



CAPÍTULO V

**Acceso abierto a la
información científica**

Acceso abierto a la información científica

La llegada de la tecnología digital y de las redes de comunicación ha propiciado una evidente mejora de los canales de acceso a la información y la concreción de nuevos modelos de comunicación tanto personal como científica.

Una de las propuestas que más incidencia están teniendo en todo este proceso son las iniciativas del movimiento para el acceso abierto a la información científica, que se ha visto concretado en la creación de más de 1500 repositorios y más de 4500 revistas OA⁵ en todo el mundo durante los últimos años. En la actualidad casi todas las universidades y centros de investigación disponen de un repositorio institucional como reflejo y garante de la memoria académica de su entidad.

La importancia de éste fenómeno viene determinada por la implicación no sólo de los gestores de la información – bibliotecas, centros de documentación-, sino también por primera vez en la historia de investigadores, entidades

⁵ Registry of Open Access Repositories (ROAR) <http://roar.eprints.org/> y
DOAJ Directory of Open Access Journals (DOAJ) <http://www.doaj.org/>

financieras, e incluso estados que promueven y favorecen iniciativas de esta índole.

La Comunicación científica y el acceso abierto

La comunicación científica, es decir la manera cómo los investigadores de cualquier campo utilizan y difunden información a través de canales formales e informales, es esencial a la naturaleza y práctica de la ciencia, y está presente en todas las etapas del proceso de investigación. La investigación tiene unas características propias como son el razonamiento lógico, la universalidad y la interdisciplinariedad. La universalidad de la ciencia deriva de la Declaración Universal de Derechos Humanos, y plantea que la investigación científica como un proceso de construcción colectiva cuyo patrimonio es universal.

Durante muchos siglos este proceso ha permanecido en unos cánones estancos, que se han visto modificados por la llegada de las tecnologías de la información. La expansión de Internet ha puesto en manos de los propios científicos la

posibilidad de reapropiarse de parte de las condiciones de producción, difusión y uso de los productos de la ciencia (Russell, 2001) Sin embargo esta situación aparentemente ideal se encuentra con obstáculos de todo tipo como son barreras de acceso de tipo económico, institucional y legal que limitan el libre desarrollo de estos procesos. Esta *brecha digital* se une a las múltiples brechas que a diferencia de quienes estimaban que el acceso a la sociedad de la información iba a paliar las diferencias entre países pobres y ricos por las posibilidades de disponer de información para el desarrollo, las está ahondándolas aún más.

Hay que decir que esta brecha no existe únicamente entre países ricos y pobres sino que también se extiende a las sociedades desarrolladas entre individuos con diferentes niveles de acceso y capacidad para utilizar la información.

En el sistema de suscripción tradicional, el poder disponer de publicaciones relevantes y tener un buen índice de impacto depende entre otras cuestiones de la disponibilidad económica de nuestra organización para acceder a la literatura científica de calidad; por lo tanto la barrera económica es una barrera de acceso-impacto (Suber,

2009), pues limitar por cuestiones legales o económicas el acceso a la misma, está limitando a su vez las posibilidades de impacto de esa información.

La aparición de la World Wide Web (www) proporcionó a los científicos la potenciación en la información que habían estado buscando, lo que ha supuesto:

- La disponibilidad de un amplio espectro de fuentes de información... PERO por otro lado también la dificultad de acceso a las mismas ya que la información se gestiona por entidades privadas con intereses económicos particulares.
- Los canales individuales son cada vez más eficaces y efectivos... PERO tenemos dificultad de saber elegir los más adecuados.
- Mucha información disponible,... PERO muy dispersa.
- Mucha información disponible,... PERO sólo una pequeña parte se puede definir como académica.

Se estima que los motores de búsqueda solamente indexan un 25 % de todos los contenidos de la red, a esta

porción de información que es accesible, pero no está indexada por los buscadores se le denomina “*Internet Invisible*” El problema es importante, no sólo por el volumen estimado, más aún lo es por la calidad de estos recursos que aparecen en la cara oculta y que precisamente son los más relevantes para investigación. Por ello es importante generar servicios de información especializada que organicen y faciliten el acceso efectivo as la información.

Otro de los aspectos que han contribuido a la renovación del modelo de comunicación científica ha sido la edición electrónica, pues sería impensable la concreción del modelo de acceso abierto en un sistema de publicación impresa.

La edición electrónica tiene unas ventajas notables que hoy día les hace imprescindibles para cualquier centro de investigación como son: la rápida distribución, la facilidad de almacenamiento y recuperación, su uso desde cualquier lugar y en cualquier momento. Pero también han cambiado las condiciones como las empresas multinacionales comercializan las publicaciones electrónicas; frente a la suscripción individual la revista electrónica se vende en paquetes por las que las entidades pagan una licencia de

acceso, frente al precio unitario de cada revista con la edición electrónica las entidades tienen que negociar un precio que será variable en función de múltiples factores como son las características de centro, número de usuarios potenciales, niveles de acceso, capacidad de comercialización, etc. Por otra parte no se compra el producto en sí mismo, sino que se paga por una licencia de acceso. Pero también la llegada de la edición electrónica ha supuesto el incremento de los presupuestos de las entidades disponían para su adquisición en detrimento de otros productos o servicios como es la compra de monografías. Según un estudio desarrollado por la ARL entre los años 1986-2002 la subida acumulativa ha sido de un 227% con un incremento medio anual de un 15%; esto ha desembocado en la imposibilidad de las instituciones para mantener sus suscripciones, y ha llevado a la cancelación de determinados paquetes de revistas, y por lo tanto a una pérdida de acceso a la información publicada que afecta a la capacidad investigadora de la entidad.

Esta situación denominada *crisis de la comunicación científica* ha sido el factor desencadenante del movimiento para el Acceso Abierto, en un modelo no sostenible, en el que se plantea una situación irregular donde el monopolio

comercial lo tienen unas pocas empresas comerciales que concentran la gestión de las más prestigiosas revistas, y en la que los organismos públicos dedican una buena parte de sus recursos a financiar investigaciones, cuyos resultados terminan siendo comercializados por empresas privadas que obtienen unos amplios márgenes de beneficios por la venta de licencias de acceso a esta información que ha sido generada por entidades públicas. Es decir, las instituciones pagan esa investigación dos veces, una cuando financian a las investigaciones, y otra cuando compran esa información.

