

Programa  
Información  
Científica



datos científicos abiertos

**CONICYT**

# Datos Científicos Abiertos

La Ciencia la hacemos entre todos

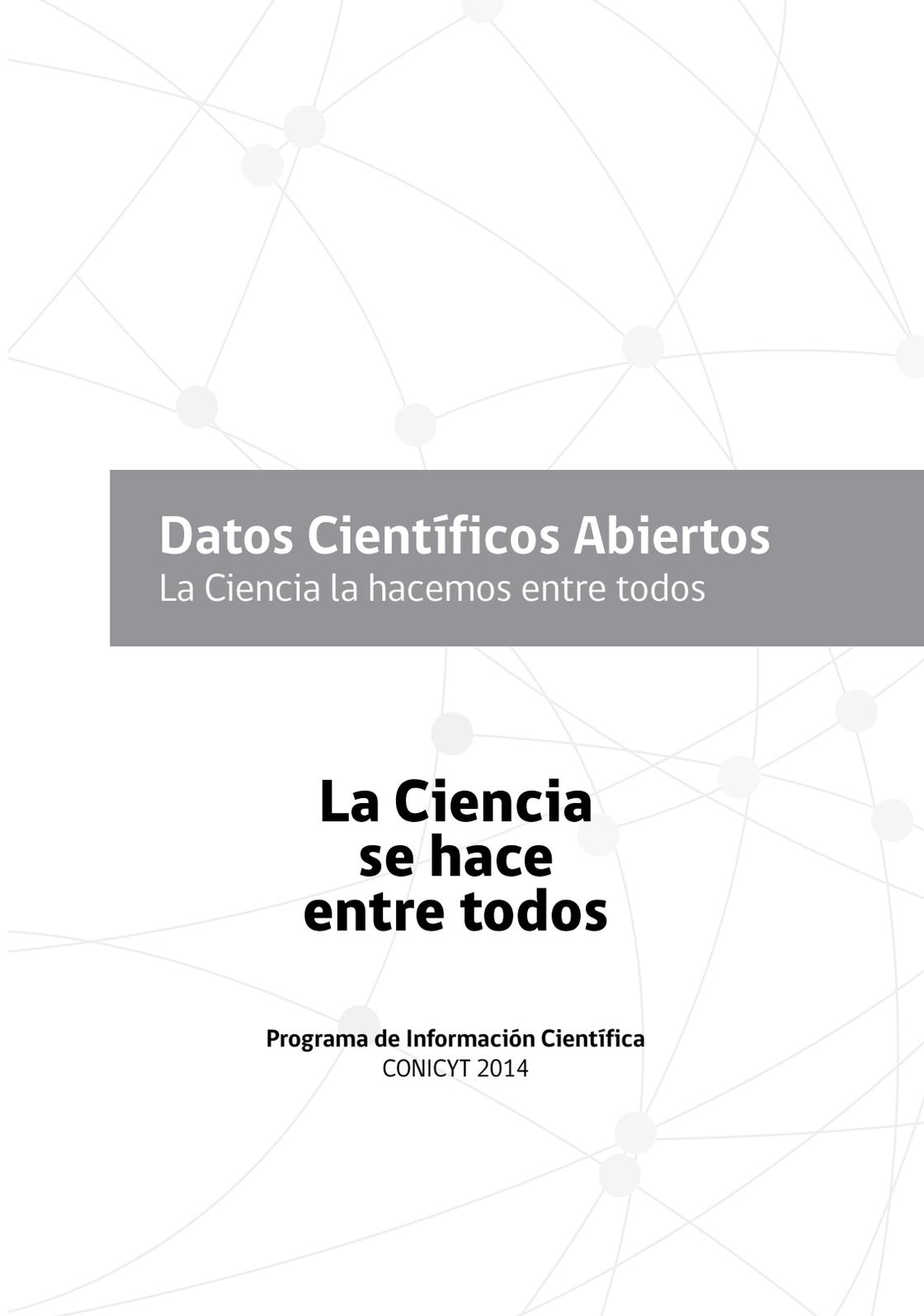
## La Ciencia se hace entre todos

Programa de Información Científica  
CONICYT 2014



**CONICYT**  
Ministerio de Educación

Gobierno de Chile



# Datos Científicos Abiertos

La Ciencia la hacemos entre todos

**La Ciencia  
se hace  
entre todos**

Programa de Información Científica  
CONICYT 2014

**Datos Científicos Abiertos**  
**Programa de Información Científica**  
**Comisión Nacional de Investigación Científica y Tecnológica**

**Equipo consultor**

Representante: Daniel Hernández H.  
Edición: Roxana Donoso P.  
Mónica Salvador T.  
Claudio Gutiérrez G.  
Daniel Hernández H.

Diseño y Diagramación: Salomé Alcaíno C.

**Otros**

Asesoría en aspectos legales de la ONG Derechos Digitales.

©2014 CONICYT

Este documento se encuentra bajo la Licencia Creative Commons  
Atribución-Compartir Igual 4.0 Internacional.



