

La Gestión de Datos de Investigación

Marisa Pérez Aliende

Universidad Autónoma de Madrid

mp.aliende@uam.es

Sevilla, 13 de Junio 2017





Archivo de los datos de investigación

- Licencias de depósito
- Licencias de uso
- Repositorios de datos
- Los metadatos

Archivo de los datos de investigación

Compartir los datos de investigación tiene beneficios y desafíos tales como:

- la propiedad de los datos
- los datos sensibles
- la información confidencial

Permitiendo el acceso
o
Restricciones

Embargos:

Período de tiempo durante el cual el acceso a los datos está restringido.

- Suele durar entre 12-24 meses.
- Período con el que cuentan los investigadores para publicar usando sus datos antes de darles publicidad.
- Influenciados por los *stakeholders*: editores, agencias de financiación.

Tecnológicas:

Basados en requerimientos de afiliación de los usuarios:

- En una institución, autenticarse para acceder a los datos.
- Loguearse en un sistema particular.
- Pertenecer a un cierto de grupo de investigación.

Acuerdos de uso:

Acuerdo explícito entre el productor de datos y los usuarios de datos secundarios.

- Puede imponer normas, reglas para reusar, almacenar, y divulgar los datos.
- Suele requerirse cuando los datos tienen el potencial de identificar aspectos humanos directa o indirectamente.
- Antes de firmarlo, leerlo detenidamente pues el hecho de no cumplirlo puede acarrear consecuencias.

Licencias de depósito

El acceso a los datos de investigación está sujeto a la aceptación de unos **términos de uso** que puede tener información sobre la limitación al acceso, estableciendo barreras (autenticación), la privacidad, confidencialidad, la suspensión de uso y la aplicación de una licencia estándar que puede impactar sobre su potencial reutilización.

Los investigadores deben discutir:

- Restricciones de acceso
- Licencias

Para el repositorio elegido

De depósito

De uso

Es un documento por el que un investigador / grupo de investigación o personal académico que deposita los datos y materiales está de acuerdo con los términos y condiciones que se indican en la misma, pudiendo hacerlo personalmente o de forma delegada.

- Se firma para que el administrador del repositorio tenga el derecho de realizar acciones tales como: almacenar, copiar, modificar los datasets sin cambiar el contenido, preservar.
- y al mismo tiempo confirmar que quien deposita tiene la obligación de suministrar los datasets al repositorio.

Licencias de depósito: Ejemplos

Technische Universität Berlin
 Universitätsbibliothek
 Hauptabteilung 3 Elektronische Dienste
 Fasanenstraße 88
 10623 Berlin
 E-Mail: team@srf.tu-berlin.de

Deposit License

Agreement for the publication of research data on the institutional repository of TU Berlin*

Transfer of rights to the repository operator

1. I transfer to the Technische Universität Berlin – represented by the Universitätsbibliothek, Hauptabteilung 3 Elektronische Dienste – the non-exclusive, irrevocable right to store electronically the submitted research data and its metadata and to make it publicly available on the Internet. This includes all known and as yet unknown types of use.
2. I give permission to TU Berlin to pass the research data on to other repositories, archives, and if necessary, aggregators for the purpose of digital archiving and/or making it publicly available. This includes, for example, the digital archive of the German National Library, as well as other appropriate subject or regional repositories.
3. I give permission for the publication of the metadata, including the abstract, under the Creative Commons Zero license (CC0, see <https://creativecommons.org/publicdomain/zero/1.0/>).
4. I agree with the creation of copies and the conversion of the research data to other electronic or physical formats for the purpose of archiving while preserving the integrity of the content.

Declaration of copyright and personal rights

1. I assure that I own the copyright of all parts of the research data and that the publication of the research data will not infringe on third-party personal rights or third-party copyright. If alleged or actual infringement of the rights of third parties is asserted, I will notify the repository operator immediately. I exempt the Technische Universität Berlin from any claims of third parties.
2. I assure that the electronic version is identical in form and content to potentially existing print versions.
3. If the document or its metadata contain any personal information (e. g. CV), I agree with their publication. I assure that no personal rights will be violated with the publication.