Pero por sí solo esto no podría explicar el fenómeno, en cuyo fondo subyace el desarrollo y aprovechamiento de las redes de información que posibilitan compartir el conocimiento de manera inmediata, a través de un nuevo concepto como es el de la *web social*, que conlleva un nuevo modelo de interactividad y de cooperación.

El acceso abierto

Open Access es el término utilizado para describir el libre acceso a la literatura científica en línea, la definición que se utiliza habitualmente para conceptualizar el término es

la emitida en la Declaración de Budapest sobre el Acceso Abierto (BOAI), en el año 2001(BOAI, 2001).

Por “acceso abierto” a esta literatura queremos decir su disponibilidad gratuita en Internet público, permitiendo a cualquier usuario leer, descargar, copiar, distribuir, imprimir, buscar o usarlos con cualquier propósito legal, sin ninguna barrera financiera, legal o técnica, fuera de las que son inseparables de las que implica acceder a Internet mismo. La única limitación en cuanto a reproducción y distribución y el único rol del copyright en este dominio, deberá ser dar a los autores el control sobre la integridad de sus trabajos y el derecho de ser adecuadamente reconocidos y citados.

De ella podemos extraer unas cuestiones esenciales para caracterizar que se entiende por Acceso Abierto:

- Las obras en acceso abierto están libremente disponibles para todos.
- Se refiere fundamentalmente a documentos que están en línea, es decir digitales y accesibles a través de Internet.

- Se trata exclusivamente de obras científicas.
- Los artículos de revista son el principal documento objeto del movimiento.
- Los autores no perciben dinero de las entidades comerciales por su esfuerzo.
- Hay un amplio rango de modalidades de uso para los documentos, pero esencialmente se garantizan la autoría y la integridad de la obra.

El primer repositorio de acceso abierto lo crea en 1991, Paul Ginsparg, físico del Laboratorio Nacional de Los Álamos (Estados Unidos), que puso en marcha un servidor gratuito llamado ArXiv⁶ al que los científicos podían enviar sus artículos, a veces incluso antes de su publicación en una revista científica.

El éxito del proyecto fue inmediato, en la actualidad ArXiv⁷ contiene cerca de 600.000 textos de Física,

⁶ ArXiv <http://arxiv.org/> [consultado el 22 de diciembre de 2009]

⁷ Originalmente fue alojada en el Laboratorio Nacional de los Álamos (at xxx.lanl.gov, de ahí su antiguo nombre, LANL preprints archive) y ahora

Matemáticas, Informática, Biología y Estadística, y recibe unas 200.000 consultas diarias. ArXiv retoma la tradición que tenían los investigadores de los campos científicos de intercambiar *preprints* antes de publicar un artículo para sondear la opinión de sus colegas; lo que hace Ginsparg es proporcionar un depósito para almacenar los *e-prints* – nombre que reciben estos documentos en los repositorios-, aunque de una manera más o menos consciente lo que está llevando a cabo es una renovación del modelo de comunicación científica.

A partir de entonces otras iniciativas que estaban trabajando en este sentido como *CogPrints*⁸ en el campo de la Psicología y desarrollado por Stevan Harnad, o *RePec*⁹ del ámbito de la Economía y dirigido por Thomas Kritchel, junto a ArcXiv, toman conciencia de la situación y deciden reunirse en octubre de 1999 en Santa Fé (Nuevo México, Estados Unidos) para empezar a trabajar de manera coordinada y establecer un acuerdo minimalista que será disponer de un formato de metadatos y un sistema

está alojado y administrado por la Universidad Cornell, con mirrors en todo el mundo.

⁸ CogPrints <http://cogprints.org/> [consultado el 22 de diciembre de 2009]

⁹ RePec <http://repec.org/> [consultado el 22 de diciembre de 2009]

identificación común que permitiera el establecimiento de un protocolo de interoperatividad entre los diferentes proyectos (Arencibia Jorge, Santillán Aldana, & Subirats Coll, 2005) que se verá concretado en la publicación en el año 2001 de *The Open Archives Initiative Protocol for Metadata Harvesting* (OAI-PMH)¹⁰ protocolo que posibilita la interoperabilidad de repositorios.

Entre 2001 y 2003 tiene lugar una serie de declaraciones institucionales que se conocen como la triple B (BBB) que perfilan las bases ideológicas de la iniciativa:

- 2001. Declaración de Budapest.
- 2003 Declaración de Bethesda.
- 2003 Declaración de Berlín.

En ellas se establece los aspectos conceptuales del acceso abierto, que es, que significa y que objetivos tiene; y se

¹⁰ *Open Archives Initiative Protocol for Metadata Harvesting*
<http://www.openarchives.org/pmh/> [consultado el 23 de diciembre de 2009]

establecen algunas cuestiones relativas a la interoperabilidad de los *archivos abiertos*, así como las vías para conseguir el Acceso Abierto:

- Ruta dorada: La publicación en revistas OA.
- Ruta verde: Autoarchivo de artículos en Repositorios.

También en el año 2001 tiene lugar la iniciativa de PloS cuando un grupo numeroso de científicos piden que todo artículo sea accesible públicamente en línea seis meses después de su publicación, quienes firman esta carta son alrededor de treinta mil científicos, que se comprometen a no publicar en las revistas que no pongan en práctica esta idea. Aunque no consiguen su propósito, sí consiguen organizarse para publicar sus propias revistas en Acceso abierto.

En 2004 los miembros que forman parte de los comités científicos se retiran de un importante grupo editorial, ante las duras condiciones que éste impone a los autores (*Journal Declaration of Independence, 2009*). Ante esta situación, Elsevier modifica los acuerdos de derechos de autor de sus productos. NATURE también modifica las condiciones

permitiendo el *autoarchivo* pasados 6 meses de la publicación. Desde entonces Biomed Central ha incluido el 10% de sus títulos en OA en el Journal Citation Report y PubMed posibilita que los documentos estén en libre acceso si las entidades o los autores pagan un plus por ello.

Desde entonces el movimiento para el Acceso Abierto ha desarrollado en muy poco tiempo un alto nivel de madurez en la renovación de los modelos de comunicación científica y ha planteado una alternativa sostenible a los modelos existentes.