**Versión en línea**

<http://datoscientificos.cl>

# Índice

<b>1. ¿Qué es el acceso abierto a la información científica y a los datos de investigación? .....</b>	<b>8</b>
<b>2. ¿Por qué abrir los datos? .....</b>	<b>11</b>
<b>3. ¿Cómo abrir la información científica? .....</b>	<b>13</b>
3.1. Utilizar una licencia abierta .....	13
3.2. Asesorarse por especialistas .....	14
3.3. Crear repositorios institucionales .....	14
3.4. Utilizar repositorios interoperables .....	14
<b>4. ¿Cómo abrir los datos? .....</b>	<b>16</b>
4.1. Utilizar una licencia abierta .....	16
4.2. Observar cómo lo hacen los pares .....	16
4.3. Asesorarse con especialistas .....	17
4.4. Describir los datos .....	17
4.5. Asignar un identificador a los datos .....	18
4.6. Asegurar la perdurabilidad de los datos .....	19
4.7. Considerar el uso de bases de datos .....	19
4.8. Consultar el sitio web <a href="http://datoscientificos.cl">datoscientificos.cl</a> .....	20
<b>5. Aspectos legales y operativos .....</b>	<b>21</b>
<b>6. ¿Cómo avanzar hacia una política? .....</b>	<b>26</b>



# Introducción

El aumento de las capacidades de las tecnologías de la información permite que los investigadores puedan capturar, almacenar, explorar y manipular información y datos en volúmenes que antaño no habría sido posible. La relevancia de los datos es tal, que algunos creen que nos encontramos frente a un cuarto paradigma para la investigación científica<sup>1</sup>.

Históricamente la comunidad científica ha compartido los resultados de sus investigaciones a través de artículos, reportes, libros y otros recursos de información científica. En cambio, los datos que sustentan las hipótesis y modelos, solían recibir una atención menor.

Hoy, con el surgimiento de tecnología que permite capturar y procesar grandes volúmenes de datos, los datos de investigación cobran un rol fundamental en la ciencia y son necesarios para validar modelos y chequear hipótesis. Incluso, al encontrarse disponibles pueden ser reutilizados para proponer otras hipótesis o servir como herramienta en los procesos de evaluación.

La importancia del acceso abierto a los datos de investigación científica es ampliamente reconocida. Varios gobiernos y organismos internacionales han realizado investigaciones al respecto, que han motivado acuerdos para compartir y preservar los datos. El informe sobre Ciencia y Tecnología 2012 de la OECD<sup>2</sup> reconoce el modelo de

---

<sup>1</sup>El término "cuarto paradigma" lo acuñó Jim Gray para referirse a los cambios en las metodologías de investigación, producto de la disponibilidad de grandes volúmenes de datos. Las ideas de Gray se exponen en el libro *The Fourth Paradigm: Data Intensive Scientific Discovery*. Ver Tansley, S. y Tolle, K. M. (Eds.). (2009). *The Fourth Paradigm: Data-Intensive Scientific Discovery*. <http://research.microsoft.com/en-us/collaboration/fourthparadigm/>.

<sup>2</sup>OECD (2012). *OECD Science, Technology and Industry Outlook*. En <http://www.oecd.org/sti/oecd-science-technology-and-industry-outlook.htm>.

ciencia abierta como una de las principales estrategias para fortalecer el desarrollo de la ciencia en el mundo:

**“Tiene el potencial de mejorar la eficiencia y la calidad de la investigación mediante la reducción de los costos de recolección de datos, facilitando la explotación de datos inactivos o accesibles a bajo costo y mediante el aumento de las oportunidades de colaboración en la investigación y en la innovación”.**

La OECD ha elaborado directrices que orientan estrategias para el acceso a los datos de información científica generados con fondos públicos<sup>3</sup>. A ello se suman los acuerdos alcanzados por la Unión Europea<sup>4</sup>, el G8<sup>5</sup> y la UNESCO<sup>6</sup>.

Lo planteado por estos organismos internacionales ha motivado a la Comisión Nacional de Investigación Científica y Tecnológica, CONICYT, a invitar a la comunidad científica chilena para intercambiar ideas en torno al diseño e implementación de una política nacional para el acceso y la preservación de datos de investigación financiados con fondos públicos, en el entendido de que los científicos y las instituciones de investigación desempeñan un rol fundamental a la hora de garantizar la preservación y facilitar el acceso a los datos. Contar con su colaboración

---

<sup>3</sup>OECD (2007). Principles and Guidelines for Access to Research Data from Public Funding. <http://www.oecd.org/dataoecd/9/61/38500813.pdf>.

<sup>4</sup>Unión Europea (2012). Recomendación de la Comisión, de 17 de julio de 2012, relativa al acceso a la información científica y a su preservación. <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:32012H0417:ES:NOT>

<sup>5</sup>G8 (2013). G8 Open Data Charter and Technical Annex. <https://www.gov.uk/government/publications/open-data-charter/g8-open-data-charter-and-technical-annex>.

<sup>6</sup>UNESCO (2013). Directrices para políticas de desarrollo y promoción del acceso abierto. <http://www.unesco.org/ulis/cgi-bin/ulis.pl?catno=222536&gp=0>



no sólo aumentará el impacto de sus propias investigaciones, sino que también generará un activo que estará disponible para toda la comunidad científica, lo que potenciará a la investigación en general.

## 1. ¿Qué es el acceso abierto a la información científica y a los datos de investigación?

Existen diversas definiciones sobre el significado de “abierto”. En general, ellas acuerdan que “una pieza de contenido o un conjunto de datos es abierto cuando cualquiera es libre de usarlo, reutilizarlo y distribuirlo, sujeto solamente al requerimiento de atribución<sup>7</sup> o de compartir bajo las mismas reglas”<sup>8</sup>.