Conclusion of the contract

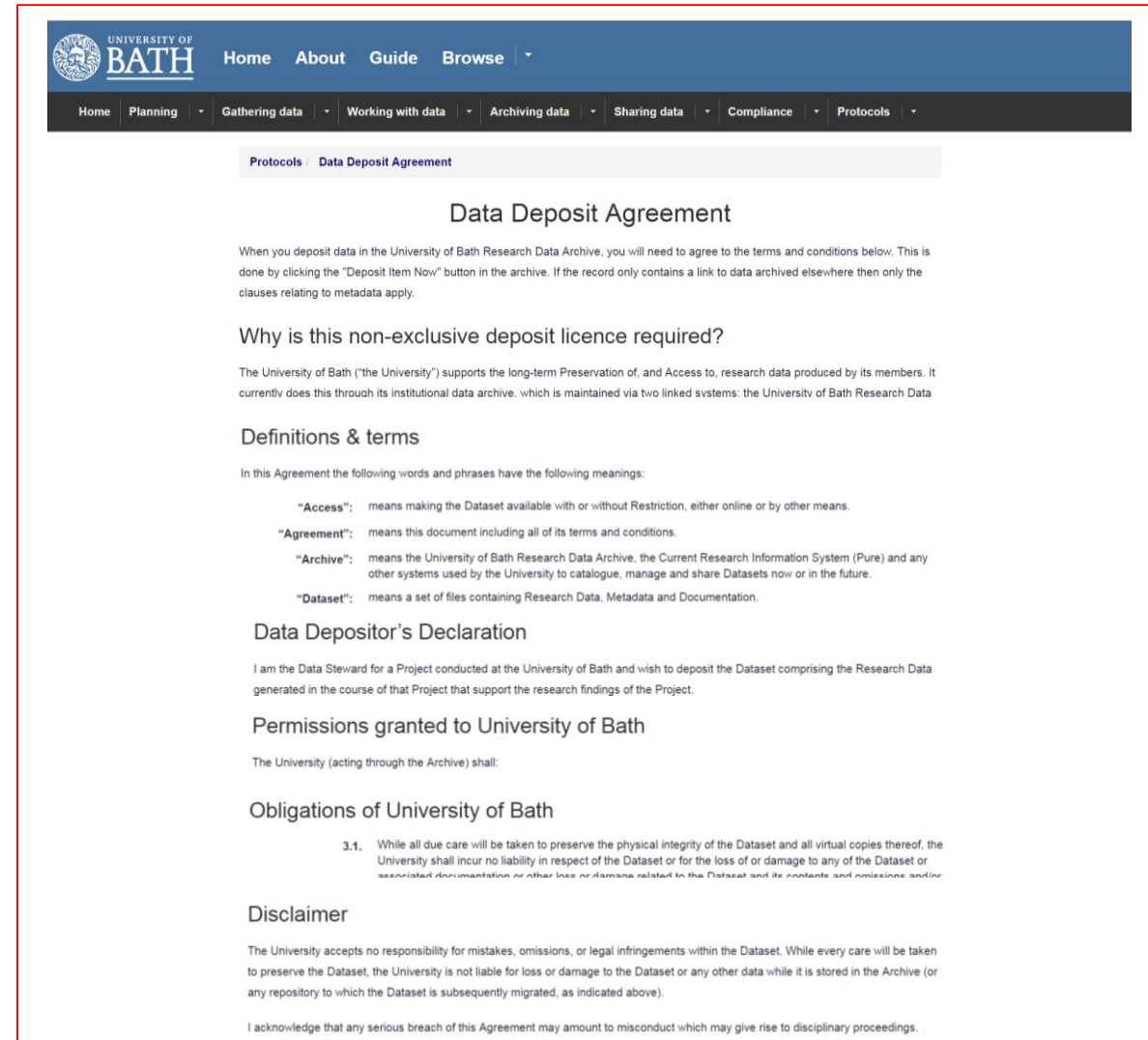
This contract comes into force once the right holder accepts these terms by clicking the link "I grant the licence" immediately prior to uploading the work. The right holder can download, save and print the text of the contract.