Objetivos del Acceso abierto

El objetivo del movimiento para el acceso abierto es mejorar el sistema de comunicación científica mediante la optimización del acceso y la maximización del impacto de la investigación a través del *autoarchivo* (Harnad, Carr, Brody, & Oppenheim, 2003). Parece evidente que si una publicación es accesible libremente lógicamente será más veces leída y consecuentemente más citada. En este sentido se han desarrollado estudios que relacionan el número de descargas

con el número de citas que recibe un documento, los datos arrojados estiman que este comportamiento difiere de unas asignaturas a otras, el rango de citas recibidas oscilaría entre 0,54 a 0,11 (M. Kurtz & Brody, 2006). Otro dato que corrobora lo expuesto es que los artículos de astrofísica depositados en ArXiv reciben el triple de citas que los de la misma revista accesibles sólo a través de suscripción.

La Ruta verde: Autoarchivo

Se denomina auto-archivo al proceso de depósito de los documentos en un repositorio por parte de sus autores para facilitar su acceso libre y gratuito en Internet. El proceso de auto-archivo conlleva el registro previo del autor en el repositorio a través del cual acepta las condiciones y se le otorgará un espacio propio desde el cual puede incorporar sus documentos.

Los documentos depositados en el repositorio se denominan *eprints* (electronic print), aquellos que han pasado un proceso de revisión por un comité científico se llaman *preprints* y los

que no lo han hecho *postprint* Una vez depositados no son públicos inmediatamente, sino que deberán ser aprobados por los *editores* del proyecto-personas que revisa en un archivo OA-

Debe quedar claro que un repositorio no es un medio de publicación, simplemente es un medio de comunicación pública de un trabajo de investigación. Los *editores* simplemente aprueban los documentos en función de una política determinada por el repositorio (es decir, que se adecuen a los contenidos, revisión de metadatos, un cierto nivel de *cientificidad*).

Cuando hacemos una consulta al repositorio lo habitual es que incluya una nota diciendo si se trata de un preprint o un postrint para que el usuario sepa a que atenerse. Además Acceso Abierto no es incompatible con publicar en una revista, se pueden hacer las dos cosas, es decir publicar en los canales normales y depositar el documento en un repositorio para que este libremente accesible y visdible a la comunidad científica.

Existen diversas acciones para promover el auto-archivo:

- Mandato de auto-archivo¹¹ (obligación a depositar a los propios autores).
- Promoción, marketing y apoyo (visibilidad).
- Incentivos a la obtención de becas, proyectos si tienen documentos depositados.
- Colección inicial de documentos de los que generalmente la entidad impulsora del proyecto posee los derechos de autor (tesis doctorales, servicio de publicaciones).

Una de las cuestiones que habitualmente se plantean los autores que desean auto-archivar sus documentos en un repositorio es si están autorizados a hacerlo, respecto a ello el autor deberá saber si la revista en la que ha publicado un artículo retiene el derecho de difusión pública de su obra; para ello puede consultar dos bases de datos que le ayudarán a conocer en que condiciones han publicado y si tienen

¹¹ ROARMAP (Registry of Open Access Repository Material Archiving Policies) <http://www.eprints.org/openaccess/policysignup/> [consultado el 23 de diciembre de 2009]

derecho a depositar el preprint, el postprint o ambos a través de un sistema sencillo de códigos de colores en:

- *ROME0/Sherpa*¹² para revistas internacionales.
- *Dulcinea*¹³ par revistas españolas.

El autor es quien deposita los documentos y es, en última instancia, el responsable de la vulneración de los derechos de autor así como el garante de la titularidad e integridad. En el caso de los artículos de revista el autor no percibe derechos económicos por la explotación, con lo cual la cuestión más discutible es la cesión del derecho de comunicación pública de la obra, por lo tanto un autor, cuando deposita un documento, debe conocer qué derechos tiene cedidos sobre la misma.

Aunque el concepto de divulgación resulta útil para tutelar el interés de los autores a dar a conocer su obra, su

¹² SHERPA/RoMEO - Publisher copyright policies & self-archiving
<http://www.sherpa.ac.uk/romeo/> [consultado el 23 de diciembre de 2009]

¹³ DULCINEA - Derechos de copyright y las condiciones de auto-archivo de revistas científicas españolas
<http://www.accesoabierto.net/dulcinea/> [consultado el 23 de diciembre de 2009]

coexistencia con el de publicación no es fácil, en el caso de España el concepto de divulgación engloba el de publicación, aunque poner una obra a disposición del público en un sitio web supone una divulgación, pero no se puede considerar una publicación en su sentido jurídico.

La Ruta dorada: Revisitas OA

La vía dorada también ha manifestado un importante avance, con el impulso de políticas institucionales y estatales para la creación de revistas en Acceso Abierto, actualmente el Directorio de Revistas en Acceso Abierto (DOAJ)¹⁴ recoge 4500, y el Proyecto Scielo contiene más de 300 revistas.

Las revistas en libre acceso son un modelo más sostenible que hace posible que los editores puedan continuar asumiendo la publicación, recuperando costes y obteniendo ciertos beneficios (Alonso Arévalo, Subirats i Coll, & Martínez Conde, 2008). Además hay que señalar que las revistas en libre acceso tienen unas características generales

¹⁴ Directory of Open Access Journal <http://www.doaj.org/> [consultado el 23 de diciembre de 2009]

que comparten con las revistas digitales e impresas (mecanismos de control de calidad, revisión por pares, coste de publicación) y otras específicas de este modelo relativas a sus niveles de accesibilidad (son libremente accesibles y retienen sus derecho a través de Licencias CC¹⁵).

Las revistas OA provienen de iniciativas públicas a través de instituciones, universidades, asociaciones profesionales y centros de investigación; o del sector privado con modelos que intentan compatibilizar el acceso abierto con el modelo de negocio denominado “el autor paga” (término confuso, ya que quien realmente paga es la institución a la que pertenece el autor) o a través de embargos temporales (6 meses a 2 años).

La mayor dificultad de esta ruta es diseñar planes de negocio como una alternativa viable y sostenible. Una creencia común es considerar que las revistas OA son gratuitas, que de hecho lo son para el usuario final, pero que conllevan unos costes de producción para las entidades o editores que la publican, que oscila entre 300 y 1500 euros por artículo:

¹⁵ Creative Commons España <http://es.creativecommons.org/>
[consultado el 24 de diciembre de 2009]

- *BioMed*, 750 libras
- *Blacwell on line open* 1250 libras
- *Plos Medicine*, 1500 dólares
- *Springer Open Choice*, 300 dólares

En cuanto a su origen tenemos diferentes tipos de revistas OA:

- **Revistas OA** nacidas con vocación de acceso abierto (BioMed Central)
- **Editores convencionales**, que ofrecen también revistas OA, tales como Springer Open Choice Program
- **Editores no convencionales**, revistas editadas por asociaciones, departamentos e investigadores.