Para el caso de la información científica, los principales esfuerzos para lograr el acceso abierto a la literatura científica comenzaron en 2002 con la declaración de Budapest<sup>9</sup> de la Iniciativa OpenAccess, ratificada en las posteriores declaraciones de Bethesda<sup>10</sup> (2003) y Berlín<sup>11</sup> (2003). La Iniciativa OpenAccess, entiende que una publicación es de acceso abierto cuando cumple las siguientes dos condiciones<sup>12</sup>:

1. Los autores y los titulares de los derechos de propiedad intelectual otorgan a los usuarios un derecho libre, irrevocable, universal y perpetuo de acceso y licencia para copiar, utilizar, distribuir, transmitir y presentar el trabajo públicamente y hacer y distribuir obras derivadas, en cualquier soporte digital para cualquier finalidad

---

<sup>7</sup>Se denomina requerimiento de atribución al deber de indicar quién es el creador de una obra cuando ésta es redistribuida o citada.

<sup>8</sup>Esta es la definición abreviada que aparece en la portada del sitio <http://opendefinition.org>. La versión completa puede leerse en <http://opendefinition.org/od/>.

<sup>9</sup>Iniciativa de Budapest para el Acceso Abierto (2002). Ver en <http://www.budapestopenaccessinitiative.org/translations/spanish-translation>.

<sup>10</sup>Declaración de Bethesda sobre Publicación de Acceso Abierto (2003). Ver en <http://legacy.earlham.edu/~peters/fos/bethesda.htm>.

<sup>11</sup>Berlin Declaration on Open Access to Knowledge in the Sciences and Humanities. Ver en <http://openaccess.mpg.de/286432/Berlin-Declaration>.

<sup>12</sup>Estas condiciones corresponden a las planteadas en la declaración de 2003 en Bethesda.

responsable, sujeto a la apropiada atribución de la obra, así como el derecho de hacer una pequeña cantidad de copias impresas para su uso personal.

2. Una versión completa de la obra y todos los materiales suplementarios, incluyendo una copia de los permisos citados anteriormente, en un formato electrónico estándar apropiado se depositará de forma inmediata a la publicación inicial en al menos un repositorio en línea apoyado por una institución académica, una sociedad de intelectuales, una agencia gubernamental, o cualquier otra organización debidamente establecida que persiga facilitar el acceso abierto, la distribución sin restricciones, la interoperabilidad y el archivado a largo plazo.

Para los datos se suelen agregar otros requerimientos. Por ejemplo, que los formatos usados en su disciplina sean abiertos. El Open Data Handbook define los datos abiertos con tres principios<sup>13</sup>:

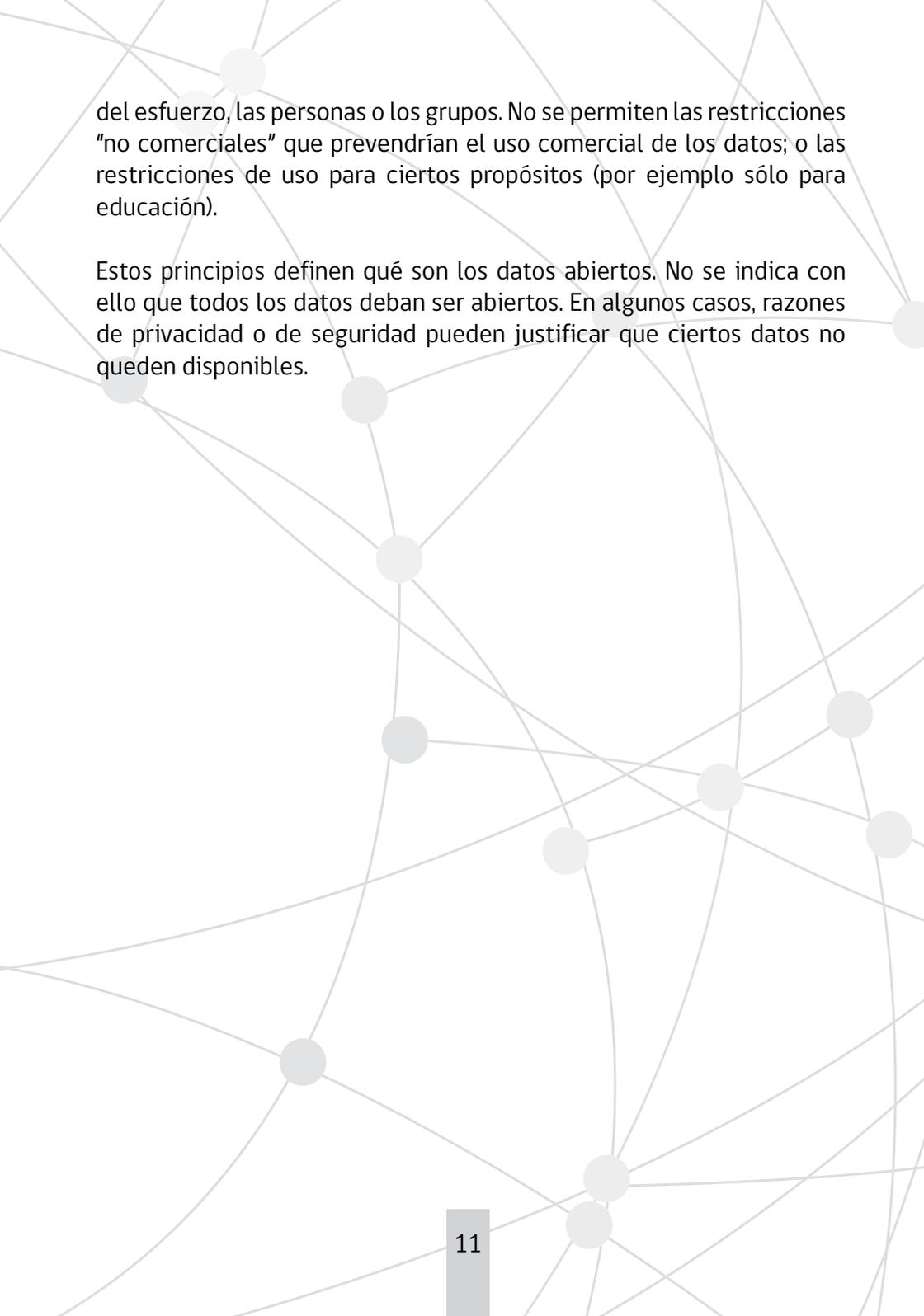
**Disponibilidad y acceso:** los datos deben estar disponibles como un todo y a un costo razonable de reproducción, preferiblemente descargándolos de Internet. Además, los datos debe estar disponibles en una forma conveniente y modificable.

