este artículo se presenta una compilación y un análisis de algunas de las revistas especializadas en datos, de las que se describe su origen, la evolución que están experimentando y sus características.

Editoriales como *Nature* o *Wiley* apuestan por las revistas de datos

2. Metodología

Este trabajo se sustenta metodológicamente sobre dos pilares: la revisión bibliográfica y la recopilación de los títulos de revistas con artículos de datos. En la revisión se han consultado no sólo artículos de investigación sino también normas como por ejemplo de la Unión Europea, informes derivados de proyectos como *ODE (Opportunities for data exchange)* o declaraciones de organismos como la OCDE.

<http://ec.europa.eu/research/swafs/index.cfm?pg=policy&lib=science>
<http://www.alliancepermanentaccess.org/index.php/community/current-projects/ode>
<http://www.oecd.org/sti/outlook/e-outlook/stipolicyprofiles/interactionsforinnovation/openscience.htm>

La compilación de revistas se efectuó entre octubre de 2014 y abril de 2015. Se registraron los títulos que exclusivamente publican artículos de datos. Se obtuvo información a partir de varias fuentes:

- lecturas comentadas en el apartado anterior;
- proyecto *ODiSEA*
<http://odisea.ciepi.org>
- información hecha pública por el proyecto *JoRD (Journal research data)*:
<http://www.webarchive.org.uk/wayback/archive/20140701173800>
<http://jordproject.wordpress.com>
- blog de Sarah Callaghan
<http://proj.badc.rl.ac.uk/preparde/blog/DataJournalsList>
- contactos personales.

El principal filtro fue diferenciar entre revistas que aceptan datos como material adicional, y revistas que casi exclusiva-

mente se dedican a publicar artículos de datos, que son el dominio de estudio de esta investigación. Por ello, del total de 75 revistas recopiladas, finalmente sólo se identificaron 20 como revistas de datos.

La información se sistematizó en un documento en *Drive* con los siguientes campos: título de la revista, año de aparición, periodicidad, categoría temática, relación con los repositorios de datos, editorial, url y otros aspectos de descripción (tabla 1).

Los datos recogidos han sido objeto de análisis estadísticos básicos, mediante tablas dinámicas de *MS Excel*, y observación cualitativa individualizada con el fin de identificar:

- el crecimiento anual de las revistas de datos, tomando como fecha de aparición el año que reseñan;
- las disciplinas más representadas, agrupadas manualmente según las categorías de la base de datos *Essential indicators* de *Thomson Reuters*. Consideramos que su grado de desagregación en 22 categorías es adecuado para este estudio.
- las características que conforman una descripción general de qué es una revista de datos.

3. Data journals y data papers

Chavan y Penev (2013) definen un *data paper* (artículo de datos) como “una publicación en una revista cuyo propósito es describir datos en vez de informar de una investigación o sus conclusiones”. Así, contendría los datos sin las hipótesis ni los argumentos, los resultados o la discusión que se haya llevado a cabo. Por extensión, un *data journal* sería una revista especializada en la publicación de *data papers*. Para Whyte et al. (2013):

“Un artículo de datos describe una colección de datos, un proceso, un software, formatos de archivo, etc., sin el requisito de un análisis novedoso o de conclusiones innovadoras. Describe cuándo, cómo y por qué se recogieron los datos y en qué consiste el producto”.

Esta diferenciación es importante porque en ciencia y tecnología existían hasta ahora tres tipos principales de artículos de investigación: artículos normales de investigación, artículos de revisión, y notas breves (Bosch; Mas; Moyano, 1999; López-Borrull, 2003). Un artículo de datos sería un nuevo tipo de comunicación científica (Newman; Corke, 2009).

A lo largo de este trabajo se han identificado 20 títulos. El más antiguo, *Atomic data and nuclear data tables*, apareció en 1969, publicado por *Elsevier*; y el segundo de 1956, *Journal of chemical and engineering data*, por la *American Chemical Society (ACS)*. A partir de entonces su aparición es bastante irregular, tal y como muestra la figura 1. Se concentra en 2013 con cinco títulos, siendo el último de 2015, publicado por *Inderscience Publishers*. En los últimos diez años han aparecido 14 de las 20 revistas estudiadas.

Con respecto a las editoriales, observamos que algunas con larga tradición han creado un título, como *Nature* o *Wiley & Sons*. La editorial con mayor cantidad es *Pensoft Publishers*, con 3, a la que siguen *Metajournals*, *Elsevier* y *Ubiquity* con 2 títulos cada una. Es importante reseñar, pues, que algunas de las principales editoriales científicas apuestan por este