Los autores

El autor es el agente principal y máximo beneficiario del acceso abierto en términos de visibilidad, y en la posibilidad de poder disponer de la literatura científica sin ningún tipo de restricción. El autor se beneficia de ser más leído, más citado y consecuentemente obtendrá mayor reconocimientos de su comunidad científica, lo que le va a proporcionar más ayudas, méritos y financiación para futuros proyectos. Sin embargo, esto contrasta con la baja proporción de autores que tienen depositado al menos un documento en un archivo abierto (10%), las causas que ellos argumentan (Carr & Swan, 2007):

- Desconocimiento. No conocen el acceso abierto, ni sus beneficios.
- No sabe cómo autoarchivar o no tienen tiempo de hacerlo.
- Temen infringir la legislación.
- Falta de conciencia de la sostenibilidad de la ciencia.

El principal argumento para convencer a los autores es el de la alta visibilidad los resultados y esfuerzos de su trabajo (estadísticas proporcionadas por los repositorios). Ya que sus

trabajos van a estar mundialmente accesibles a través de los recolectores globales y sus ideas universalmente difundidas a través de los motores de búsqueda

Repositorios

El término repositorio procede del latín *Repositorium* que significa “armario”. El diccionario de la Real Academia Española lo define como el “lugar donde se guarda algo” aunque el término tal y como lo conocemos actualmente proviene del inglés *Repository* que designa a un depósito o archivo centralizado donde se almacena y mantiene información digital.

Un repositorio es una colección de objetos digitales basada en la Web, de material académico producido por los miembros de una institución (o varias), con una política definida que tiene unas características específicas:

- Auto-archivo. El contenido es depositado por el creador, propietario de los derechos.
- Interoperabilidad.

- Libre accesibilidad.
- Preservación a largo plazo frente a la baja perdurabilidad de internet.

Los repositorios pueden ser de diferentes tipos en función de que recojan los documentos de una entidad, una disciplina o un lugar geográfico:

- Archivos institucionales (Gredos, Eprints Complutense)
- Archivos disciplinares o temáticos (ArXiv, E-LIS, RePec..)
- Archivos centralizados (recolectores) (OAIster, Scirus)

Independientemente de la plataforma utilizada un archivo abierto tiene dos componentes comunes básicos:

- Mecanismos de depósito (Autoarchivo, Ayudas y tutoriales).

- Mecanismo de acceso y recuperación (Buscador, Browser).

OAI-PMH solamente es una interfaz sumamente sencilla para acceder a la información bibliográfica disponible en un archivo o repositorio. Por lo tanto cualquiera puede realizar una implementación del mismo para poner a disposición de la comunidad Internet los datos que hasta ahora estaban escondidos en bases de datos o catálogos. Para facilitar esta tarea han aparecido una serie de programas que permiten a cualquier institución (universidad o centro de investigación) crear su propio archivo al tiempo que hacerlo compatible con OAI-PMH (Silló, T., 2005).

- Eprints University of Southampton.

<http://www.eprints.org>

El más popular de todos. Es un software desarrollado en el seno del Open Citation Project dirigido por Stevan Harnad en la Universidad de Southampton (UK). Eprints puede

funcionar en cualquier ordenador con sistema operativo Linux.

➤ Dspace. <http://www.dspace.org>

Está desarrollado por la empresa HP y las bibliotecas del MIT. También es un software con las fuentes disponibles públicamente (open source). Muy extendido en España.

➤ CDSware (CERN). <http://cdsware.cern.ch>

➤ Fedora Cornell University <http://fedoraproject.org/>

Un repositorio institucional es el reconocimiento de la vida intelectual y académica de una entidad, que permite que ésta sea representada documentalmente y difundida en formato digital.

La implementación de un repositorio tiene un bajo coste y unos beneficios inmediatos para la institución, cuyo objetivo es poner a disposición de la sociedad y del resto de

investigadores toda la producción científica generada por la institución y dar visibilidad a la entidad y a sus autores.

Los documentos que se almacenan son fundamentalmente artículos de revistas, pero cada vez es más frecuente encontrar también documentos institucionales, patrimoniales y objetos de aprendizaje.

Recolectores

Frente a la estructura distribuida de Internet donde existen millones de servidores en todo el mundo que recogen información que está hipervinculada formando un gran documento mundial que es la www.

La estructura abierta consiste en un concepto de interoperatividad de sistemas a través de un protocolo de comunicaciones común OAI-PMH que facilita que estos sistemas individuales puedan funcionar como un recurso común recopilando los metadatos libremente expuestos por los proveedores de servicios (repositorios).

Un recolector de recursos digitales es una herramienta para acceder a los documentos a texto completo contenidos en bibliotecas digitales o en repositorios que cumplan con el protocolo *Open Archives Initiative – Protocol for Metadata Harvesting* (OAI-PMH).

El recolector actualiza de forma periódica los nuevos documentos depositados en los repositorios que compila, ampliando así su base de datos de forma continua.

Existen recolectores en ámbitos sectoriales, internacionales, regionales y nacionales:

- **DRIVER** (Europeo) <http://www.driver-repository.eu/>
- **E-ciencia** (regional)
<http://www.madrimasd.org/informacionIDI/e-ciencia/>
- **Google Scholar** (Internacional)
<http://scholar.google.es/>
- **Hispana: Recolector de Recursos Digitales MCU**
(España) <http://roai.mcu.es/es/inicio/inicio.cmd>
- **OAIster** (Internacional) <http://www.oaister.org>

- **OpenDOAR** (Internacional)
<http://www.opendoar.org/>
- **RECERCAT** (Diposit de la Recerca de Catalunya)
<http://www.recercat.net/>
- **Recolecta** (España)
<http://www.recolecta.net/buscador/>
- **Scientific Commons** (Internacional)
<http://www.scientificcommons.org/>
- **Scirus** (Internacional) <http://www.scirus.com/>
- **SDL Search Digital** <http://drtc.isibang.ac.in/sdl/>