**Reutilización y redistribución:** los datos deben ser provistos bajo términos que permitan reutilizarlos y redistribuirlos, e incluso integrarlos con otros conjuntos de datos.

**Participación universal:** todos deben poder utilizar, reutilizar y redistribuir los datos. No debe haber discriminación alguna en función

---

<sup>13</sup>Ver en <http://opendatahandbook.org/es/what-is-open-data/>.



del esfuerzo, las personas o los grupos. No se permiten las restricciones “no comerciales” que prevendrían el uso comercial de los datos; o las restricciones de uso para ciertos propósitos (por ejemplo sólo para educación).

Estos principios definen qué son los datos abiertos. No se indica con ello que todos los datos deban ser abiertos. En algunos casos, razones de privacidad o de seguridad pueden justificar que ciertos datos no queden disponibles.

## 2. ¿Por qué abrir los datos?

Los principios de los datos abiertos concitan un gran respaldo entre los gobiernos porque se advierte que es una manera de facilitar el intercambio de información entre las distintas organizaciones, junto con promover la participación ciudadana y el escrutinio público. Ese es la motivación principal para permitir el acceso a los datos gubernamentales y, por extensión, a los datos científicos financiados con fondos públicos.

En el caso de la ciencia, los datos abiertos poseen un valor adicional al permitir la reproducción de los resultados alcanzados en las investigaciones.

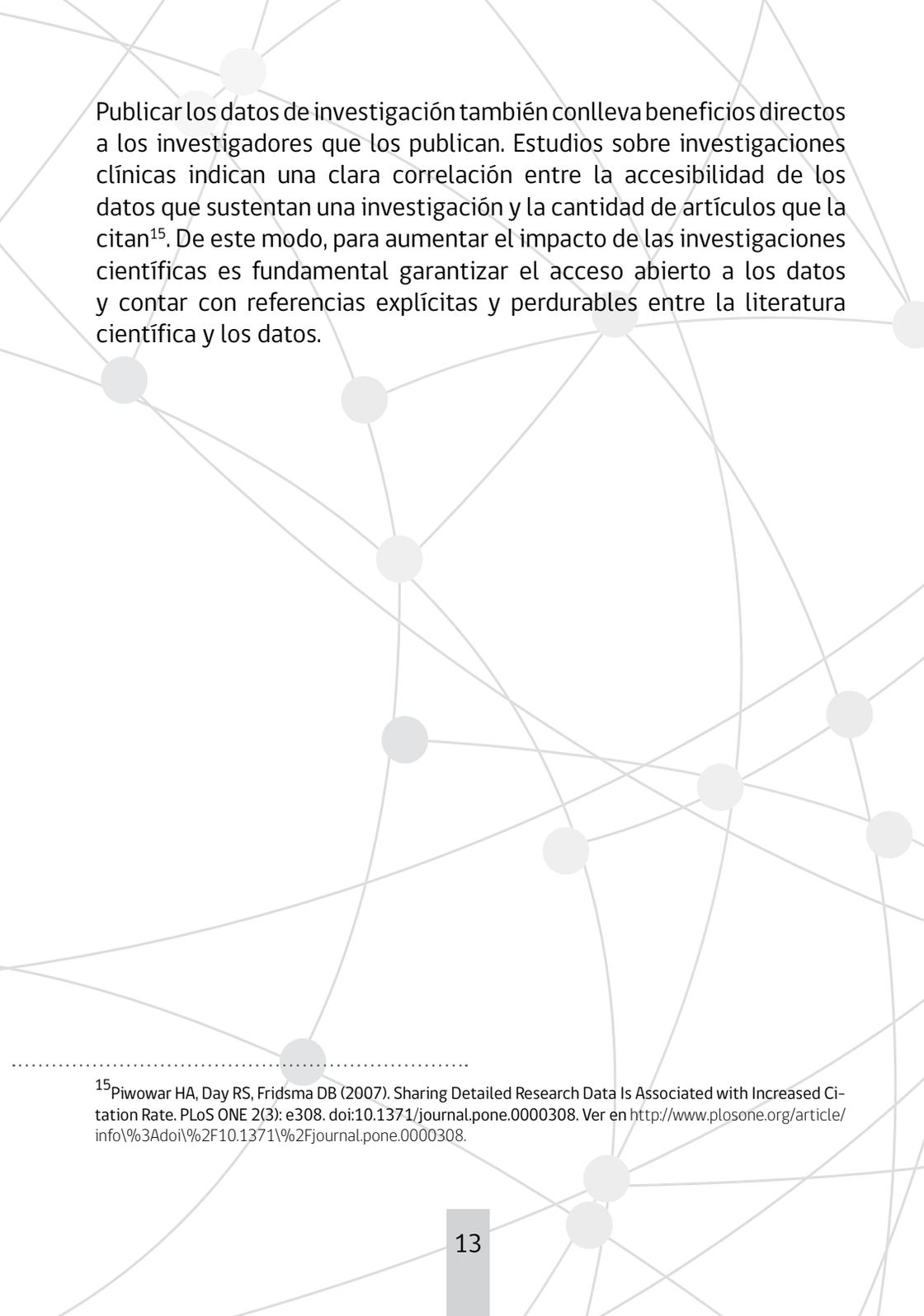
No se puede apoyar o refutar una hipótesis sin contar con el acceso a los datos utilizados en la investigación. La ciencia sin datos se convierte en un edificio con cimientos débiles. Así, resulta crucial para el trabajo científico garantizar el acceso y la preservación tanto de la literatura científica como de los datos que la sustentan.