Tabla 1. Revistas de datos identificadas

Revista	Año de aparición	Periodicidad	Disciplina	Disciplina Essential Indicators	Editorial	Url
<i>Atomic data and nuclear data tables</i>	1969	bimestral	Atomic physics, nuclear physics	Physics	Elsevier	http://www.journals.elsevier.com/atomic-data-and-nuclear-data-tables
<i>Biodiversity data journal</i>	2013	asap	Biodiversidad	Biology & biochemistry	Pensoft Publishers	http://biodiversitydatajournal.com
<i>Biomedical data journal</i>	2014	semestral	Biomedicina	Molecular biology and genetics	Procon	http://biomed-data.eu
<i>Codata data science journal</i>	2002	asap	Science and technology	Multidisciplinary	Codata	http://datascience.codata.org/about
<i>Dataset papers in science</i>	2013	variable	Science and medicine	Clinical medicine	Hindawi Publishing Corporation	http://www.hindawi.com/journals/dpis
<i>Earth system science data</i>	2009	bianual	Earth system science	Geosciences	Copernicus Publications	http://www.earth-system-science-data.net
<i>Ecological research</i>	1986	bimestral	Ecology, plant sciences, zoology, evolutionary Biology, Behavioral sciences, forestry	Plants, animal science	Springer Japan	http://link.springer.com/journal/11284
<i>Geoscience data journal</i>	2014	semestral	Geoscience	Geosciences	John Wiley & Sons	http://onlinelibrary.wiley.com/journal/10.1002/%28ISSN%292049-6060
<i>Gigascience</i>	2012	asap	Life and biomedical sciences	Molecular biology and genetics	BMC Publishers	http://www.gigasciencejournal.com
<i>International Journal of data science</i>	2015	trimestral	Big data cloud, mining and management, Business analytics, intelligence and mathematics, Computer science	Computer science	Inderscience Publishers	http://www.inderscience.com/jhome.php?jcode=ijds
<i>Journal of chemical and engineering data</i>	1956	asap	Chemical and Biomolecular Engineering	Biology & biochemistry	ACS Publications	http://pubs.acs.org/journal/jceaax
<i>Journal of open archaeology data</i>	2012	asap	Archaeology	Social sciences, general	Metajournals	http://openarchaeologydata.metajnl.com/articles?filter=all&order=date&pp=10&page=1&current_page=0
<i>Journal of open psychology data</i>	2013	asap	Psychology	Psychiatry / psychology	Metajournals	http://openpsychologydata.metajnl.com/articles
<i>Journal of open research software</i>	2013	asap	Open source research software	Computer science	Ubiquity Press	http://openresearchsoftware.metajnl.com
<i>Journal of physical and chemical reference data</i>	1972	anual	Physical science and engineering	Engineering	AIP Publishing	http://scitation.aip.org/content/aip/journal/jpcrd
<i>Nuclear data sheets</i>	1971	mensual	Nuclear physics	Physics	Elsevier	http://www.journals.elsevier.com/nuclear-data-sheets
<i>Open health data</i>	2013	asap	Health and medical data	Clinical medicine	Metajournals	http://openhealthdata.metajnl.com/about/submissions#authorGuidelines
<i>Phytokeys</i>	2010	variable	Taxonomy, phylogeny, biogeography and evolution of plants	Environment / ecology	Pensoft Publishers	http://www.pensoft.net/journals/phytokeys
<i>Scientific data</i>	2014	asap	Interdisciplinar	Multidisciplinary	Nature	http://www.nature.com/sdata/about
<i>Zookeys</i>	2008	variable	Biodiversidad	Environment / ecology	Pensoft Publishers	http://zookeys.pensoft.net

asap = los artículos se van publicando en modo continuo, a medida que se terminan

Esta tabla está accesible como "dataset" en:

<http://dx.doi.org/10.6084/m9.figshare.1549666>

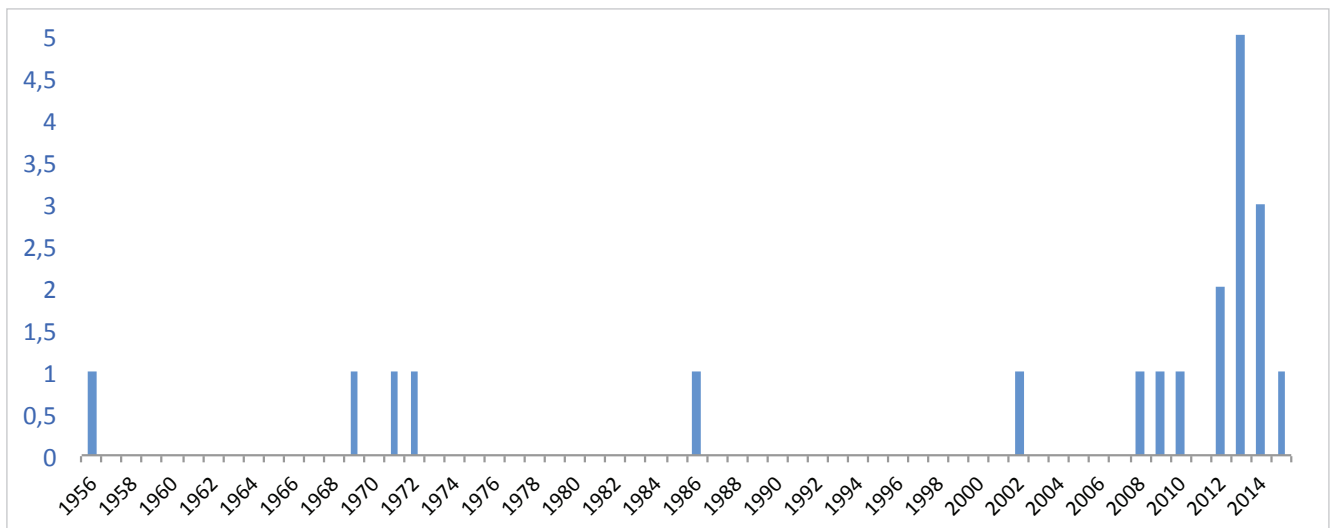


Figura 1. Evolución de la aparición de las revistas de datos

tipo de revistas, lo que permite suponer que pueden tener continuidad y un posterior crecimiento.

Del conjunto de revistas estudiado 17 seleccionan un repositorio donde depositar los conjuntos de datos, que pueden ser generales, como por ejemplo *Figshare*, o de la misma disciplina de la revista. En la mayoría de los casos las revistas son de *open access* vía dorada (con pago por publicar o *article processing charge*, APC, aunque como estrategia de promoción algunas aún no lo exigen).

El análisis por disciplinas muestra el predominio de la medicina y las ciencias de la tierra, si bien toda clasificación tiene su sesgo. Aparece una situación equilibrada, con doce materias representadas, aunque sorprendentemente no hay ninguna que pueda asignarse a un campo tan activo como las ciencias agrícolas. La principal característica, pues, sería que pertenecen a múltiples disciplinas, como puede apreciarse en la tabla 2.

4. Estudio de algunas revistas de datos

A continuación, se presentan breves estudios de cuatro casos representativos de este tipo de revistas:

Tabla 2. Categorías temáticas de las revistas de datos, según las de *Essential indicators* de Thomson-Reuters

Categoría temática	Núm. de revistas
Physics	2
Biology & biochemistry	2
Molecular biology and genetics	2
Multidisciplinary	2
Clinical medicine	2
Geosciences	2
Plants, animal science	1
Computer science	2
Social sciences, General	1
Psychiatry & psychology	1
Engineering	1
Environment & ecology	2

- una de una editorial de gran prestigio que ha apostado por la publicación de conjuntos de datos en forma de artículo;
- una que ya se publicaba en formato papel en los años 70 y que ha evolucionado hasta convertirse en revista de datos;
- una de una editorial de las que innovan en los métodos de publicación y realiza un proceso de revisión semiabierto;
- la recientemente relanzada revista del *Committee on Data for Science and Technology (Codata)* del *International Council for Science (ICSU)*.

En los últimos diez años, han aparecido 14 de las 20 revistas de datos existentes

Se comentarán las políticas de publicación y la información que ofrecen a los autores, para poder estudiar los criterios de selección y admisión. La revisión es lo que diferencia principalmente a una revista de datos de un repositorio de datos, además de su periodicidad y formato de presentación. La revisión de un artículo de datos se hace a dos niveles (**Mayernik et al.**, 2015), técnico y de contenido. Para algunos autores este segundo podría no producirse, pues es un tipo documental que no tiene por qué aportar preguntas de investigación, discusión o conclusiones (**Parsons; Duerr; Minster**, 2010).

Scientific data

<http://www.nature.com/sdata>

Esta revista del grupo *Nature* apareció en 2014 con los objetivos de “maximizar la reutilización de los datos y permitir la búsqueda, el enlace y su análisis con técnicas de *data-mining*”. Su objetivo es publicar de la forma más rápida posible, sin periodicidad, conjuntos de datos a los que denomina *data descriptors*. En noviembre de 2015 cuenta con 114 artículos. Cobra un APC de 1.000 US\$.

En sus criterios de publicación, destaca que en la información relativa a las personas debe evaluarse tanto su reutilización como su importancia y su privacidad. Asimismo, entre sus recomendaciones a los revisores prioriza los conjuntos de datos que puedan tener interés para un número mayor de científicos.

Se centra en las ciencias de la vida, pero acepta materias relacionadas:

“Estamos enfocados a conjuntos de datos de ciencias de la vida, y a comunidades científicas biomédicas y ambientales. Se ruega que los científicos con conjuntos de datos de otros campos se pongan en contacto con nosotros previamente al envío. Estamos dispuestos a considerar conjuntos de datos de ciencias sociales, particularmente los que puedan ser de utilidad para un análisis integrador a través de las fronteras disciplinarias tradicionales de la biomedicina, las ciencias ambientales y las ciencias sociales”

Finalmente, como la propia editorial comunica, ya se encuentra indizada en *PubMed*, y recientemente ha sido admitida en *Medline*.

Earth system science data

<http://www.earth-system-science-data.net>

ESSD apareció en 2009, publicada por *Copernicus Publications*. Está recogida en múltiples índices, entre los que destaca *Science Citation Index Expanded* y *Scopus*. Actualmente no cobra por publicar.