* The institutional repository of TU Berlin, DepositOnce, is a service of the [Servicezentrum Forschungsdaten und Publikationen \(SZF\)](#). The Contact partner for all matters concerning the research data submitted to DepositOnce is the Universitätsbibliothek of TU Berlin, Hauptabteilung 3 Elektronische Dienste.

2015-11-23

TU Berlin. Deposit License

14 June, 2017



UNIVERSITY OF BATH Home About Guide Browse

Home Planning Gathering data Working with data Archiving data Sharing data Compliance Protocols

Protocols / Data Deposit Agreement

Data Deposit Agreement

When you deposit data in the University of Bath Research Data Archive, you will need to agree to the terms and conditions below. This is done by clicking the "Deposit Item Now" button in the archive. If the record only contains a link to data archived elsewhere then only the clauses relating to metadata apply.

Why is this non-exclusive deposit licence required?

The University of Bath ("the University") supports the long-term Preservation of, and Access to, research data produced by its members. It currently does this through its institutional data archive, which is maintained via two linked systems: the University of Bath Research Data

Definitions & terms

In this Agreement the following words and phrases have the following meanings:

- "Access": means making the Dataset available with or without Restriction, either online or by other means.
- "Agreement": means this document including all of its terms and conditions.
- "Archive": means the University of Bath Research Data Archive, the Current Research Information System (Pure) and any other systems used by the University to catalogue, manage and share Datasets now or in the future.
- "Dataset": means a set of files containing Research Data, Metadata and Documentation.

Data Depositor's Declaration

I am the Data Steward for a Project conducted at the University of Bath and wish to deposit the Dataset comprising the Research Data generated in the course of that Project that support the research findings of the Project.

Permissions granted to University of Bath

The University (acting through the Archive) shall:

Obligations of University of Bath

- 3.1. While all due care will be taken to preserve the physical integrity of the Dataset and all virtual copies thereof, the University shall incur no liability in respect of the Dataset or for the loss of or damage to any of the Dataset or associated documentation or other loss or damage related to the Dataset and its contents and omissions and/or

