OFERTA académica

Introducción a la **Ciencia Abierta**

8va ESCUELA VIRTUAL INTERNACIONAL CAVILA
Un espacio Multidisciplinario

EDUCACIÓN PÚBLICA Y GRATUITA | UNIVERSIDAD NACIONAL DE LA PLATA

8va ESCUELA VIRTUAL INTERNACIONAL CAVILA
Un espacio Multidisciplinario

DIRECCIÓN GENERAL
de educación a distancia y tecnologías
UNIVERSIDAD NACIONAL DE LA PLATA

# Ciencia abierta: Datos abiertos

**Dra. Marisa R. De Giusti**
**PREBI-SEDICI** Universidad Nacional de La Plata
**CESGI** Comisión de Investigaciones Científicas

Agosto de 2021



PREBI SEDICI
prebi.unlp.edu.ar sedici.unlp.edu.ar

UNIVERSIDAD NACIONAL DE LA PLATA

CESGI Centro de Servicios en Gestión de Información

COMISIÓN DE INVESTIGACIONES CIENTÍFICAS

# Objetivos del aprendizaje

1. Conocer las características de los datos abiertos, incluyendo su empaquetado y la documentación apropiada para lograr que otros los entiendan, reproduzcan y reutilicen.
2. Ser capaz de diferenciar y trabajar con datos confidenciales y datos abiertos.
3. Ser capaz de transformar un conjunto de datos en uno que pueda compartirse en abierto (formato no propietario), cumpla con los estándares de los principios FAIR y esté diseñado para maximizar el acceso, la transparencia y la reutilización, y proporcione metadatos suficientes.
4. Ser capaz de crear un Plan de Gestión de Datos.

# Datos de investigación

Información cuantitativa o cualitativa obtenida por los investigadores en el curso de su trabajo y que puede provenir de experimentos, observaciones, modelos, encuestas, entrevistas o a partir de la transformación de información ya existente.

**Sirven para validar los resultados o conclusiones de la investigación**

# Datos abiertos

Los datos de investigación abiertos son datos de acceso gratuito que pueden ser reutilizados, mezclados y redistribuidos para la investigación académica y la docencia, entre otros usos. Idealmente, los datos abiertos no tienen restricciones para su reutilización y redistribución y cuentan con licencias acordes a ello.

Se debe atender a la privacidad de los “datos sensibles” para proteger la identidad de las personas: se pueden incluir restricciones de acceso

Compartir los datos de manera abierta facilita su consulta, es la base para la reproducibilidad y verificación de la investigación, y abre un camino para promover la colaboración.

Deben estar acompañados de licencias de uso abiertas.


**Fuente:** Strasser, C., Cook, R., Michener, W., & Budden, A. (2012). *Primer on Data Management: What you always wanted to know*. Recuperado a partir de http://escholarship.org/uc/item/7tf5q7n3.pdf

Toda **investigación** tiene un ciclo vital. y atraviesa distintas **fases**:

- Concepción inicial
- Planificación
- Creación de la propuesta
- Comienzo de la investigación
- Recolección de datos
- Finalización de la investigación

En las distintas etapas, las actividades asociadas al manejo y uso de los datos van cambiando.

Vilches, C. (s. f.). *Biblioguias: Gestión de datos de investigación: 1.4 El ciclo de vida de los datos*. Recuperado 30 de agosto de 2021, de https://biblioguias.cepal.org/c.php?g=495473&p=4994826

# Principios FAIR

**El énfasis está en mejorar la capacidad de las máquinas para encontrar y utilizar los datos**


Imagen: Australian National Data Services.

Son un conjunto de principios rectores para hacer que los datos de investigación sean fáciles de encontrar (**F**indable), accesibles (**A**ccesible), interoperables (**I**nteroperable y reutilizables (Wilkinson et al., 2016). Estos principios proporcionan una guía para la gestión de datos científicos y se dirigen directamente a los productores de datos y a los editores de datos para promover el uso máximo de los datos de investigación.

# El problema de los datos

**Dada la diversidad de disciplinas y métodos de investigación:** ss necesario conocer el ciclo de vida de cada investigación.