Destacan dos características que las diferencian de otras revistas:

- Mantiene secciones de artículos de longitud normal, así



Figura 2. *Scientific data*
<http://www.nature.com/sdata>

como comunicaciones breves (por ejemplo, sobre las adiciones a los conjuntos de datos) y comentarios, así como artículos de revisión y temas especiales. Es decir, es una revista de datos pero publica también artículos tradicionales.

- Tiene un sistema de revisión en dos etapas: en la primera, los artículos que han pasado una rápida revisión por pares se publican inmediatamente en el apartado *ESSD discussions (Essdd)*, donde se inicia un debate público interactivo, con presencia de revisores anónimos o públicos que hacen comentarios adicionales, y en donde los propios autores responden. Una vez se ha superado esta etapa, se publican en *ESSD*. Para garantizar la primacía de los derechos de descubrimiento para los autores y para proporcionar un registro permanente, ambos apartados tienen un ISSN propio y permiten su citabilidad.

Atomic data and nuclear data tables

<http://www.journals.elsevier.com/atomic-data-and-nuclear-data-tables>

Es un ejemplo de revista creada con anterioridad a la aparición de internet, que ya tenía los datos como núcleo de su publicación. Aparece en 1969. Su editorial es *Elsevier* y está indizada en la *WoS* y en *Scopus*. Publica tablas y gráficos de utilidad general para los investigadores tanto en las áreas básicas como aplicadas.

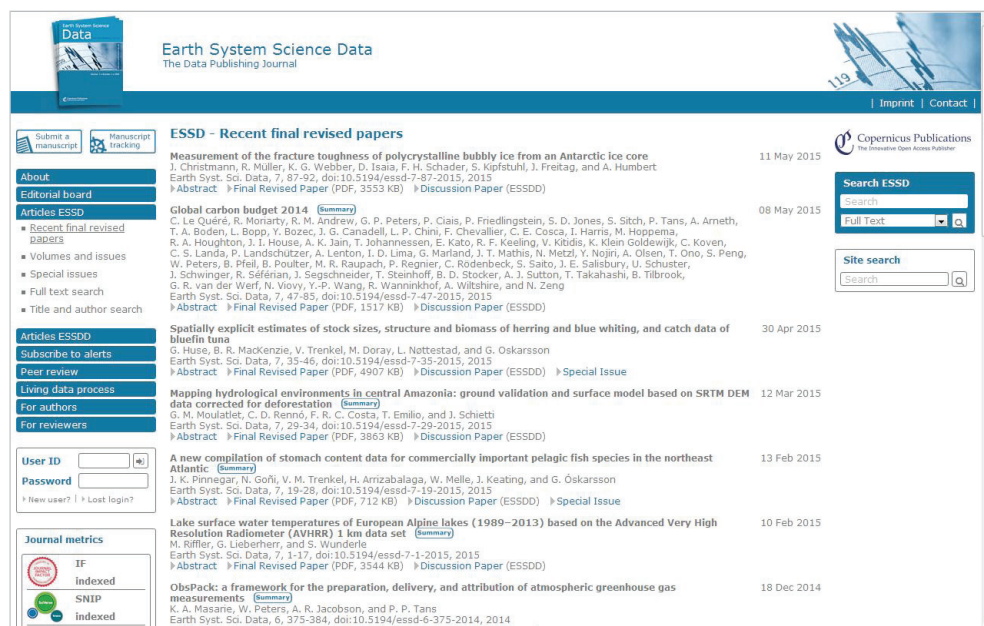


Figura 3. Vista de los artículos que ya han pasado los dos etapas de revisión en *Earth system science data*

Son interesantes las indicaciones a los autores en varios aspectos, como por ejemplo los éticos y de conflictos de intereses.

El acceso es por suscripción (vía *ScienceDirect*, por ser del grupo editorial *Elsevier*). Tiene la opción *open choice* para liberar artículos bajo pago de 2.200 US\$.

Data science journal

<http://datascience.codata.org>

Creada en 2002 ha sido relanzada en 2014, bajo la edición de *Ubiquity Press* y la dirección de Sarah Callaghan. Es multidisciplinar y se publica en acceso abierto. Publica artículos de investigación, ensayos y artículos de datos. Experimenta con otras formas de publicación, como los *overlay articles*, que refunden información de varias fuentes.

Los conjuntos de datos tienen que depositarse en un repositorio con un enlace persistente para asegurar su preservación. En su licencia explícita que el autor conserva los derechos para redistribuir su trabajo o liberarlo en un repositorio, incluso antes de ser publicado.