En muchos casos la generación de datos es costosa. Por consiguiente, dejarlos disponibles para oportunidades adicionales de reutilización ahorra recursos y rentabiliza la primera inversión. Un buen ejemplo es el Proyecto Genoma Humano, cuya inversión de USD 3,8 billones entre 1990 y 2003 generó un retorno de USD 796 billones<sup>14</sup>. Esto equivale a un retorno de la inversión (ROI) de 141:1.

Compartir los datos también permite emprender investigaciones que en muchos casos no serían posibles, al reutilizar los datos para verificar otras hipótesis o probar otras metodologías. Es decir, los datos pueden llegar a tener un impacto no previsto en la investigación original.

---

<sup>14</sup>Simon Tripp y Martin Grueber (2011). Economic Impact of the Human Genome Project. Battelle Memorial Institute.



Publicar los datos de investigación también conlleva beneficios directos a los investigadores que los publican. Estudios sobre investigaciones clínicas indican una clara correlación entre la accesibilidad de los datos que sustentan una investigación y la cantidad de artículos que la citan<sup>15</sup>. De este modo, para aumentar el impacto de las investigaciones científicas es fundamental garantizar el acceso abierto a los datos y contar con referencias explícitas y perdurables entre la literatura científica y los datos.

---

<sup>15</sup>Piwovar HA, Day RS, Fridsma DB (2007). Sharing Detailed Research Data Is Associated with Increased Citation Rate. *PLoS ONE* 2(3): e308. doi:10.1371/journal.pone.0000308. Ver en <http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0000308>.

### 3. ¿Cómo abrir la información científica?

Existen algunas prácticas que se recomiendan para la publicación de información científica en la modalidad “acceso abierto”, que se describen a continuación.

#### 3.1. Utilizar una licencia abierta

Se recomienda usar alguna de las licencias Creative Commons con atribución<sup>16</sup> para la literatura científica<sup>17</sup>. Estas licencias permiten al titular de los derechos autorizar el uso de sus obras por terceras personas sin perder los derechos de atribución.

Estas licencias no sólo son documentos legales, sino que también incluyen una explicación simple de los derechos y obligaciones que confieren al usuario, además de una aplicación en lenguaje de máquina para identificar la obra en Internet mediante el empleo de opciones avanzadas de búsqueda en buscadores como Google y Yahoo!. Además, las licencias Creative Commons están disponibles en más de una treintena de idiomas, incluidos el castellano, francés, inglés, japonés, chino y ruso, entre otros. De esta forma, el idioma no es una traba para los usuarios de las obras licenciadas con licencias Creative Commons. Estas ventajas justifican la recomendación del uso de las licencias Creative Commons. Sin embargo, otras licencias abiertas también se podrían ser usar<sup>18</sup>.

---

<sup>16</sup>Las licencias Creative Commons con atribución poseen las siguientes variantes: CC-BY, CC-BY-SA, CC-BY-NC, CC-BY-NC-SA, CC-BY-ND o BY-NC-ND. Cada una de estas variantes se explica con mayor detalle en el sitio de Creative Commons <http://www.creativecommons.cl/tipos-de-licencias/>.

<sup>17</sup>En <http://creativecommons.org/science> se enumeran proyectos científicos que utilizan estas licencias.

<sup>18</sup>Un resumen de las licencias consideradas abiertas por la Open Definition se puede encontrar en <http://opendefinition.org/licenses/>.

### **3.2. Asesorarse por especialistas**

La publicación de información científica requiere del asesoramiento por especialistas de las ciencias de la información, que puedan dar directrices para la organización de la literatura científica publicada y la agregación de metadatos para describirla.

### **3.3. Crear repositorios institucionales**

Si bien, publicar la literatura científica en un sitio web personal es un primer paso para el acceso abierto, tiene el riesgo de que el sitio sea borrado por descuido o que la información deje de estar disponibles si algo le ocurre al dueño del sitio. Para que el acceso a la información perdure, se recomienda publicarla en lugares más estables como repositorios institucionales.

Además de que facilitan la perdurabilidad de la información publicada, los repositorios institucionales aumentan la visibilidad de las publicaciones y permiten establecer vínculos entre ellas a través de los metadatos.

### **3.4. Utilizar repositorios interoperables**

Las principales plataformas de repositorios abiertos permiten la búsqueda federada de documentos y la cosecha de registros entre ellos de manera automática. Para que ello sea posible se requieren estándares de interoperabilidad para los metadatos de los documentos y para la comunicación entre los repositorios.

El principal estándar de metadatos es Dublin Core<sup>19</sup>, que define un conjunto de atributos mínimos (título, resumen, autores, materias, fechas, etc.) que debiera describir cada documento de un repositorio.

---

<sup>19</sup>Ver en <http://dublincore.org/>.

La Open Archive Initiative (OAI) es la principal autoridad en la definición de estándares de interoperabilidad entre repositorios abiertos.