La naturaleza de los datos: cualitativos o cuantitativos o mixtos que se generan.

Lo tipos y las cantidades de los datos recogidos.

Los formatos de guardado de los datos según su especie.

Es necesario conocer el proceso de recolección y el contexto.

Es necesario describirlos.

★ **Es importante contar con un relato claro sobre el proceso, el contexto y el momento de recolección de los datos. Lo ideal es contar con un Plan de gestión de datos debidamente comunicado y compartido.**

# Depósito y curación de datos

La curación de datos incluye: organizar, describir, preservar publicar, proteger la privacidad y la confidencialidad, facilitar el descubrimiento, el acceso y la reutilización de los datos.

El depósito y posterior publicación en un repositorio precisa de expertos del área y de personal del repositorio con capacidades de catalogación e implica un diálogo que comienza en el reconocimiento y familiarización de los quehaceres de una y otra parte.

Supone conocer o aprender estándares de descripción de los datos según su área.

# Publicación de datos

La mayoría de los docentes e investigadores está familiarizado con la publicación de artículos, libros o presentaciones en congresos en acceso abierto.

Los datos que recoge en el ciclo de la investigación generalmente quedan guardados en máquinas personales, pero hoy cada vez más se precisa publicar los datos asociados a un artículo y que confirman la investigación e incluso es posible hoy en día publicar un artículo sobre los datos que fundamentan una investigación.

- Siempre es importante guardar los datos en un repositorio
- La editorial vinculada a la publicación de un artículo puede querer alojarlos
- Además hay repositorios específicos para datos o institucionales que los albergan.

# Publicación de datos

Publicar datos es equivalente a publicar un artículo:

- los datos se deben **describi**r (metadatos),
- **revisar** en cuanto a su **calidad** (metodología, relevancia… )
- **permitir** su **búsqueda**,
- **identificación** unívoca
- **localización** y
- **los datos deben** poder **ser citables** como los artículos.
    - Esto implica la asignación de un identificador persistente: DOI por ejemplo.

# Citación de datos

La correcta citación de los conjuntos de datos, hace más fácil su identificación, recuperación, indexación e inclusión en indicadores de impacto, plataformas y recursos y, en definitiva, su difusión y visibilidad en general.

Los datos básicos que se necesitan son: la referencia bibliográfica, con los suficientes metadatos para su identificación, así como un identificador persistente tipo Handle o DOI.

Ejemplo de citación de un set de datos, extraído de ZENODO:

https://zenodo.org/record/5234181#.YSTxF4hKg2w

Banda, Juan M., Tekumalla, Ramya, Wang, Guanyu, Yu, Jingyuan, Liu, Tuo, Ding, Yuning, Artemova, Katya, Tutubalina, Elena, & Chowell, Gerardo. (2021). A large-scale COVID-19 Twitter chatter dataset for open scientific research - an international collaboration (Version 76) [Data set]. Zenodo. https://doi.org/10.5281/zenodo.5234181

**Todas las versiones se citan con el mismo DOI. Siempre resuelve a la última.**

# Gestión de datos de investigación





## ¿Por qué citar datos de investigación?

Como sostiene el DCC "la motivación para citar datos surge del reconocimiento que los datos generados en el transcurso de una investigación son tan valiosos para el discurso académico como lo son las monografías y los artículos." (Bal & Duke, 2015)

El citar apropiadamente los datos de investigación ofrece muchos **beneficios para investigadores e instituciones de investigación**. Por ejemplo:

- Dar mayor **publicidad** a la investigación realizada
- Recibir el **debido crédito** por su rol en la recolección de datos
- Hacer un **seguimiento del impacto** de un determinado set de datos, a través de métricas que apliquen específicamente a este tipo de resultados de investigación
- Potencialmente, recibir **incentivos por la publicación y uso de los sets de datos**, como ocurre con otros productos de la investigación académica
- Asegurar que los datos puedan ser efectivamente **localizados** y, con ello, **reutilizados por otros**
- Al vincular las publicaciones directamente con los sets de datos (tanto desde revistas académicas como en los repos datos), **dar sustento a las afirmaciones que se realizan en la investigación, advertir faltas y proveer información contextual** para comprender un set de datos específico (documentación)