Disclaimer

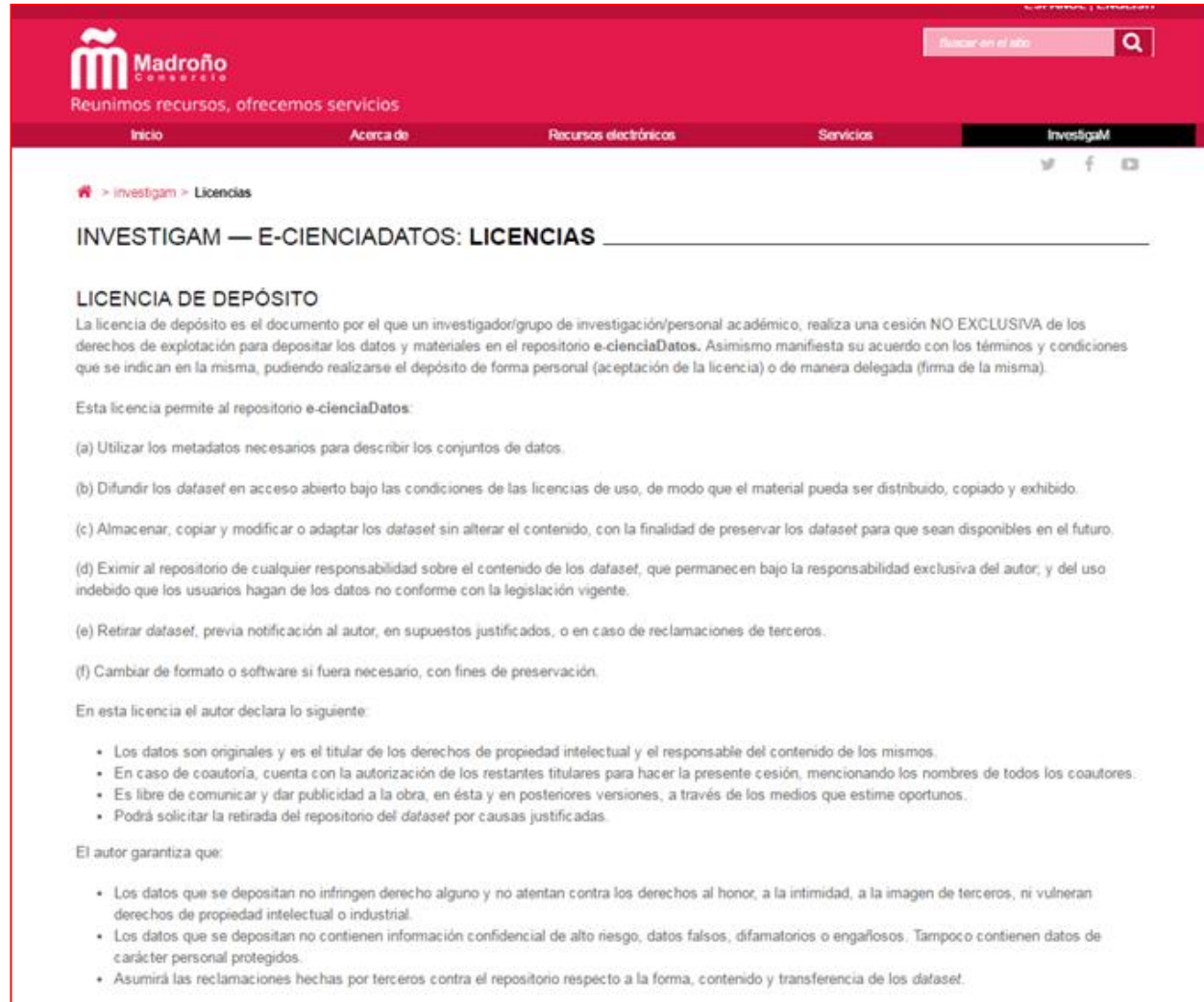
The University accepts no responsibility for mistakes, omissions, or legal infringements within the Dataset. While every care will be taken to preserve the Dataset, the University is not liable for loss or damage to the Dataset or any other data while it is stored in the Archive (or any repository to which the Dataset is subsequently migrated, as indicated above).

I acknowledge that any serious breach of this Agreement may amount to misconduct which may give rise to disciplinary proceedings.

University of Bath. Data Deposit Agreement

<http://researchdata.bath.ac.uk/protocols/data-deposit-agreement/>

Licencias de depósito: Ejemplos



The screenshot shows the website for Madroño Consorcio, specifically the 'InvestigAM' section. The page title is 'INVESTIGAM — E-CIENCIADATOS: LICENCIAS'. The main heading is 'LICENCIA DE DEPÓSITO'. The text explains that this license is a non-exclusive assignment of rights for depositing data and materials in the e-cienciaDatos repository. It lists six conditions (a-f) that the license allows, such as using metadata, open access, preservation, and format changes. It also states what the author declares and guarantees in this license.

INVESTIGAM — E-CIENCIADATOS: LICENCIAS

LICENCIA DE DEPÓSITO

La licencia de depósito es el documento por el que un investigador/grupo de investigación/personal académico, realiza una cesión NO EXCLUSIVA de los derechos de explotación para depositar los datos y materiales en el repositorio e-cienciaDatos. Asimismo manifiesta su acuerdo con los términos y condiciones que se indican en la misma, pudiendo realizarse el depósito de forma personal (aceptación de la licencia) o de manera delegada (firma de la misma).

Esta licencia permite al repositorio e-cienciaDatos:

- Utilizar los metadatos necesarios para describir los conjuntos de datos.
- Difundir los dataset en acceso abierto bajo las condiciones de las licencias de uso, de modo que el material pueda ser distribuido, copiado y exhibido.
- Almacenar, copiar y modificar o adaptar los dataset sin alterar el contenido, con la finalidad de preservar los dataset para que sean disponibles en el futuro.
- Eximir al repositorio de cualquier responsabilidad sobre el contenido de los dataset, que permanecen bajo la responsabilidad exclusiva del autor; y del uso indebido que los usuarios hagan de los datos no conforme con la legislación vigente.
- Retirar dataset, previa notificación al autor, en supuestos justificados, o en caso de reclamaciones de terceros.
- Cambiar de formato o software si fuera necesario, con fines de preservación.