Los recolectores contribuyen a dar una alta visibilidad a la producción científica, ya que aunque hayamos depositado un documento en un archivo institucional o temático éstos van a ser recogidos por los diferentes recolectores y accesibles desde los mismos, con lo cual los canales de comunicación y descubrimiento de información científica se multiplican, por ello a veces es baladí la discusión que habitualmente se entabla sobre si

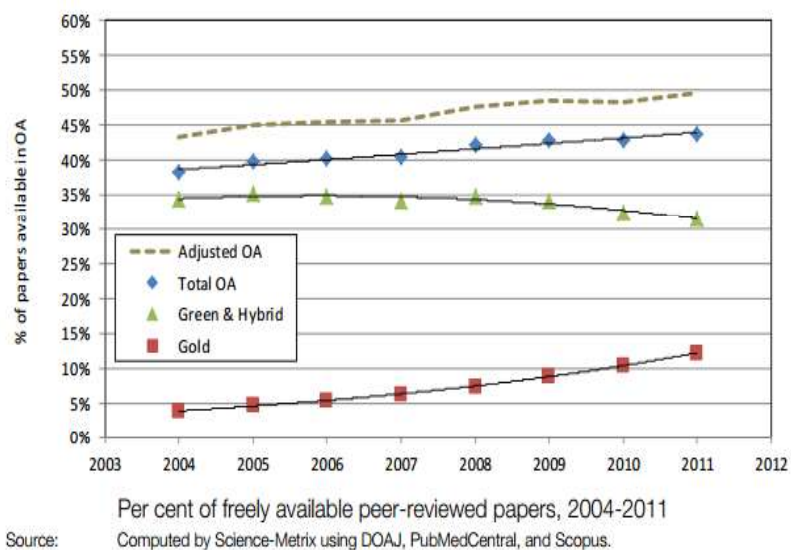
es mejor depositar en un repositorio institucional o temático, ya que lo importante y fundamental para que un documento de un investigador tenga visibilidad es que este libremente accesible en un repositorio OA.

La visibilidad no la va a proporcionar directamente el repositorio sino los recolectores que compilan esa información, y especialmente Google académico que indexa los textos completos hasta un máximo de 500 k.

Evolución del Acceso Abierto

El acceso abierto es un instrumento fundamental para mejorar la circulación del conocimiento y, con ella, la innovación. Son varios los países y distintas las disciplinas donde el número de artículos a los que puede accederse hoy gratuitamente supera el 50 %. El informe “Proportion of Open Access Peer-reviewed Papers at European and World Levels, 2004-2011”, un estudio europeo cuyo objetivo era cuantificar la proporción de los artículos revisados disponibles en acceso abierto tomando un área geográfica

correspondiente a Europa, Brasil, Canadá, Japón y Estados Unidos, estimó que en 2011 el 40% de los artículos *peer review* estaban en acceso abierto (Archambault et al., 2013; Gargiulo, 2013).



Proporción de artículos revisados por pares en acceso abierto en el ámbito europeo y mundial, 2004-2011

La libre disponibilidad de la mayor parte de los artículos es ya una realidad en el campo de la ciencia y la tecnología en general, en el de la investigación biomédica, en

el de la biología y en el de las matemáticas y la estadística. Por el contrario, los campos en los que el acceso abierto sigue estando más restringido son el de las ciencias sociales y humanidades y el de las ciencias aplicadas, la ingeniería y la tecnología.(Archambault et al., 2013) Pues existen distintas formas de comportamiento de los autores en diferentes disciplinas científicas, en Ciencias prácticamente el total de los autores deposita el *preprint* en Arxiv, el repositorio de la ciencia creado hace 20 años por Paul Ginsparg en los Álamos, hoy en Standford. Esto forma parte de la dinámica de trabajo de esta área que requiere de la rápida difusión de los resultados, ya que de otro modo otro investigador puede adelantarse en su publicación.

Por ello lo que encontramos en este repositorio son *preprints*, y no *postprint* como ocurre en las Ciencias Sociales y Humanidades, que se deposita el artículo una vez publicado.

El efecto del "Open Access" (OA) en la visibilidad o impacto de las publicaciones científicas es uno de los temas más importantes y controvertidos en el campo de la bibliometría y ciencias de la información. Durante los últimos

10 años se han publicado numerosos estudios empíricos que examinan esta cuestión utilizando diversas metodologías y puntos de vista. Las dos principales metodologías que se aplicaron en estos estudios relacionados con OA, tienen sus propias potencialidades y limitaciones. El primer método se basa en el análisis de la citación, y la segunda en el análisis de usos y descargas (Moed, 2012).

El debate en torno al impacto del Acceso Abierto OA se inició con la publicación por parte de Steve Lawrence del artículo, titulado “Free online availability substantially increases a paper's impact”(Lawrence, 2001) en la revista Nature, el análisis se hace en el campo de la informática. Si bien en este estudio no se refiere exactamente al "Acceso abierto", sino más en general, en el sentido de los artículos que están en línea en Internet. Desde un punto de vista metodológico, el debate se centra en los prejuicios, los grupos de control, muestreo y el grado en que las conclusiones de los estudios de caso puedan ser tomadas como generales.

En 2004 Stevan Harnad y Tim Brody dos de los padres del acceso abierto, afirmaron en un estudio de D-Lib Magazine titulado “Comparing the impact of open access

(OA) vs. non-OA articles in the sane journals” que los artículos de física que se depositaron como *preprints* en ArXiv, y que fueron publicados posteriormente en revistas revisadas por pares, generaron un impacto de citas de hasta el 400 por ciento superiores a los documentos en las mismas revistas que no habían sido publicados en ArXiv. Michael Kurtz (M. J. Kurtz et al., 2005) encontró en un estudio de astronomía que los autores publican sus mejores artículos libremente en la web, si bien es cierto que los artículos depositados como *preprints* son publicadas anteriormente y por lo tanto su periodo de citación es más amplio y por lo tanto su frecuencia de cita es más alta.