Vilches, C. (s. f.). *Biblioguias: Gestión de datos de investigación: Citación de los datos*. Recuperado 30 de agosto de 2021, de https://biblioguias.cepal.org/gestion-de-datos-de-investigacion/citacion

# ¿Cómo citar datos de investigación?

Existen **dos mecanismos** que deben ser utilizados al citar los datos de investigación dentro de una publicación académica. Primero, debe incluirse siempre una **declaración sobre acceso a los datos**, y luego, deben incluirse las **referencias en texto** que sean necesarias.

Declaración de acceso a los datos | Citas en el texto

La **declaración de acceso a los datos** debe incluir:

- La definición de qué datos están disponibles y en qué repositorio, o bien especificar qué persona debe ser contactada para solicitar acceso a los datos;
- Una URL (idealmente debe ser un identificador persistente, por ejemplo un DOI) o la dirección de correo electrónico de la institución o departamento correspondiente (nunca una dirección de correo personal);
- Una mención a la base legal o ética sobre la cual se restringen los datos, cuando aplique, y
- Cuando los datos no estén disponibles de forma abierta, debe incluirse un enlace persistente a la información sobre condiciones que rigen el acceso a los datos.

Vilches, C. (s. f.). *Biblioguias: Gestión de datos de investigación: Citación de los datos*. Recuperado 30 de agosto de 2021, de https://biblioguias.cepal.org/gestion-de-datos-de-investigacion/citacion

Aunque algunas revistas exigen la ubicación de esta declaración dentro de una sección específica del artículo, podrían existir otros requerimientos o quedar al criterio de quien ha elaborado la publicación. En este último caso, puede ubicarse junto a mención de reconocimiento al organismo financiador.

A continuación damos algunos ejemplos de declaraciones de acceso a los datos, dependiendo del tipo de acceso que se trate:

- *"Todos los datos creados en esta investigación están disponibles en el archivo de la Universidad del Valle en [enlace persistente] bajo licencia Creative Commons Atribución 3.0"*
- *"Este estudio ha sido elaborado en base a los datos que están disponibles de forma pública en [incluir enlace persistente]. Mayor información sobre el procesamiento de datos disponible en el Archivo de la Universidad del Valle en [enlace persistente]"*
- *"Esta publicación ha sido elaborada en base múltiples conjuntos de datos que se encuentran disponibles de forma abierta en las ubicaciones citadas en la sección de referencias"*
- *"Los datos que apoyan esta investigación están disponibles a petición de las personas interesadas. Por favor solicitar acceso a dependencia@institución.com"*

Vilches, C. (s. f.). *Biblioguias: Gestión de datos de investigación: Citación de los datos*. Recuperado 30 de agosto de 2021, de https://biblioguias.cepal.org/gestion-de-datos-de-investigacion/citacion

# Cita tus datos de investigación

## Por qué es importante citar los datos:

- Los conjuntos de datos también son resultados de investigación como los artículos, monografías, etc.
- Facilita la identificación y el acceso a los datos y de esta forma su localización, validación y reutilización.
- Permite reconocer la autoría de sus creadores.
- Facilita la métrica e impacto de los datos.
- Favorece la transparencia de la investigación científica.

## Buenas prácticas para citar datos:

- Se debe facilitar la identificación, localización y el acceso a los datos mediante un identificador único y persistente (DOI, Handle, etc.).
- Cada conjunto y subconjunto de datos (dataset) debe citarse de forma independiente.
- Las citas de los datos utilizados han de aparecer en la sección de referencias bibliográficas de la publicación resultante.
- Se recomienda incluir un identificador único de autor (ORCID, etc.).

**Enlaza los datos con los documentos resultado de investigación y viceversa, y crea las referencias bibliográficas de los mismos.**

## Elaboración de la cita

- Existen elementos mínimos obligatorios (O) y otros recomendados (R) que se combinan para elaborar la cita en cualquier estilo estándar (APA, MLA, Chicago, etc.) o los propuestos por los principales repositorios de datos (Dataverse, Dryad, etc.).