Entre los estándares definidos se encuentran un protocolo de cosecha de metadatos<sup>20</sup>, un estándar para describir e intercambiar descripciones de agregación entre componentes de recursos<sup>21</sup> y un marco para sincronización de archivos<sup>22</sup>.

Entre las plataformas de repositorios documentales existentes se recomienda la utilización de DSpace<sup>23</sup> por su robustez, flexibilidad y satisfactorio soporte de los estándares abiertos de metadatos y comunicación entre repositorios.

---

<sup>20</sup>OAI Protocol for Metadata Harvesting (OAI-PMH). Ver en <http://www.openarchives.org/pmh/>.

<sup>21</sup>OAI Object Reuse and Exchange (OAI-ORE). Ver en <http://www.openarchives.org/ore/>.

<sup>22</sup>OAI ResourceSync Framework Specification (OAI-RS). Ver en <http://www.openarchives.org/rs/toc>.

<sup>23</sup>Ver en <http://www.dspace.org>.

## 4. ¿Cómo abrir los datos?

Existen algunas prácticas que se recomiendan para la publicación de datos científicos abiertos, que se describen a continuación.

### 4.1. Utilizar una licencia abierta

Se recomienda usar la licencia Creative Commons Cero (CC0). CC0 es un modelo de licenciamiento que dedica e identifica una obra como parte del dominio público. Permite al titular de los derechos autorizar el uso de las obras (los datos, en este caso) por terceras personas.

En la actualidad, la Licencia CC0 se emplea en diversas iniciativas para facilitar el acceso. La British Library ha liberado cerca del 20 % de su información de catalogación con más de tres millones de registros.

La Organización Europea para la Investigación Nuclear, OpenLibrary y OpenJurist han usado CC0 para liberar datos bibliográficos, al igual que la Biblioteca, Archivo y Museo digital Europea; el repositorio de datos genéticos de Genomes Unzipped; el repositorio de fotografía e imágenes aéreas y satelitales de MichiganView y el portal de datos abiertos del gobierno de Holanda. Otros organismos internacionales consideran usar la Licencia CC0 como la herramienta legal para poner sus contenidos a disposición de los usuarios.

### 4.2. Observar cómo lo hacen los pares

Las ciencias son diversas y heterogéneas. Los formatos y procedimientos difieren entre las comunidades. Existen formatos para codificar series de tiempo para sensores meteorológicos, para modelar redes sociales, para codificar secuencias de ADN, etc. Es posible que en su área exista un formato estándar y abierto especializado para los datos de su investigación. Al publicar en ese formato facilitará a sus pares la reutilización de sus datos. Cuando no exista ese formato, lo más

conveniente es basarse en un formato estándar y genérico como XML<sup>24</sup> o HDF<sup>25</sup>, que poseen una gran variedad de herramientas para su manipulación.

Además, están bien documentados lo que permite ahorrar esfuerzo al momento de describir los datos.

### 4.3. Asesorarse con especialistas

La publicación y gestión de datos es una tarea intrínsecamente interdisciplinaria. Por ello puede ser útil pedir asesoría a profesionales del área de las ciencias de la información e informática, en particular a quienes se dedican a la gestión de datos y que tienen vocación por el acceso abierto<sup>26</sup>.

### 4.4. Describir los datos

Para ayudar a usar los datos publicados, hay que proveer una descripción de los mismos: lo que contienen, quién los obtuvo, cuándo, qué metodología se usó, el identificador del artículo donde se generaron, cómo se estructuran y cómo se pueden usar. Estas descripciones (o metadatos) pueden ir incrustadas en los datos (como parte del archivo descargable) o en la ficha de un catálogo de datos. Siempre que un formato de archivo permita incrustar metadatos, es recomendable hacerlo.

---

<sup>24</sup>XML (Extensible Markup Language) es un formato desarrollado por el Consorcio de la Web para codificar datos semi-estructurados. Más información se puede encontrar en <http://www.w3.org/XML/>.

<sup>25</sup>HDF (Hierarchical Data Format) es un formato desarrollado en la Universidad de Illinois para la codificación de datos jerárquicos. Más información se puede encontrar en 26.

<sup>26</sup>En el sitio <http://datoscientificos.cl> hay enlaces actualizados a comunidades de datos abiertos y a personas con conocimiento sobre cómo publicar datos en áreas específicas.

La Web es el estándar mundial para el intercambio de información y de datos. En ella también existe un estándar denominado RDF<sup>27</sup> que permite describir datos y enlazarlos (en el nivel de los metadatos) con otros datos publicados en la Web. Usar RDF para describir los datos es una buena práctica.

#### **4.5. Asignar un identificador a los datos**

Para poder referenciar los datos hay que identificarlos con un sistema estándar. De este modo, ese identificador se podrá usar tanto en el artículo para el que se generaron los datos como en artículos de otros investigadores que reutilicen esos datos. En la Web, un buen identificador son las URL, por su globalidad y porque permiten acceder a información sobre el objeto identificado. De este modo, la manera más sencilla de publicar los datos es asignarles una URL que permita descargar el archivo y acceder a su descripción.

Si bien las URL pueden facilitar el acceso a los datos, en el largo plazo pueden presentar limitaciones. Un problema común es la desaparición del sitio donde se publicaban archivos de datos o cambios en su dominio. Una solución son los identificadores generados a partir de los datos y que son independientes del lugar donde éstos se publiquen.

El uso adecuado de identificadores en los datos y en los artículos facilitará la formación de una red virtuosa entre la literatura científica y los datos.

---

<sup>27</sup> RDF (Resource Description Framework) es un formato desarrollado por el Consorcio de la Web para implementar una red de datos global. Más información sobre RDF se puede encontrar en <http://www.w3.org/RDF/>.