5. Requerimientos de un data journal

Los criterios básicos en los que coinciden la mayoría de las editoriales que publican revistas de datos son:

- publicar *datasets* con un alto potencial de reutilización;
- ser plenamente *open access* para que los *datasets* puedan ser citados y reutilizados;
- Poseer *copyright* libre: los *datasets* deben poderse copiar, distribuir, transmitir y reutilizar, pero proporcionando reconocimiento a los autores;
- asegurar el proceso *peer review*;
- exigir la preservación en repositorios públicos de acceso abierto y de larga duración;
- asegurar la identificación mediante un *digital object identifier* (doi);
- frecuencia de publicación *as soon as possible* (asap)

Estas características se definen en las políticas y guías de publicación de algunas de las editoriales que publican mayor número de revistas de datos, entre las que se encuentra *Pensoft Publishers* con revistas como

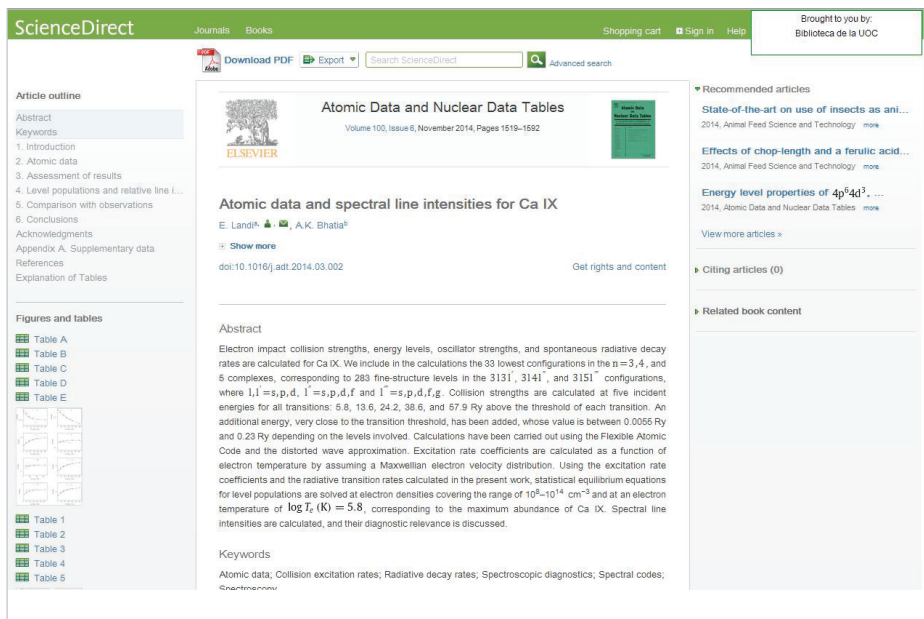


Figura 4. Atomic data and nuclear data tables

Zookeys, *PhytoKeys*, *Biodiversity data journal*, y *Nature conservation*. Algunos extractos literales de sus indicaciones:

“*Pensoft* admite varios métodos para la publicación de datos, tales como conjuntos de datos descargables complementarios a un artículo de investigación, o bien alojados en repositorios de datos externos y vinculados.

[...]

Publicando los datos, los autores y las instituciones quedan acreditados por su trabajo de crearlos y mantenerlos, quedando registrada su prioridad y autoría. Los datos se indexan, y así pueden ser descubiertos y citados”

<http://www.pensoft.net/page.php?P=23>

Copernicus Publications en su revista de datos *Earth system science data* (*ESSD*) propone las siguientes pautas:



Figura 5. Data science journal

“*Copyright* libre: cualquier persona puede copiar, distribuir, transmitir y adaptar los conjuntos de datos, siempre y cuando acredite los autores originales (equivalente a la licencia *Creative Commons Attribution*)”

[...]

Disponibilidad a largo plazo: el repositorio tiene que cumplir con los más altos estándares para garantizar la disponibilidad a largo plazo de los conjuntos de datos” http://www.earth-system-science-data.net/about/publication_policy.html

BioMed Central Publishers recoge:

“La revista *GigaScience* tiene como objetivo aumentar la transparencia y la reproducibilidad de la investigación, haciendo hincapié en la calidad y utilidad de los datos en contraste con las evaluaciones subjetivas de impacto inmediato. Para habilitar el acceso y los análisis futuros, se requiere que los metadatos y el código fuente estén disponibles públicamente. Proporcionamos una base de datos y un repositorio en la nube que pueden albergar datos, información complementaria y herramientas”

<http://www.gigasciencejournal.com/about>

Finalmente, *Ubiquity Press* en su *Journal of open research software (JORS)* expresa:

“*JORS* presenta *Software Metapapers* que describe software de investigación con alto potencial de reutilización. Trabajamos con varios repositorios especializados e institucionales para garantizar que el software esté archivado profesionalmente, se conserve, y esté disponible en abierto. Igualmente importante es que el software y los artículos sean citables y se pueda rastrear su reutilización”.

<http://openresearchsoftware.metajnl.com/about>

6. Conclusiones

Una de las principales ventajas de las revistas de datos es que se centran en los datos de investigación abiertos, lo cual fomenta la reutilización ya que con ellos se describe la metodología de obtención y análisis de los datos –con una profundidad similar a la que se exige en las patentes-. En muchos artículos convencionales, la descripción de la recogida de datos se reduce para ceder el protagonismo a los resultados y su discusión.