Autor(es) **(O)** | Identificador autor **(R)** | Fecha **(O)** | Título **(O)**

Identificador único persistente **(O)** | Tipo de recurso **(O)**

Versión y/o Edición **(O)** | Repositorio de datos **(R)** | Publicación **(R)**

Productor **(R)** | Ámbito geográfico **(R)** | Ámbito temporal **(R)**

### Ejemplo de cita estilo APA

Autoría | Título

Remesar Betlloch, X., Antelo, A., Llivina, C., Albà, E., Berdié, L., Agnelli, S.,... Alemany, M. (2015). *Influence of a hyperlipidic diet on the composition of the non-membrane lipid 6 pool of red blood cells of male and female rats.* [Dataset]. Versión de 22 de junio de 2015 Dipòsit digital de documents de la UAB. http://hdl.handle.net/2445/66010

Fecha | Repositorio | Tipo de recurso | Identificador único y persistente | Versión

El personal de tu Biblioteca te puede asesorar

crue Universidades Españolas Red de Bibliotecas REBIUN

# Empaquetamiento de datos

“Los paquetes de datos son contenedores para describir y compartir archivos de datos, y generalmente incluyen un archivo de metadatos que describe las características y el contexto de un conjunto de datos. Esto puede incluir aspectos tales como información de su creación, procedencia, tamaño, tipo de formato, definiciones de campos, así como cualquier otro archivo contextual relevante, como scripts de creación de datos o documentación textual”.

Bezjak, S., Conzett, P., Fernandes, P. L., Görögh, E., Helbig, K., Kramer, B., Labastida, I., Niemeyer, K., Psomopoulos, F., Ross-Hellauer, T., Schneider, R., Tennant, J., Verbakel, E., Clyburne-Sherin, A., Brinken, H., & Heller, L. (2019). *Manual de Capacitación sobre Ciencia Abierta*. Zenodo. p. 26 https://doi.org/10.5281/zenodo.2588214 (original) https://book.fosteropenscience.eu/es

Soloaga, I., Fernández, E. C., & De Giusti, M. R. (2020, diciembre 3). *Cómo crear paquetes de información para repositorios digitales*. Mini Curso «Cómo crear paquetes de información para el repositorio» (Brasilia, 2020). http://sedici.unlp.edu.ar/handle/10915/110561

# Empaquetamiento de datos

“Los datos son para siempre: los conjuntos de datos sobreviven a su propósito original. Las limitaciones de los datos pueden ser obvias dentro de su contexto original, como un catálogo de biblioteca, pero pueden no ser evidentes una vez que los datos se han separado del entorno para el que se crearon”.

“Los datos no son autosuficientes: la información sobre el contexto y la procedencia de los datos (cómo y por qué se creó, qué objetos y conceptos del mundo real representan, las limitaciones de sus valores) es necesaria para ayudar a los consumidores a interpretarlos”.

“La estructuración de metadatos sobre conjuntos de datos de una manera estándar y legible por máquinas fomenta la promoción, la capacidad de intercambio y la reutilización de los datos”.

Bezjak, S., Conzett, P., Fernandes, P. L., Görögh, E., Helbig, K., Kramer, B., Labastida, I., Niemeyer, K., Psomopoulos, F., Ross-Hellauer, T., Schneider, R., Tennant, J., Verbakel, E., Clyburne-Sherin, A., Brinken, H., & Heller, L. (2019). *Manual de Capacitación sobre Ciencia Abierta*. Zenodo. p. 26 https://doi.org/10.5281/zenodo.2588214 (original) https://book.fosteropenscience.eu/es

# Privacidad y confidencialidad

## ¿Qué se entiende por privacidad?

“El ´derecho a la privacidad´ se refiere al estar libre de intrusiones o perturbaciones en la vida privada o en los asuntos personales. Toda investigación debe esbozar estrategias para proteger la privacidad de los sujetos involucrados, y también sobre cómo el investigador tendrá acceso a la información”.

“Los conceptos de privacidad y confidencialidad están relacionados pero no son lo mismo. La privacidad se refiere al individuo o al sujeto, mientras que la confidencialidad se refiere a las acciones del investigador”.