#### **4.6. Asegurar la perdurabilidad de los datos**

Publicar los datos en un sitio web personal tiene el riesgo de que el sitio sea borrado por descuido o que los datos dejen de estar disponibles si algo le ocurre al dueño del sitio. Para que los datos perduren se recomienda publicarlos en lugares más estables como repositorios institucionales de datos o en repositorios abiertos<sup>28</sup>. Si en su organización no existe un repositorio para datos, se debe fomentar la creación de uno.

#### **4.7. Considerar el uso de bases de datos**

En muchos casos puede que los datos estén contenidos en un archivo y sean procesados mediante alguna aplicación disponible o que haya sido creada durante la investigación. En estos casos, cuando se trate de archivos pequeños, publicarlos en un repositorio pareciera ser el camino adecuado. En cambio, cuando los investigadores necesitan gestionar los datos de manera colaborativa, resulta más apropiado usar una base de datos.

Garantizar el acceso a datos contenidos en una base de datos implica aspectos adicionales a tomar en cuenta. Entregar acceso de lectura a una base de datos puede ser algo complejo cuando hay operaciones que implican mucho costo de cómputo. Además, gestionar una base de datos requiere definir permisos de lectura y de escritura para los distintos recursos albergados. Por ello, no resulta claro cómo extender la definición de “abierto” a las bases de datos, pero un requisito razonable es permitir a todos la posibilidad de replicar los datos en una base de datos propia.

---

<sup>28</sup>En el sitio <http://datoscientificos.cl> hay una lista actualizada de repositorios de datos abiertos donde publicar.



#### **4.8. Consultar el sitio web [datoscientificos.cl](http://datoscientificos.cl)**

Las prácticas recomendadas en este manual son sólo un primer apresto para la publicación de los datos. En la web y, en especial, en <http://datoscientificos.cl> hay variados recursos y personas que pueden asesorar la publicación de los datos.

## 5. Aspectos legales y operativos

A continuación se presentan preguntas y respuestas para clarificar algunas dudas en torno a los aspectos legales y operacionales de la publicación de datos de investigación.

### **¿Protege la propiedad intelectual los datos contenidos en una base de datos?**

El sistema legal de derechos de autor chileno sólo protege los datos que integran una base de datos cuando en sí mismos son una obra intelectual original.

### **¿Qué datos contenidos en una base de datos no están protegidos por derechos de autor?**

Los datos no tienen protección cuando no son originales; por ejemplo, cuando la base de datos se compone de un registro de temperaturas mínimas y máximas en un determinado lugar, o un simple listado de nombres de comunas y sus superficies y números de habitantes.

### **¿Qué bases de datos se protegen por derechos de propiedad intelectual?**

Según el sistema de derecho de autor chileno sólo se protegen las bases de datos cuando, por razones de la selección o disposición de sus contenidos, puedan considerarse originales y, por consiguiente, constituyan creaciones de carácter intelectual. Sin embargo, es discutible qué hace una obra original, pero en general se entiende como tal el hecho de que sea distinguible formalmente de otras obras similares. Por ejemplo, un tabla de datos meteorológicos podría considerarse una obra si es publicada en una aplicación que visualice

las temperaturas con colores sobre un mapa y permita desplazarse por superficie o por el tiempo. No obstante, aunque esta publicación sea considerada una obra, los datos podrían ser extraídos de ella, puestos sobre una tabla y redistribuidos libremente, ya que por sí solos no son obras originales.

### **¿Quién es el dueño de los datos?**

De acuerdo con la ley chilena, el dueño o titular de los datos que componen la base de datos es quien haya materialmente creado la obra original. En el caso que los datos no sean obras, los datos no tendrán titular y podrán ser libremente utilizados por terceros.

### **¿Qué derecho tiene el fabricante de una base de datos?**

En Chile, a diferencia de otros países, el fabricante no tiene derechos sobre una base de datos. Incluso si la base de datos es original, sea por la originalidad de la selección o por la disposición de sus contenidos, el fabricante no tiene derechos sobre ella. Quien sí podría eventualmente tener derechos es el creador de la base de datos.

### **¿El acceso abierto implica renunciar a la propiedad sobre los datos?**

No. El acceso abierto está estrechamente vinculado con la titularidad de los datos cuyo acceso se ofrece. Si existen derechos de autor sobre los datos compilados, ellos se mantienen intactos incluso si la compilación construida a partir de ellos ha sido liberada.

## **¿El acceso abierto implica renunciar a la propiedad sobre la base de datos?**

No. Una política de acceso abierto no atenta contra la titularidad o propiedad de los datos susceptibles de protección. El acceso abierto es un conjunto de garantías que permite a terceros poder acceder y manipular ciertas obras intelectuales, en este caso datos, en determinadas condiciones que respeten la titularidad del autor.

## **¿Qué significa dominio público?**

Dominio público es el estado legal en que están ciertas obras intelectuales, ya sea por el paso del tiempo o por la decisión de su titular. En el dominio público las obras pueden ser utilizadas y reproducidas sin necesidad de que el titular entregue su permiso.

## **¿Qué pasa con las autorizaciones o licencias extendidas antes de hacerse obligatorio el acceso abierto?**

Dichas autorizaciones se mantienen vigentes, a menos que las partes decidan modificar sus condiciones para hacerlas compatibles, si es que no lo son, con las políticas de acceso abierto.