En esta licencia el autor declara lo siguiente:

- Los datos son originales y es el titular de los derechos de propiedad intelectual y el responsable del contenido de los mismos.
- En caso de coautoría, cuenta con la autorización de los restantes titulares para hacer la presente cesión, mencionando los nombres de todos los coautores.
- Es libre de comunicar y dar publicidad a la obra, en ésta y en posteriores versiones, a través de los medios que estime oportunos.
- Podrá solicitar la retirada del repositorio del dataset por causas justificadas.

El autor garantiza que:

- Los datos que se depositan no infringen derecho alguno y no atentan contra los derechos al honor, a la intimidad, a la imagen de terceros, ni vulneran derechos de propiedad intelectual o industrial.
- Los datos que se depositan no contienen información confidencial de alto riesgo, datos falsos, difamatorios o engañosos. Tampoco contienen datos de carácter personal protegidos.
- Asumirá las reclamaciones hechas por terceros contra el repositorio respecto a la forma, contenido y transferencia de los dataset.

Licencias de uso

Licencias de uso: son el modo más efectivo de comunicar al usuario del repositorio los términos en los que puede utilizar, reproducir, modificar y los derechos de propiedad intelectual de los datos que están depositados.

Diferentes tipos de datos  diferentes tipos de licencias

Entre las licencias estándar están:

- Creative Commons
- Open Data Commons
- Dominio Público
- Licencias múltiples



OpenDataCommons.org

Legal tools for Open Data



(Ball, A. "How to license Research Data". DCC How to Guides. Edinburg: DCC, 2014. <http://www.dcc.ac.uk/resources/how-guides>)

Licencias de uso: CC

Licencias Creative Commons: CC organización sin ánimo de lucro que se ha convertido en líder en el desarrollo de herramientas legales para compartir trabajos en internet.

- **Versión 4.0 o posteriores** (las anteriores no son específicas para datos)
- Incluye explícitamente el derecho sui generis de las bases de datos
- Todas las versiones tratan a los datasets y a las bbdd como un todo, lo que genera problemas con colecciones en las que hay trabajos con diferentes tipos de copyright
- Útil para datasets simples
- Datos que se van a usar automáticamente

<http://creativecommons.org>

- (CC BY 4.0):** Permite hacer uso completo del dataset. Mayor reutilización manteniendo el reconocimiento.
- (CC BY-SA 4.0):** Adaptaciones o nuevos trabajos a partir de los datos, compartir con licencia similar. Si se combinan recursos con diferentes licencias, puede ser incompatible.
- (CC BY-ND 4.0):** Más flexible que la anterior. Permite compartir datasets, pero no distribuir obras derivadas o adaptaciones. Es restrictiva en la investigación, pero adecuada cuando el dataset contiene materiales que pueden perder su valor o ser alterados ej.: imágenes
- (CC BY-NC 4.0)**
- (CC BY-NC-SA 4.0)**
- (CC BY-NC-ND 4.0)**



Licencias de uso: ODC

Licencias Open Data Commons: Son más específicas para bases de datos

<https://opendatacommons.org/>

- **Attribution License (ODC-By)**. Reconocimiento para datos/bases de datos. Permite: compartir, copiar, distribuir y usar la bbdd, crear trabajos a partir de la misma y modificar, transformar y construir a partir de la bbdd. Tiene que acompañar el texto:

This {DATA(BASE)-NAME} is made available under the Open Data Commons Attribution License:

<http://opendatacommons.org/licenses/by/{version}>.

- **Open Database License (ODC-ODbL)**. Reconocimiento, compartir igual para datos y bbdd. Permite todo lo anterior, y añade una condición de *copyleft*, cualquier adaptación de la bbdd o trabajos derivados tendrán que tener la misma licencia y permanecer en abierto.