Vilches, C. (s. f.). *Biblioguias: Gestión de datos de investigación: Privacidad de los datos y confidencialidad*. Recuperado 30 de agosto de 2021, de https://biblioguias.cepal.org/c.php?g=495473&p=4398114

# Datos sensibles

Una serie de requisitos éticos se aplican a la gestión de los datos de investigación, particularmente cuando ésta involucra a personas.

Es posible compartirlos. Hay broker (intermediarios) de datos que se encargan de mantener la confidencialidad, están en su mayoría en la temática de la salud.

Los datos sensibles pueden extraerse o hacerse la anonimización y publicar el “subset”, abriendo la parte confidencial en caso que se considere adecuado.

Los datos anonimizados no requieren consentimiento para compartir o publicar, pero se considera ético informar a los sujetos sobre el uso y destino de los datos.

# Confidencialidad

La confidencialidad se refiere al acuerdo del investigador con el participante acerca de cómo se manejará, administrará y difundirá la información privada de identificación. La propuesta de investigación debe describir las estrategias para mantener la confidencialidad de los datos identificables, incluidos los controles sobre el almacenamiento, la manipulación y el compartir datos personales.

Para minimizar los riesgos de divulgación de información confidencial:

- Si es posible, recopile los datos necesarios sin utilizar información de identificación personal.
- Si se requiere información de identificación personal, reitre la identificación de los datos después de la recolección o tan pronto como sea posible.
- Evite transmitir electrónicamente datos personales no cifrados.

Comisión Económica para América Latina y el Caribe. (s. f.). *Biblioguias de la CEPAL*. Recuperado 30 de agosto de 2021, de https://biblioguias.cepal.org/

# Confidencialidad

- Otras consideraciones incluyen la retención de instrumentos originales de recolección, tales como cuestionarios o grabaciones de entrevistas. Una vez que estos se transfieren a un paquete de análisis o se realiza una transcripción y la calidad es asegurada o validada, puede que ya no haya razón para retenerlos.
- Preguntas sobre qué datos conservar y por cuánto tiempo deben ser planificados con antelación y dentro del contexto de sus capacidades para mantener la confidencialidad de la información

Comisión Económica para América Latina y el Caribe. (s. f.). *Biblioguias de la CEPAL*. Recuperado 30 de agosto de 2021, de https://biblioguias.cepal.org/

# Etiquetado de datos

DataTags es un marco digital que permite asesorar sobre las restricciones legales, contractuales y políticas que rigen las decisiones sobre el intercambio de datos. El sistema DataTags hace a un usuario una serie de preguntas para determinar las propiedades clave de un conjunto de datos y aplica reglas inferenciales para determinar qué leyes, contratos y buenas prácticas son aplicables. El resultado es un conjunto de DataTags recomendados, o etiquetas simples e icónicas que representan una política de datos legible por humanos y utilizable por máquinas, y un acuerdo de licencia a medida para el conjunto de datos. El sistema DataTags está diseñado para integrarse con el software del repositorio de datos y también funciona como una herramienta independiente. DataTags se ha desarrollado en la Universidad de Harvard. En Europa, DANS está trabajando para ajustar DataTags a la legislación europea General Data Protection Regulation (GDPR) (cf. DANS GDPR DataTags).

Bezjak, S., Conzett, P., Fernandes, P. L., Görögh, E., Helbig, K., Kramer, B., Labastida, I., Niemeyer, K., Psomopoulos, F., Ross-Hellauer, T., Schneider, R., Tennant, J., Verbakel, E., Clyburne-Sherin, A., Brinken, H., & Heller, L. (2019). *Manual de Capacitación sobre Ciencia Abierta*. Zenodo. p. 28 https://doi.org/10.5281/zenodo.2588214 (original) https://book.fosteropenscience.eu/es

# Compartir los datos

Como se mencionó anteriormente, el objetivo final de compartir tus datos de investigación es hacerlos reutilizables al máximo. Para ello, antes de compartir datos debe gestionarlos de acuerdo con las mejores prácticas. Esto incluye, por ejemplo, la documentación y la elección de formatos de archivos y licencias abiertos. Puedes leer más sobre estos temas en Sección 4: Investigación reproducible y análisis de datos así como Sección 6: Licenciamiento abierto y formatos de archivo.