## **¿Qué condiciones resultan análogas al dominio público?**

Todas aquellas condiciones que supongan libertad para hacer uso de las obras sin necesidad de permiso ni pago de regalías a los creadores. Por ejemplo, son análogas al dominio público las condiciones mediante las cuales un autor decide que su obra se puede reproducir, utilizar y ejecutar públicamente sin previo pago y siempre que se respete su calidad de autor en las obras derivadas.

## **¿Puedo ofrecer licencias pagadas sobre los datos de investigación a entidades interesadas?**

Esta respuesta depende de la política de datos abiertos vigente. Mientras los derechos de autor se sigan radicando en los beneficiarios de los fondos y la política de acceso abierto no sea excluyente, es posible que los titulares de los derechos ofrezcan licencias en paralelo, que podrán ser a título gratuito u oneroso. En particular, la propuesta de política de CONICYT<sup>29</sup> cumple con estos dos requisitos, por lo que si esta condición se mantuviese en la política que finalmente se adopte, la respuesta sería sí.

## **¿Puede otro centro de investigación utilizar los datos que pongo a disposición en cumplimiento del acceso abierto?**

Sí. Una política de acceso abierto supone precisamente permitir a otros centros de investigación y terceros en general acceder a los datos.

## **¿Puede una empresa copiar y vender los datos que pongo a disposición en cumplimiento del acceso abierto a datos?**

Sí. Cuando los datos se encuentran en dominio público, una empresa u organización con fines de lucro podrá copiar y vender dichos datos, sin que ello suponga que la empresa se los apropie. Los datos siguen estando en dominio público. Incluso, si la empresa los vende en una disposición original que les agregue valor, los datos por sí mismos no dejarán de ser abiertos, sólo podrá ser una obra protegida la disposición original de los datos creada por la empresa.

## **¿Puedo impedir el uso indebido o incorrecto de los datos que pongo a disposición del público en cumplimiento del acceso abierto a datos?**

No. El cumplimiento del estándar de dominio público respecto de los datos supone que éstos tienen un ciclo de vida distinto de aquél por el cual inicialmente se crearon. Así, quien generó dichos datos no puede controlar los resultados que el tratamiento alternativo de los mismos pueda suponer en el futuro.

## **¿Puedo revocar mi permiso de brindar acceso abierto a datos?**

El permiso es, en principio, irrevocable. Esto no significa que no puedan cambiar las condiciones de acceso entregadas en un primer momento. Significa que el titular de los derechos no puede impedir que quienes accedieron a los datos cuando se encontraban en acceso abierto puedan continuar haciendo uso de los mismos de acuerdo con esas condiciones.

## **¿Puedo obtener una exención al acceso abierto a datos cuando existan patentes involucradas o considerando mis compromisos internacionales?**

Esto depende de la política de datos abiertos. Una política puede considerar un mecanismo para otorgar exenciones por razones fundadas, tales como la existencia de patentes involucradas, modelos de negocios o compromisos internacionales. Sin embargo, una política de datos abiertos también puede poner plazos máximos a tales exenciones.

## 6. ¿Cómo avanzar hacia una política?

De acuerdo con las directrices de la OECD<sup>29</sup>, las estrategias para instalar una política de datos abiertos deben cumplir con los principios de apertura, flexibilidad, transparencia, conformidad legal, protección de la propiedad intelectual, responsabilidad, profesionalismo, interoperabilidad y sustentabilidad.

Chile forma parte de la OECD desde el año 2010, lo que además de las ventajas que trae la membresía, implica el compromiso de promover la política de acceso libre y la preservación de los datos de investigación científica.

Desde esta perspectiva, se propone discutir con la comunidad una política de acceso y preservación de datos generados con fondos públicos consistente con los estándares internacionales.

Para abrir la conversación en torno al diseño e implementación de una política para fomentar el acceso y la preservación de los datos de investigación, CONICYT ha elaborado la siguiente propuesta de política de datos abiertos, que sería aplicable a todos los proyectos financiados con fondos públicos:

1. Los beneficiarios de todas las líneas de investigación financiadas por CONICYT deberían publicar los informes, documentos y conjuntos de datos resultantes de sus investigaciones en repositorios abiertos y en formatos adecuados.
2. La publicación podrá hacerse efectiva, tanto en repositorios centralizados en CONICYT, como en repositorios pertenecientes a las instituciones a las que pertenezcan los

---

<sup>29</sup>OECD (2007), 6p. cit.

investigadores, siempre que estos últimos cumplan con los estándares acordados. En este último caso, los investigadores debieran comprometerse a informar a CONICYT las URL o identificadores de los documentos y conjuntos de datos publicados.

3. El plazo para la publicación de los datos en los repositorios abiertos no debiera superar el año después de rendido el informe final, aún cuando no se hayan publicado los respectivos artículos.

4. El plazo de publicación de los informes finales de las investigaciones por parte de CONICYT no debiera ser mayor a tres meses, una vez aprobados. El carácter de este informe debiera estar en concordancia con las bases de la línea de financiamiento que se esté reportando.

5. El plazo de publicación en repositorios abiertos de los artículos no debiera ser superior a seis meses después de publicados en revistas especializadas. Para cumplir con esta obligación será suficiente publicar el preprint.

6. Cuando la investigación genere la tramitación de una patente, esté vinculada a un modelo de negocios o existan compromisos internacionales que lo requieran, se podrá omitir la publicación de los datos más sensibles. Esta excepción no debiera exceder a los dos años después de aprobado el informe final.

7. Al aceptar el financiamiento, los investigadores asumen los términos y condiciones de esta política.

Invitamos a la comunidad científica a expresar su opinión, dejando sus comentarios y sugerencias en <http://datoscientificos.cl>.



<http://datoscientificos.cl>