Los artículos de datos hacen salir a la luz conjuntos de datos que por distintos motivos no fueron publicados en artículos de investigación: datos que subyacen en los informes, en las tesis doctorales u otros trabajos académicos, los que no obtuvieron los resultados esperados, datos antiguos que permiten una comparación o una evolución posterior... De esta manera, al tiempo que se recuperan datos, se da publicidad a las metodologías utilizadas, lo que evita redundar en la misma investigación.

Es decir, una primera conclusión es que se refuerza la visibilidad de los aspectos metodológicos que van asociados a la investigación y se hacen visibles datos que nunca verían la luz en el sistema tradicional de la comunicación científica.

A diferencia de un repositorio, la revista de datos asegura la realización de un proceso de revisión explícito, descrito, centrado en la coherencia de los datos y el método de recolección (**Mayernick et al.**, 2015). Aunque los repositorios acreditan contar con una alta calidad como depósitos (por ej., estabilidad de los URI), de momento no cuentan con revisión por pares (**Whyte et al.**, 2013).

La revisión del material suplementario es muy discutida por algunos autores (**Parsons; Duerr; Minster**, 2010) y ha sido descartada por revistas como *Nature*, pues consume recursos humanos. Para suplirla, *ESSD* ha optado por un sistema mixto: una evaluación rápida interna que después se completa con la efectuada en abierto por la comunidad.

“ Un *data paper* es un artículo de revista que describe datos en vez de informar de una investigación o de sus conclusiones (Chavan; Penev, 2013)

Steinhart (2013) abordó la relación entre los repositorios institucionales y los editores. Hace notar que bibliotecas – generalmente responsables de los repositorios- y editores tienen objetivos comunes en comunicación científica: almacenar los conjuntos de datos de forma inteligible para que sean reutilizables.

Sin embargo se introduce en el sistema editorial un riesgo evidente: engrosar la burbuja de publicaciones académicas. Algunas revistas se presentan abiertamente como una forma de aumentar el curriculum del investigador con un artículo más. Es, por tanto, una capa más que si se extrema podría llevar a acrecentar la práctica de “fragmentar artificialmente una investigación en unidades publicables mínimas”, de forma que aumentaría la inflación académica (**Baiget; Torres-Salinas**, 2013). Así, los artículos de comunicaciones breves se justifican en ámbitos con mucha competencia permitiendo apuntar las primeras conclusiones de una investigación y asignar la autoría. Un *data paper* cubriría el método de recogida de los datos, pero faltaría aún el artículo completo con la necesaria discusión y contextualización de resultados. Aun así, este riesgo no puede desmerecer el conjunto de ventajas de las revistas de datos. No hay que olvidar que inicialmente el movimiento *open access* y los repositorios también debieron lidiar con acusaciones y percepciones respecto a la baja calidad (**Warr**, 2003). Se está trabajando en sistemas de certificación de los repositorios (**Callaghan et al.**, 2014).

Nuestro estudio muestra que editoriales como *Nature* o *Wiley* apuestan por dicho tipo de revista como producto comercial solvente, al tiempo que cada disciplina percibe ya su utilidad.

Por tanto, la primera conclusión con respecto al nuevo modelo de revista es que hace más confiable la comunicación científica. Pero hay que vigilar que el sistema no se desvirtúe, tanto porque se utilice como forma de hinchar el curriculum de los investigadores como porque ha de mantener la calidad en la revisión por pares. Finalmente, apuntamos una

responsabilidad añadida a dichas revistas, a sus revisores y gestores. ¿Quién se responsabiliza de los aspectos éticos y legales de los datos incluidos? Alguien debe garantizar que se preserve el anonimato de los individuos, sus datos personales y otros aspectos. Algunas revistas lo recogen de forma clara, otras no.

Como conclusión final cabe insistir en que se ha descrito cómo se ha incrementado en los últimos años este tipo de revistas. Desde 2013 el número de revistas ha crecido con decisión, pasando de una cierta excepcionalidad a una mayor presencia en varias disciplinas. Dos hechos pueden explicar parte del fenómeno. En primer lugar, en diciembre de 2013 el programa H2020 de la Unión Europea puso en marcha un proyecto piloto sobre datos que obliga a sus participantes a hacerlos públicos; y en marzo de 2015 entró en el Congreso de los EUA la ley *Fair access to science and technology research act* (Fastr).

Es aún temprano, consideramos, para comprobar si las revistas de datos van a encontrar su espacio en el ecosistema de la ciencia. Parece oportuna una investigación más profunda que considere, por disciplinas, cuántas de estas revistas sobreviven y cómo se relacionan con el resto de elementos del sistema que está haciendo públicos los datos, incidiendo especialmente en aspectos como la interoperabilidad y la preservación.

Nota

Como material complementario a este artículo, en: <http://dx.doi.org/10.6084/m9.figshare.1549666>

figura un “dataset” con las revistas especializadas en datos tratadas en este artículo y otras que eventualmente aparezcan en el futuro.