This {DATA(BASE)-NAME} is made available under the Open Database License:

<http://opendatacommons.org/licenses/odbl/1.0/>. Any rights in individual contents of the database are licensed under the Database Contents License: <http://opendatacommons.org/licenses/dbcl/1.0/>

Licencias de uso: Dominio Público

Licencias de Dominio Público:

- Las más permisivas.
- Se renuncia al copyright y a los derechos sobre las bdd.
- Falta de protección ante la competencia desleal.
- Resuelve ambigüedades sobre la reutilización de los datos.

CC0 1.0 Universal (CC0 1.0) Public Domain Dedication



Implica renuncia a los derechos personales sobre el trabajo.

CC Public Domain Mark (CCPDM) Etiqueta de Dominio Público 1.0



Se recomienda para trabajos libres de copyright en todo el mundo, es decir de dominio público. No se debe usar para renunciar a derechos.

Open Data Commons Public Domain Dedication and License (PDDL)

Dedicada específicamente para las bdd y su contenido (datos). Deja fuera a los programas informáticos (software) utilizados para generar la bdd, a las patentes, y a las marcas asociadas a la bdd.

Licencias múltiples:

Permite a los destinatarios de los datos elegir entre una serie de licencias bajo la que quieren usar los datos. Se utiliza sobre todo en el ámbito del software libre: controlar más, y permitir o prohibir el uso del software en aplicaciones comerciales o propietarias y por resolver los problemas de compatibilidad que existen entre licencias *copyleft* (permitir la libre distribución de copias y versiones modificadas de una obra u otro trabajo, exigiendo que los mismos derechos sean preservados en las versiones modificadas).

Otras licencias aprobadas por la Open Source Initiative (OSI):

- [GNU. General Public License](#)
- [MIT license](#)
- [Mozilla Public License 2.0](#)

Licencias de uso

MÁS PERMISIVA	Icono	Nombre y Ejemplos	Permisos	Otros requisitos
Top (Green)		Domínio Público (CC0): Europeana, Figshare, Open Goldberg V.	• Compartir • Copiar • Remezclar • Ganar dinero	• Menciono al autor (en algunas jurisdicciones)
Second (Green)		Reconocimiento (by): PLOS, Saylor.org.	• Compartir • Remezclar • Ganar dinero	• Menciono al autor
Third (Orange)		Reconocimiento – CompartirIgual (by-sa): Wikipedia, Wikimedia, Arduino, P2PU	• Compartir • Remezclar • Ganar dinero	• Menciono al autor • Mantengo la misma licencia (by-sa)
Fourth (Yellow)		Reconocimiento – SinObraDerivada (by-nd): Drupal, Behance, GNU, Free Software Foundation.	• Compartir • Ganar dinero	• Menciono al autor • No hago remezclas
Fifth (Orange)		Reconocimiento – NoComercial (by-nc): Brooklyn Museum, Wired.com Photography	• Compartir • Remezclar	• Menciono al autor • No gano dinero
Sixth (Red)		Reconocimiento – NoComercial – CompartirIgual (by-nc-sa): MIT Open CourseWare	• Compartir • Remezclar	• Menciono al autor • No gano dinero • Mantengo la misma licencia (by-nc-sa)
Bottom (Purple)		Reconocimiento – NoComercial – SinObraDerivada (by-nc-nd): Videos TED Talks, Propublica	• Compartir	• Menciono al autor • No hago remezclas • No gano dinero • Mantengo la misma licencia (by-nc-nd)

Licencias conformes con la *Open Definition* <http://opendefinition.org/licenses>

License	Domain	By	SA	Comments
Creative Commons CCZero (CC0)	Content, Data	N	N	Dedicate to the Public Domain (all rights waived)
Open Data Commons Public Domain Dedication and Licence (PDDL)	Data	N	N	Dedicate to the Public Domain (all rights waived)
Creative Commons Attribution 4.0 (CC-BY-4.0)	Content, Data	Y	N	
Open Data Commons Attribution License (ODC-BY)	Data	Y	N	Attribution for data(bases)
Creative Commons Attribution Share-Alike 4.0 (CC-BY-SA-4.0)	Content, Data	Y	Y	
Open Data Commons Open Database License (ODbL)	Data	Y	Y	Attribution-ShareAlike for data(bases)