Además de compartir los datos, la apertura de la investigación depende del intercambio de materiales. Los materiales que utilizan los investigadores son específicos de cada disciplina y, a veces, exclusivos de un laboratorio: reactivos, protocolos, cuadernos, software y hardware

Bezjak, S., Conzett, P., Fernandes, P. L., Görögh, E., Helbig, K., Kramer, B., Labastida, I., Niemeyer, K., Psomopoulos, F., Ross-Hellauer, T., Schneider, R., Tennant, J., Verbakel, E., Clyburne-Sherin, A., Brinken, H., & Heller, L. (2019). *Manual de Capacitación sobre Ciencia Abierta*. Zenodo. p. 28 https://doi.org/10.5281/zenodo.2588214 (original) https://book.fosteropenscience.eu/es

Revistas de datos sobre biodiversidad: https://www.gbif.org/data-papers

**The Journal of Open Humanities Data** (JOHD) features peer reviewed publications describing humanities data or techniques with high potential for reuse. https://openhumanitiesdata.metajnl.com/about/

Datasets de **Ciencias de la Tierra**: https://www.earth-system-science-data.net/

**Biomedical Data Journal**: https://www.biomed-data.eu/

Energía Nuclear https://www.nndc.bnl.gov/nds/

Datasets sobre Medioambiente: https://www.soils.org/publications/journals/author-resources/jeq-instructions/dataset/

**Geoscience Data Journal** https://rmets.onlinelibrary.wiley.com/journal/20496060

**Open Humanities Data:** https://openhumanitiesdata.metajnl.com

**Journal of Open Psychology Data** (JOPD) https://openpsychologydata.metajnl.com/

**The Astrophysical Journal** tiene una serie de suplementos dedicados a documentar datasets: https://iopscience.iop.org/journal/0067-0049/page/Scope

**Data in brief** https://www.journals.elsevier.com/data-in-brief

**Journal of Open Archaeology Data** (JOAD) https://openarchaeologydata.metajnl.com/

# Ley de Creación de Repositorios Digitales Institucionales de Acceso Abierto Nº 26.899

**Artículo 3°-** Todo subsidio o financiamiento proveniente de agencias gubernamentales y de organismos nacionales de ciencia y tecnología del SNCTI, destinado a proyectos de investigación científico-tecnológica que tengan entre sus resultados esperados la generación de datos primarios, documentos y/o publicaciones, deberá contener dentro de sus cláusulas contractuales la presentación de un plan de gestión acorde a las especificidades propias del área disciplinar, en el caso de datos primarios y, en todos los casos, un plan para garantizar la disponibilidad pública de los resultados esperados según los plazos fijados en el artículo 5º de la presente ley. **A los efectos de la presente ley se entenderá como dato primario a todo dato en bruto sobre los que se basa cualquier investigación y que puede o no ser publicado cuando se comunica un avance científico pero que son los que fundamentan un nuevo conocimiento.**

# Ley de Creación de Repositorios Digitales Institucionales de Acceso Abierto Nº 26.899

**Artículo 5°- Los investigadores, tecnólogos, docentes, becarios de posdoctorado y estudiantes de maestría y doctorado** cuya actividad de investigación sea financiada con fondos públicos, **deberán depositar o autorizar expresamente el depósito de una copia de la versión final de su producción científico-tecnológica publicada o aceptada para publicación y/o que haya atravesado un proceso de aprobación por una autoridad competente o con jurisdicción en la materia, en los repositorios digitales de acceso abierto de sus instituciones**, en un plazo no mayor a los seis (6) meses desde la fecha de su publicación oficial o de su aprobación. **Los datos primarios de investigación deberán depositarse en repositorios o archivos institucionales digitales** propios o compartidos y estar disponibles públicamente **en un plazo no mayor a cinco (5) años del momento de su recolección**, de acuerdo a las políticas establecidas por las instituciones, según el artículo 2º.