Agradecimiento

Este trabajo se ha desarrollado gracias a la financiación parcial del Plan Nacional de I+D+i del MINECO OPENDATAS-CIENCIA CS02012-39632-C02-02.

7. Bibliografía

Baiget, Tomás; Torres-Salinas, Daniel (2013). *Informe APEI sobre publicación en revistas científicas*. Gijón: Asociación Profesional de Especialistas en Información. <http://www.apei.es/wp-content/uploads/2013/11/InformeAPEI-Publicacionescientificas.pdf>

Bosch, Elisabeth; Mas, Francesc; Moyano, Albert (1999). *Documentació química*. Barcelona: Edicions de la Universitat de Barcelona (Textos docents: 114). ISBN: 84 8338 153 2

Callaghan, Sarah; Tedds, Jonathan; Lawrence, Rebecca; Murphy, Fiona; Roberts, Timothy; Wilcox, Will (2014). “Cross-linking between journal publications and data repositories: A selection of examples”. *International journal of digital curation*, v. 9, n. 1, pp. 164-175 <http://www.ijdc.net/index.php/ijdc/article/view/9.1.164>

Chavan, Vishwas; Penev, Lyubomir (2011). “The data paper: a mechanism to incentivize data publishing in biodiversity science”. *BMC Bioinformatics*, n. 12 (Suppl. 15, S2). <http://dx.doi.org/10.1186/1471-2105-12-S15-S2>

Cronin, Blaise (2013). *Scientific journals. Challenges and trends*. Conferencia en Universitat de Barcelona, 18 de abril. <http://bd.ub.edu/grups/ccd/sites/bd.ub.edu.grups.ccd/files/Jornada%20Revistes%202013/BlaiseCronin.pdf>

Halbert, Martin (2013). *Research data management: Prospects for research data management*. Washington DC: Council on Library and Information Resources. ISBN: 978 1 932326 47 5 <http://www.clir.org/pubs/reports/pub160>

Hildyard, Christopher J.; Whitaker, Benjamin J. (1996). “Chemical publishing on the internet: Electronic journals - Who needs them?” *Online information 96. Procs of the intl online information meeting*, Londres, 3-5 diciembre. <http://files.eric.ed.gov/fulltext/ED411825.pdf>

Kratz, John; Strasser, Carly (2014). “Data publication consensus and controversies”. *F1000Research*, v. 3, p. 94. <http://f1000research.com/articles/3-94/v1>

Lawrence, Bryan; Jones, Catherine; Matthews, Brian; Pepler, Sam; Callaghan, Sarah (2011). “Citation and peer review of data: Moving towards formal data publication”. *Intl journal of digital curation*, v. 6, n. 2, pp. 4-37. <http://dx.doi.org/10.2218/ijdc.v6i2.205>

López-Borrull, Alexandre (2003). *Síntesi i caracterització estructural de complexos d'Ag(I) amb lligands de tipus tiourea o tioamida. Estudi de fonts d'informació electrònica especialitzades en química*. Tesis doctoral, Universitat Autònoma de Barcelona. <http://www.tdx.cat/handle/10803/3158;jsessionid=7AD923DD1C780AEE1F451431552642EA.tdx2>

López-Borrull, Alexandre (2012). “Física vs química: dos modelos de publicación científica”. *El profesional de la información*, v. 21, n. 2, pp. 167-172. <http://hdl.handle.net/10760/18736> <http://dx.doi.org/10.3145/epi.2012.mar.07>

López-Borrull, Alexandre (2014). “Retos de la comunicación científica”. *Anuario ThinkEPI*, v. 8, pp. 203-207.

Mayernik, Matthew S.; Callaghan, Sarah; Leigh, Roland; Tedds, Jonathan; Worley, Steven (2015). “Peer review of datasets: When, why, and how”. *Bulletin of the American Meteorological Society*, v. 96, n. 2, pp. 191-201. http://www.researchgate.net/publication/267392725_Peer_Review_of_Datasets_When_Why_and_How <http://dx.doi.org/10.1175/BAMS-D-13-00083.1>

Newman, Paul; Corke, Peter (2009). “Editorial: Data papers - peer reviewed publication of high quality data sets”. *The intl journal of robotics research*, v. 28, n. 5, pp. 587. <http://dx.doi.org/10.1177/0278364909104283>

Organisation for Economic Co-operation and Development (OCDE) (2006). *Recommendation of the Council concerning access to research data from public funding*. 14 Dec. 2006 - C(2006)184. <http://acts.oecd.org/Instruments/ShowInstrumentView.aspx?InstrumentID=159>

Ortoll, Eva; Canals, Agustí; García-Alsina, Montserrat; Co-

barsí-Morales, Josep (2014). "Principales parámetros para el estudio de la colaboración científica en *Big science*". *Revista española de documentación científica*, v. 37, n. 4. <http://dx.doi.org/10.3989/redc.2014.4.1142>

Pampel, Heinz; Dallmaier-Tiessen, Suenje (2014). "Open research data: From vision to practice". En: Bartling, Sönke; Friesike, Sascha (eds.). *Opening science. The evolving guide on how the internet is changing research, collaboration and scholarly publishing*. Heidelberg: Springer, pp. 213-224. <http://doi.org/10.1007/978-3-319-00026-8>