Existen asistentes de licencias de datos:

- [European Data Portal. Licensing Assistant](https://www.europeandataportal.eu/en/content/show-license)
- [OER IPR Support Project](http://www.web2rights.com/OERIPRSupport/)

H2020 recomienda el uso de las CC BY 4.0 o CC0.
e-cienciaDatos prefiere utilizar: CC0, PDDL, CC BY, CC BY-SA, pudiendo el investigador solicitar otra.

Licencias de uso: Ejemplos

LICENCIA DE USO

La licencia de uso es un instrumento legal, mediante el cual, el titular de los derechos de propiedad intelectual permite a terceras personas realizar determinados usos de los datos sin infringir dichos derechos. Son una forma simple y estandarizada de otorgar permisos legales.

En estas licencias se establecen las condiciones de uso, distribución y reutilización de los *dataset*.

e-cienciaDatos, siguiendo las directrices de la [Open Knowledge International](#), recomienda el uso de licencias de dominio público como:

- *Creative Commons Zero* (CC0 1.0): implica una liberación de los derechos de propiedad intelectual al ofrecer los datos en dominio público. Se permite su utilización sin ningún tipo de restricción. Bajo esta licencia los datos se pueden copiar, modificar, distribuir y hacer públicos, incluso para fines comerciales, sin solicitar autorización.
- *Open Data Commons, Public Domain Dedication and Licence* (PDDL): es semejante a CC0 pero está dedicada específicamente para las bases de datos y su contenido (datos), en su conjunto o individualmente. Implica el dominio público para datos / bases de datos. Permite compartir, copiar, distribuir y utilizar la base de datos y/o los datos, sin ningún tipo de restricción; crear a partir de los mismos, y cambiarlos o transformarlos.

También se pueden elegir otras licencias como:

- *Reconocimiento* (CC BY 4.0): permite a los usuarios hacer uso completo del *dataset*, incluido el permiso para modificar los datos sin ningún compromiso de compartir los datos resultantes, incluso con fines comerciales. Es la que permite una mayor reutilización, siempre que sea reconocida la autoría. Se recomienda para una máxima difusión y utilización de los datos sujetos a la licencia.
- *Reconocimiento-Compartir Igual* (CC BY-SA 4.0): permite distribuir adaptaciones o crear nuevos trabajos a partir de los datos, siempre que se compartan con una licencia similar. De este modo se trata de garantizar que cualquier obra derivada se comparta.
- *Attribution License* (ODC-By): reconocimiento para datos / bases de datos. Permite compartir (copiar, distribuir y usar la base de datos), crear (generar trabajos a partir de la base de datos) y adaptar (modificar, transformar y construir a partir de la base de datos).
- *Open Data License* (ODC-OdbL): reconocimiento, compartir igual para datos / bases de datos. Permite todo lo anterior y además del reconocimiento, cualquier adaptación de la bases de datos, o trabajos derivados.
- También se podrían aplicar licencias específicas para software, aprobadas por la [Open Source Initiative](#) (OSI): GNU General Public License Versions, MIT license, BSC License, entre otras.

Más información en:

- [Creative Commons](#)
- [Open Data Commons](#)
- [Open Source Initiative](#)

e-cienciaDatos. Licencia de uso

<http://www.consorciomadrono.es/investigam/licencias/>

¿Por qué archivar los datos de investigación?

Aspectos esenciales del archivo de los datos es hacerlos:

- **ACCESIBLES**
 - **RE-USABLES**
- } a lo largo del tiempo

La preservación de los datos de investigación es una fuente importante que se puede poner en riesgo si no se preserva de manera adecuada.

Aspectos claves:

- **La gestión**
- **Compartir los datos**

Asegurarse que los datos estarán disponibles para un uso futuro hace que se convierta en un desafío debido:

- Naturaleza de los objetos digitales
- Heterogeneidad de los datos.