Para encontrar más casos similares, el proyecto OpenCit¹⁶ recoge desde 2007 toda la bibliografía publicada al respecto (The Open Citation, 2007). Entre ellos destacamos un estudio de 2010 llevado a cabo por Gargouri y otros (Gargouri et al., 2010) que realizaron un estudio en el que comprobaron que los artículos de aquellos autores que autoarchivaron sus artículos en la versión del editor en acceso abierto de las revistas de suscripción para que fuera accesible gratuitamente para todos en la web, concluyendo que estos se

¹⁶ Open Cit Project <http://opcit.eprints.org/>

citaron significativamente más que los artículos de la misma revista y años que no fueron autoarchivados.

Generalmente estos estudios evalúan la correlación entre las cifras de citación y de descarga. Se parte de la hipótesis de que los artículos con un menor número de descargas también acumulan menos citas. Los resultados difieren de una base de datos a otra. Sí existe una cierta correlación entre citas y descargas, pero éstas dependen de la disciplina o tipo de documento. En conclusión se puede decir que las descargas no son un sustituto universal de las cifras de citación, si bien el acceso abierto generalmente parece que conlleva una determinada ventaja competitiva (Nieder, Dalhaug, & Aandahl, 2013).

En el cuadro que vemos a continuación (Archambault et al., Op. Cit) se establece una puntuación superior a 1 que indica que las publicaciones en OA son más citadas que en el ámbito general, mientras que una puntuación por debajo de 1 significa que estas publicaciones se citan con menos frecuencia. Por ejemplo, los trabajos sobre Agricultura, Pesca y Silvicultura reciben más o menos el mismo nivel de la cita (1.06) en OA que si sólo estuvieran en cualquiera revista. Por

lo general prácticamente en todas las áreas de conocimiento se deriva una ventaja de citación OA, y paradójicamente, en muchos de los campos en los que la proporción de artículos OA es baja reciben una ventaja competitiva de cita considerable, como en la filosofía y teología (1,54), humanidades artes en general y ciencias sociales, los estudios de comunicación, ingeniería y artes visuales y escénicas.

Esto responde a una lógica, ya que cuanto más artículos esté en Acceso Abierto en una disciplina menos se beneficiaran que aquellas disciplinas que tienen todo en OA, ya que al haber menos, los pocos que hay se citan mucho más. Pues si el 100% de los artículos estuvieran en abierto el OA no tendría ninguna ventaja competitiva.

IMPACTO DE LAS PUBLICACIONES OA ENTRE 2008 Y 2011

Field	All Publications	Green & Hybrid	Gold	OA
Agriculture, Fisheries & Forestry	1.00	1.38	0.52	1.06
Biology	1.00	1.41	0.48	1.15
Biomedical Research	1.00	1.25	0.76	1.17
Built Environment & Design	1.00	1.33	n.c.	1.23
Chemistry	1.00	1.38	0.36	1.09
Clinical Medicine	1.00	1.56	0.54	1.34
Communication & Textual Studies	1.00	1.66	0.88	1.46
Earth & Environmental Sciences	1.00	1.30	0.82	1.22
Economics & Business	1.00	1.32	0.22	1.20
Enabling & Strategic Technologies	1.00	1.43	0.75	1.25
Engineering	1.00	1.55	0.55	1.46
General Arts, Humanities & Social Sciences	1.00	1.53	0.10	1.46
General Science & Technology	1.00	2.57	0.54	1.54
Historical Studies	1.00	1.54	0.51	1.29
Information & Communication Technologies	1.00	1.37	0.89	1.27
Mathematics & Statistics	1.00	1.22	0.71	1.16
Philosophy & Theology	1.00	1.56	n.c.	1.54
Physics & Astronomy	1.00	1.36	1.01	1.32
Psychology & Cognitive Sciences	1.00	1.37	0.69	1.29
Public Health & Health Services	1.00	1.36	0.72	1.19
Social Sciences	1.00	1.52	0.55	1.26
Visual & Performing Arts	1.00	1.93	0.11	1.40
Total	1.00	1.41	0.60	1.24

Source: Computed by Science-Metrix using DOAJ, PubMedCentral, and Scopus.

Hay que matizar que las estadísticas sobre las revistas de la ruta dorada (revista OA) requieren una cuidadosa interpretación, ya que muchas de ellas son más jóvenes y pequeñas, y estos factores tienen un efecto negativo sobre la tasa de citas y por lo tanto indican negativamente en los valores medidos de cita, pues los investigadores prefieren leer y citar con más frecuencia revistas establecidas, por lo que es un reto para una revista joven es tener autores y artículos de

alta calidad, pues se necesita tiempo para obtener una reputación y atraer a investigadores consagrados.

Si bien, la mayoría de estos estudios se hacen en relación a las citas recibidas en ISI Web of Science, donde las citas que se recogen son las de las propias revistas incluidas en este índice, y no las citas provenientes de otras revistas no incluidas en ISI. Por ello si consultamos diferentes índices como Scopus, que tiene una mayor representación de revistas de otros ámbitos geográficos o lingüísticos, o mismamente el índice Google Scholar que se beneficia de la sinergia del propio buscador, veremos que el impacto es mucho mayor, y un factor fundamental para ganar visibilidad es que este motor de búsqueda indexe una investigación como una publicación académica.

En 2016 SPARC publicó los resultados de un informe donde analizaba la ventaja competitiva y de cita del acceso abierto en base a todos los estudios sobre la cuestión recopilados por el proyecto OpCit, que desde hace años mantiene al día una lista de estudios sobre el impacto del acceso abierto. SPARC Europa completo esta lista con otros

estudios sobre si existe o no una cierta ventaja de cita para los artículos de acceso abierto a través de su página *The Open Access Citation Advantage Service* (SPARC, 2016) Haciendo un análisis de los 70 estudios recopilados por ambas páginas para determinar cuántos de ellos son concluyentes o no respecto a si existe una ventaja de citación de los artículos depositados en acceso abierto. Los resultados de los 70 estudios evaluados muestran las siguientes cifras:

- 46 estudios encontraron que el acceso abierto tenía una ventaja de cita
- 17 estudios no encontraron que existía una ventaja de cita
- 7 estudios no fueron concluyentes

En el contexto de la Unión Europea (UE), como parte de una gama más amplia de políticas para fomentar la circulación del conocimiento, la Comisión Europea (CE) hizo del Acceso Abierto una de las prioridades para el Espacio Europeo de Investigación (EEI). Recientemente, la Comisión Europea (CE) amplió y reforzó su política de Acceso Abierto exigiendo a cada beneficiario del Programa Marco de

Investigación e Innovación, Horizonte 2020 (H2020), publicar en abierto los resultados de sus investigaciones con una revisión por pares, además de los datos brutos de la investigación. Sin embargo, las publicaciones científicas y los datos de las investigaciones requieren un tratamiento diferente y su acceso está en diferentes fases de desarrollo, quedando el acceso directo a los datos en un segundo plano.