Parsons, Mark A.; Duerr, Ruth; Minster, Jean-Bernard (2010). "Data citation and peer review". *Eos, Transactions American Geophysical Union*, v. 91, n. 34, pp. 297-298. <http://doi.org/10.1029/2010EO340001>

Peset, Fernanda; Ferrer-Sapena, Antonia; Subirats-Coll, Imma (2011). "Open data y linked open data: su impacto en el área de bibliotecas y documentación". *El profesional de la información*, v. 20, n. 2, pp. 165-173. <http://recyt.fecyt.es/index.php/EPI/article/view/epi.2011.mar.06> <http://dx.doi.org/10.3145/epi.2011.mar.06>

Peset, Fernanda; González, Luis-Millán (2016, en prensa). *Ciencia abierta y gestión de datos de investigación*. Gijón: TREA.

Rodríguez-Yunta, Luis; Giménez-Toledo, Elea (2013). "Fusión, coedición o reestructuración de revistas científicas en humanidades y ciencias sociales". *El profesional de la información*, enero-febrero, v. 22, n. 1, pp. 36-45. <http://dx.doi.org/10.3145/epi.2013.ene.05>

Silva, Lyz (2014). "PLOS' New data policy: Public access to data". *PLoS Blogs*, 24 February. <http://blogs.plos.org/everyone/2014/02/24/plos-new-data-policy-public-access-data-2>

Steinhart, Gail (2013). "Partnerships between institutional repositories, domain repositories and publishers". *Bulletin*

of the Association for Information Science and Technology, August/September. http://www.asis.org/Bulletin/Aug-13/AugSep13_Steinhart.html

Vernooy-Gerritsen, Marjan (ed.) (2009). *Enhanced publications: Linking publications and research data in digital repositories*. Amsterdam: Amsterdam University Press, ISBN: 978 90 8964 188 5. <http://dare.uva.nl/cgi/arno/show.cgi?fid=150723>

Villarroya, Anna; Claudio-González, Melba; Abadal, Ernest; Melero, Remedios (2012). "Modelos de negocio de las editoriales de revistas científicas: implicaciones para el acceso abierto". *El profesional de la información*, v. 21, n. 2, pp. 129-135. <http://accesoabierto.net/sites/accesoabierto.net/files/Villarroya-Claudio-Abadal-Melero-EPI.pdf> <http://dx.doi.org/10.3145/epi.2012.mar.02>

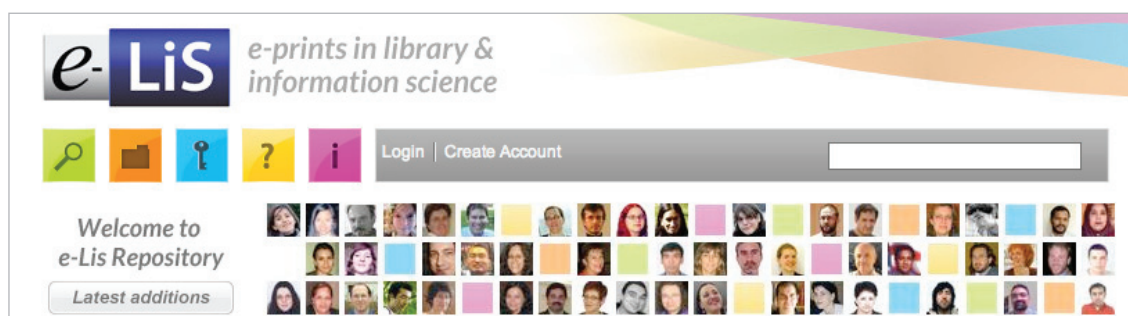
Ware, Mark; Mabe, Michael (2015). *The STM report: an overview of scientific and scholarly journal publishing*. 4th ed. The Hague: International Association of Scientific, Technical and Medical Publishers. http://www.stm-assoc.org/2015_02_20_STM_Report_2015.pdf

Warr, Wendy A. (2003). "Evaluation of an experimental chemistry preprint server". *Journal of chemical information and computer sciences*, v. 43, n. 2, pp. 362-373. <http://dx.doi.org/10.1021/ci025627a>

Wellen, Richard (2013). "Open access, megajournals, and moocs: On the political economy of academic unbundling". *SAGE Open*, October, v. 3, n. 4. <http://sgo.sagepub.com/content/3/4/2158244013507271>

Whyte, Angus; Callaghan, Sarah; Tedds, Jonathan; Mayer-nik, Matthew S. (2013). "Perspectives on the role of trustworthy repository standards in data journal publication". *Iassist conf*, Cologne, May 2013. <http://www.iassistdata.org/conferences/2013/presentation/3657>

Da visibilidad a tu trabajo depositándolo en e-LIS, el mayor repositorio internacional sobre biblioteconomía, documentación y comunicación



<http://eprints.rclis.org>