# Resolución 753- E/2016 MINCYT

## Reglamento operativo para la aplicación de la Ley nº 26.899

### ANEXO I

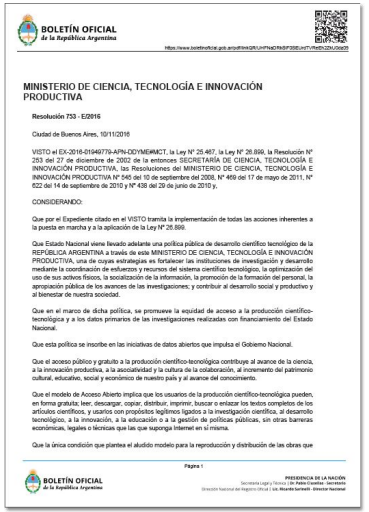
- **CAPÍTULO I**
  DE LOS PRINCIPIOS RECTORES DEL ACCESO ABIERTO Y DEL ÁMBITO DE IMPLEMENTACIÓN DE LA LEY

- **CAPÍTULO II**
  DE LOS SUJETOS ALCANZADOS POR LA LEY Y SUS RESPONSABILIDADES

- **CAPÍTULO III**
  ACERCA DE LAS POLÍTICAS INSTITUCIONALES DE ACCESO ABIERTO Y/O DEL REPOSITORIO INSTITUCIONAL

- **CAPÍTULO IV**
  ACERCA DE LOS REPOSITORIOS INSTITUCIONALES

- **CAPÍTULO V**
  ACERCA DE LAS ETAPAS DE ADECUACIÓN A LA LEY Y DE LA APLICACIÓN DE LA SANCIÓN



BOLETÍN OFICIAL
de la República Argentina

MINISTERIO DE CIENCIA, TECNOLOGÍA E INNOVACIÓN PRODUCTIVA

Resolución 753 - E/2016

Ciudad de Buenos Aires, 10/11/2016

VISTO el EX-2016-01949779-APN-DDYME#MCT, la Ley N° 25.467, la Ley N° 26.899, la Resolución N° 253 del 27 de diciembre de 2002 de la entonces SECRETARÍA DE CIENCIA, TECNOLOGÍA E INNOVACIÓN PRODUCTIVA, las Resoluciones del MINISTERIO DE CIENCIA, TECNOLOGÍA E INNOVACIÓN PRODUCTIVA N° 546 del 10 de septiembre del 2008, N° 469 del 17 de mayo de 2011, N° 622 del 14 de septiembre de 2010 y N° 438 del 29 de junio de 2010 y,

CONSIDERANDO:

Que por el Expediente citado en el VISTO tramita la implementación de todas las acciones inherentes a la puesta en marcha y a la aplicación de la Ley N° 26.899.

Que Estado Nacional viene llevado adelante una política pública de desarrollo científico tecnológico de la REPÚBLICA ARGENTINA a través de este MINISTERIO DE CIENCIA, TECNOLOGÍA E INNOVACIÓN PRODUCTIVA, una de cuyas estrategias es fortalecer las instituciones de investigación y desarrollo mediante la coordinación de esfuerzos y recursos del sistema científico tecnológico, la optimización del uso de sus activos físicos, la socialización de la información, la promoción de la formación del personal, la apropiación pública de los avances de las investigaciones; y contribuir al desarrollo social y productivo y al bienestar de nuestra sociedad.

Que en el marco de dicha política, se promueve la equidad de acceso a la producción científico-tecnológica y a los datos primarios de las investigaciones realizadas con financiamiento del Estado Nacional.

Que esta política se inscribe en las iniciativas de datos abiertos que impulsa el Gobierno Nacional.

Que el acceso público y gratuito a la producción científico-tecnológica contribuye al avance de la ciencia, a la innovación productiva, a la asociatividad y la cultura de la colaboración, al incremento del patrimonio cultural, educativo, social y económico de nuestro país y al avance del conocimiento.