Importancia y futuro del Acceso Abierto y la Ciencia Abierta

La sociedad TIC propicia y requiere un diluvio universal de datos, procesarlos, entenderlos y transformarlos en decisiones de valor es el reto del análisis big data. Casi todos los analistas consideran “Big Data” como una de las tendencias de futuro que tendrán que tener en cuenta la mayoría de las empresas e instituciones (Reitano, 2013).

En el ámbito científico, la arraigada cultura de la participación en la comunidad de investigación es una de las fuerzas impulsoras detrás de los avances científicos. Son numerosos los científicos que han señalado la ironía de que justo en el momento histórico en el que disponemos de las más avanzadas tecnologías para permitir la disponibilidad a nivel mundial y el proceso de distribución de los datos científicos, la ampliación de la colaboración y acelerar el ritmo y la profundidad del descubrimiento; estamos ocupados en las restricciones de los datos y en evitar el uso de tecnologías avanzadas en favor del conocimiento.

Ciencia abierta implica una lógica centrada en la mejora de la eficiencia en la ciencia; aumentar la transparencia y la calidad en el proceso de validación de la investigación; acelerar la transferencia de conocimientos; aumento del descubrimiento del conocimiento frente a los desafíos globales de una manera más eficaz; y promover la participación de los ciudadanos en la ciencia y la investigación (Ball, 2015).

La investigación científica innovadora tiene un papel crucial en el tratamiento de los desafíos globales, que van desde la atención de la salud y el cambio climático por medio de las energías renovables y la gestión de los recursos naturales. La velocidad y la profundidad de esta investigación dependen de fomentar los intercambios de colaboración entre las diferentes comunidades y asegurar su más amplia difusión.

Por lo que no sólo cobra importancia la disponibilidad de los resultados de la investigación, si no que cada vez es más importante poder acceder a todos los datos relacionados con la investigación, desde presentaciones, fuentes de datos,

estadísticas, resultados fallidos, etc. a este movimiento se le ha denominado datos abiertos, y forma parte de otros movimientos open como Ciencia Abierta, Educación abierta, etc.

Los llamados “Big Data”, un término utilizado para referirse a la explosión de una ingente cantidad y diversidad de datos digitales de alta frecuencia, se están convirtiendo en un elemento esencial para la competencia, y en un futuro inmediato serán clave para el crecimiento de la productividad, la innovación y la posibilidad de generar el suficiente excedente para las sostenibilidad de la sociedad. Datos de registros de llamadas, transacciones de banca móvil, contenido generado por el usuario de internet, tales como blogs y tweets, búsquedas en línea, imágenes de satélite, etc. es información procesable que requiere el uso de técnicas computacionales para dar a conocer las tendencias y patrones dentro de y entre éstos extremadamente grandes conjuntos de datos socioeconómicos. Las bibliotecas de investigación juegan un papel vital en la gestión y curación de este tipo contenido, pero requieren de mecanismos de financiación adecuados (Erway & Rinehart, 2016).

Por *Open Research Data*, se entiende aquellos datos que están disponibles gratuitamente en Internet, y que permiten que cualquier usuario pueda descargar, copiar, analizar, reprocesar, reutilizar, cambiar de formato o utilizarlos para cualquier otro propósito sin barreras financieras, legales o técnicas otras que los fundamentales de la conexión a la propia Internet (Moore, 2014).

Por ello estos datos finales relacionados con la ciencia publicada deben ser expuestos de manera explícita en el dominio público. Open Data Research se centra en materiales no textuales, tales como mapas, genomas, compuestos químicos, fórmulas matemáticas y científicas, datos de práctica médica, biociencia y biodiversidad. A menudo surgen problemas debido a que estos son de gran valor comercial o integrados en obras de alto valor como informes técnicos.

El acceso o la reutilización de los datos es controlado por las organizaciones, tanto públicas como privadas. El control puede ser a través de restricciones de acceso, licencias, derechos de autor, patentes y tasas para el acceso o la reutilización.

Los defensores de datos abiertos argumentan que estas restricciones están en contra de la comunidad y que estos datos deben estar disponibles sin restricción ni costo. Además, es importante que los datos sean reutilizables sin necesidad de permiso adicional, aunque los tipos de uso (tales como la creación de obras derivadas) pueden ser controlados por licencia.

Los objetivos del movimiento Open Data son similares a los de otros los movimientos “abiertos”

- **Acceso abierto** se refiere a las publicaciones académicas libremente disponibles en internet. En algunos casos, estos artículos también incluyen conjuntos de datos abiertos.
- **Contenido Abierto** se refiere a hacer que los recursos destinados a un público humano (como en prosa, fotos, o videos) estén totalmente accesibles.
- **Open Notebook Science** se refiere a la aplicación del concepto de datos abiertos en el proceso científico, incluidos los experimentos fallidos y los datos en bruto.

- **Conocimiento Abierto.** La Open Knowledge Foundation aboga por la apertura de una serie de cuestiones, incluyendo pero no limitado a los de Open Data. (a) científicos, históricos, geográficos o de otro tipo (b) contenidos como música, películas y libros (c) Gobierno y otros datos administrativos. Open Data está incluida en el ámbito de la Definición de Conocimiento Abierto, que se alude en el Protocolo de “Science Commons” para la aplicación de datos de acceso abierto.
- **Open Source (Software)** se ocupa de las licencias bajo las cuales los programas de ordenador pueden ser distribuidos.

En 2010 se publicaron los Principios Panton Principles for Open Data in Science¹⁷. Estos principios redundaban en la idea de que “la ciencia se basa en la construcción, reutilización y apertura crítica de la publicación de los datos del conocimiento” en los que se proporcionaba una lista breve de los fundamentos a tener en cuenta cuando se

¹⁷ <http://pantonprinciples.org/>

exponen datos abiertos. Destinado a una amplia audiencia de investigadores, editores y bibliotecarios, el libro *Issues in Open Research Data* explora las implicaciones de los principios a través de una serie de puntos de vista sobre los datos de investigación abiertos.

La OCDE ha tomado la delantera en el desarrollo de principios y normas para facilitar el acceso a los datos de la investigación generada con fondos públicos. Las bases de datos se está convirtiendo rápidamente en una parte esencial de la infraestructura del sistema de la ciencia mundial. El proyecto internacional del genoma humano no es más que un buen ejemplo con éxito en todo el mundo de un esfuerzo a gran escala en la que se utiliza información accesible en abierto con una gran variedad de propósitos.

En el ámbito más práctico la apertura de datos requiere de inversiones que si bien en un principio pueden parecer excesivas a medio plazo se estimaran como bajas en relación con los beneficios que pueden propiciar.

Otros han planteado la necesidad de la apertura de datos por defecto, y como la utilización de estándares web comunes

y la comparabilidad de los datos hacen de ellos un recurso aún más valioso.

Un aspecto a tener en cuenta a tenor de lo expuesto es que las decisiones sobre qué datos se liberan primero y cómo se presentan los datos no son neutrales. El concepto de Open Data (OD) supone proporcionar información que puede ser utilizada por cualquier persona para cualquier propósito y sin coste alguno pero esto puede tomar muchas y variadas formas, por lo que para ser considerados libres deben presentarse en un formato electrónico normalizado a través de internet y con las licencias de uso adecuadas.

Esto hace que sea adaptable y fácil de analizar y combinar con otros datos, todo lo que a su vez proporciona una mayor utilidad y valor. Además, los datos públicamente disponibles proporciona información, conocimiento y sabiduría que tienen como potencial una serie de beneficios, sociales, económicos y medioambientales. Los principales usuarios de OD incluyen la comunidad académica, el sector privado y los gobiernos.

Conclusiones

El movimiento para el acceso abierto no supone en ningún momento un sistema de comunicación científica excluyente, más bien es simplemente una alternativa a otros ya existentes, y de alguna manera es previsible que coexistirá con mayor o menor aceptación con los sistemas tradicionales. Una cuestión básica es su contribución al desarrollo sostenible de la ciencia, aunque lo ideal sería encontrar un modelo de edición científica que no dependiera de las restricciones de acceso, o al menos un modelo de negocio que fuera compatible con el libre acceso a la literatura científica.

Lo que ha puesto de manifiesto el movimiento para el libre acceso a la información es la concienciación de la comunidad científica y de los gestores de información en las cuestiones relativas al acceso a la información de calidad, existiendo un interés directo de gobiernos, instituciones internacionales y organismos públicos en este fenómeno.

La tarea fundamental que se plantea el movimiento OA es persuadir a los autores para que autoarchiven sus trabajos de investigación para la libre difusión internacional, pues va a

contribuir a dar visibilidad a sus trabajos, va a recibir más citas de otros autores, y por lo tanto mayor reconocimiento, y más capacidades de financiación de futuros proyectos.

Después del software de código abierto, estándares abiertos, el acceso abierto y el contenido abierto; el fenómeno de los datos abiertos es la última frontera y quizás incluso en la actualidad uno de los recursos más estratégicos para gobiernos y organizaciones.

Bibliografía

- Alonso Arévalo, J., Subirats i Coll, I., & Martínez Conde, M. L. (2008). *Informe APEI sobre acceso abierto* (Vol. 2).
- Archambault, E., Amyot, D., Deschamps, P., Nicol, A., Rebut, L., & Roberge, G. (2013). *Proportion of Open Access Peer-Reviewed Papers at the European and World Levels—2004-2011*. Québec, Canada: Science-Metrix Inc.

- Arencibia Jorge, R., Santillán Aldana, J., & Subirats Coll, I. (2005). Iniciativas de acceso abierto en Ciencias de la Información y Documentación. Evolución y perspectivas de E-LIS. *Revista Española de Documentación Científica*, 28(2).
- Ball, D. (2015). *Open Science, open data, open access ... A UKeIG white paper*. London: UKeIG.
- Carr, L. H. S., & Swan, A. (2007). A Longitudinal Study of the Practice of Self-Archiving. Working Paper . (Unpublished). *ECS EPrints Repository*.
- Erway, R., & Rinehart, A. (2016). *If You Build It, Will They Fund? Making Research Data Management Sustainable* OCLC.
- Gargiulo, P. (2013). ¿En qué punto se encuentra el acceso abierto? ¿Cómo medirlo? *Block de bid*.
- Gargouri, Y., Hajjem, C., Larivière, V., Gingras, Y., Carr, L., Brody, T., & Harnad, S. (2010). Self-Selected or Mandated, Open Access Increases Citation Impact for Higher Quality Research. *PLoS ONE*, 5(10). doi: 10.1371/journal.pone.0013636
- Harnad, S., Carr, L., Brody, T., & Oppenheim, C. (2003). Mandated online RAE CVs linked to university eprint

- archives: Enhancing UK research impact and assessment. *Ariadne*(34).
- Kurtz, M., & Brody, T. (2006). The impact loss to authors and research. *Open Access: Key Strategic, Technical and Economic Aspects*(5).
- Kurtz, M. J., Eichhorn, G., Accomazzi, A., Grant, C., Demleitner, M., Henneken, E., & Murray, S. S. (2005). The effect of use and access on citations. *Information Processing & Management*(41), 1395–1402.
- Lawrence, S. (2001). Online or Invisible? *Nature*, 411(6837), 521.
- Moed, H. F. (2012). Does open access publishing increase citation or download rates? *Research Trends*(28).
- Moore, S. A. (2014). *Issues in Open Research Data* Ubiquity Press.
- Nieder, C., Dalhaug, A., & Aandahl, G. (2013). Correlation between article download and citation figures for highly accessed articles from five open access oncology journals. *SpringerPlus*, 2(1). doi: 10.1186/2193-1801-2-261

- Reitano, S. (2013). *The Benefits of Open Data : An Organizational Management Project for public organizations*: Royal Roads University.
- Russell, J. M. (2001). La comunicación científica a comienzos del siglo XXI. *Revista internacional de ciencias sociales*(168).
- SPARC. (2016). *The Open Access Citation Advantage Service. SPARC, 2016*. London: SPARC.
- Suber, P. (2009). Open Access in 2008. *Journal of Electronic Publishing, 12*(1).
- The Open Citation, P. (2007). The effect of open access and downloads ('hits') on citation impact: a bibliography of studies. *The Open Citation Project*.