Que el modelo de Acceso Abierto implica que los usuarios de la producción científico-tecnológica pueden, en forma gratuita; leer, descargar, copiar, distribuir, imprimir, buscar o enlazar los textos completos de los artículos científicos, y usarlos con propósitos legítimos ligados a la investigación científica, al desarrollo tecnológico, a la innovación, a la educación o a la gestión de políticas públicas, sin otras barreras económicas, legales o técnicas que las que suponga Internet en sí misma.

Que la única condición que plantea el aludido modelo para la reproducción y distribución de las obras que

Página 1

BOLETÍN OFICIAL
de la República Argentina

PRESIDENCIA DE LA NACIÓN

*Resolución 753- E/2016 MINCYT*, (2016) (testimony of Ministerio de Ciencia, Tecnología e Innovación Productiva). https://www.argentina.gob.ar/normativa/nacional/resoluci%C3%B3n-753-2016-267833

## ANEXO I - CAPÍTULO IV

ACERCA DE LOS REPOSITORIOS INSTITUCIONALES

### Artículo 16.- Datos Primarios. Continuación

## Plan de gestión de datos

- Documento que reúne información sobre los conjuntos de datos que se generarán en el marco de un proyecto de investigación con el objetivo de lograr su organización, depósito en repositorios institucionales, preservación a largo plazo y difusión.

- Los PDG se crean antes de comenzar a ejecutar un proyecto de investigación y luego evolucionan durante todo el proyecto.

## Algunas excepciones

- **Derechos de propiedad industrial** (acuerdos de confidencialidad o publicación de una patente)
- **Datos sensibles** (especies protegidas, cuestiones de soberanía nacional)
- **Acuerdos previos con terceros**
  - El alcance de la excepción por acuerdos previos con terceros se extiende a aquellos acuerdos firmados, previamente al inicio del proyecto de investigación, con terceras partes no alcanzadas por la Ley que co-financian la investigación y han requerido **plazos diferentes** a los que establece la ley.
  - Se excluye de esta excepción, a los acuerdos con terceros que no han co-financiado la investigación, por ejemplo las editoriales. Las editoriales no son susceptibles de excepciones.

# Objetivos de un plan de gestión de datos

- Verificación del proyecto de investigación
  - Confiabilidad de que es posible reproducir los resultados
  - Validez del diseño de la investigación y los resultados
  - transparencia del gasto público
- Preservación de los datos
  - Prevenir la pérdida de los datos durante la ejecución del proyecto
  - Garantizar el acceso a los datos una vez finalizado el proyecto
- Publicación de los datos:
  - Difusión: Promoción del proyecto de investigación y de los organismos que financian.
  - Impacto: Crédito para los investigadores que publican datos curados
  - Acceso abierto: reutilización de los datos por parte de toda la comunidad para crear nuevos conocimientos y acelerar el avance científico

# ¿Qué requisitos mínimos debe contener un PGD en Argentina?

- Creadores
- Identificación del proyecto de investigación
- Identificación de la agencia u organismo que financia la investigación
- Tipología de los datos que se generarán y se recopilarán durante el proyecto
- Estándares que se utilizarán
- Cómo serán explotados y/o compartidos/accesibles los datos para su verificación, reutilización,redistribución, etc. En cuál/cuáles repositorio/s se alojarán los conjuntos de datos generados y recopilados, períodos de embargo, software necesario y otras herramientas que permitan su reutilización
- Condiciones de acceso (licencias de uso)
- Medidas de conservación y preservación que se tomarán durante el proyecto y previamente a su depósito en el repositorio

# ¿Cuáles son los requisitos mínimos debe contener un PGD en Argentina?

**Descripción general de los conjuntos de datos que serán generados o recopilados:**

- origen
- naturaleza
- escalas y métricas utilizadas
- volumen de los datos
- usuarios potenciales
- palabras clave
- idioma/s
- fechas o períodos relevantes, como ser de recolección
- cobertura geográfica
- metodología de colecta o generación
- procesamiento
- datos asociados
- extensiones de archivo y formato;
- estructura/organización de los datos;
- lista de variables;
- glosarios de códigos y abreviaturas;
- enlace a las publicaciones científicas que ésos datos respaldan
- información sobre si existen o no datos similares, y las posibilidades de su integración y reutilización
- volumen
- versiones