Los mayores riesgos de perder los datos son:

- **Negligencia benigna:** Falta de una estrategia de preservación a largo plazo, por lo que no hay garantías de que los datos se puedan usar años después. Y por otro lado, la tecnología, el software cambia rápidamente y los datos pueden acabar por deteriorarse.
- **Big rot:** Los datos se representan como bits, que pueden degradarse o corromperse con el tiempo.
- **Obsolescencia:** Los medios físicos en los que se almacenan los datos se pueden deteriorar, o quedarse obsoletos con el paso del tiempo. Para ello: se normalizan los ficheros en formatos no propietarios y más actuales.
- **Documentación insuficiente:** Falta de documentación apropiada y metadatos que dificulte su interpretación en un futuro.

Preservar los datos de investigación: Los repositorios

Preservar los datos a lo largo del tiempo requiere:

- **Autenticidad.**
Según la Society of American Archivist:
- **Integridad**
- **Asignación de metadatos**

“La cualidad de ser genuino, sin falsificar, sin manipular y se infiere desde la evidencia interna y externa, incluyendo sus características físicas, estructura, contenido y contexto”. Prevenir accesos no autorizados a los datos, para que no se alteren o copien.

Buenas prácticas de la UK Data Archives:

- Mantener un único fichero master.
- Una única persona responsable del fichero master.
- Registrar todos los cambios en el fichero master.

Es un objeto digital que no ha sido corrompido a lo largo del tiempo, o que sigue manteniendo la misma secuencia de bits de cuando se creó.

Buenas prácticas:

- Realizar copias de seguridad.
- Almacenar ficheros master.
- Verificar los *backups*.

Los repositorios de datos son útiles para encontrar, recuperar, reusar o archivar los datos para un futuro.

Preservar los datos de investigación: Los repositorios

Los datos se pueden depositar en uno o más repositorios:

- **Dominio específico.** Almacenan datos similares. Ej.: Dryad, Gen Bank, o Tree Base.
 - **Institucionales.** Respaldados por las instituciones.
-
- Lo ideal es que los investigadores se dirijan al repositorio al principio del ciclo de vida de los datos.
 - A la hora de escribir el Plan de Gestión de Datos, se puede asesorar en los tipos de ficheros aceptados, estándares de metadatos, documentación.
 - Seleccionar y depositar en un repositorio de confianza: Verificar que cumple con los estándares y realiza buenas prácticas a través de políticas y procedimientos, siendo transparentes (tener políticas de archivo y preservación) y estando bien documentadas.
 - Recurso de interés:

Whyte, A. "Where to keep research data: DCC checklist for evaluating data repositories" v1. Edinburg: DCC, 2014

Preservar los datos de investigación: Los repositorios

¿Dónde encontrar un repositorio de datos?

- **Las Agencias de Financiación**, algunas recomiendan un determinado repositorio ej.: H2020 y Zenodo, otras ofertan diferentes ej.: Wellcome Trust, proporciona una lista en su Guidance for Researches on data sharing
- **Registros de repositorios:**
 - **Re3data.org**. Repositorios de diferentes disciplinas. Gestionado bajo los auspicios de Datacite.
 - **Biosharing.org**. Incluye bases de datos de biociencias. Compilada con el apoyo de Oxford University Press y re3data.org
- **Agregadores de repositorios:**
 - **Research Data Discovery Service (Beta, abril 2017)** <http://researchdiscoveryservice.jisc.ac.uk/dataset>
Pretende descubrir los datos de investigación a lo largo de Gran Bretaña, para aumentar la visibilidad y transparencia.
 - **Research Data Australia** <http://researchdata.andis.org.au>
 - **DataMED** <http://datamed.org>
Motor de búsqueda de datos médicos. También en versión beta.
 - **B2FIND**. Recolector de metadatos de datos procedentes de los centros de la EUDAT y otros repositorios.
 - **OpenAIRE**

Los datos de investigación: Metadatos I

“Los metadatos son documentación sobre los datos. Describen el contenido, calidad, condición y otras características de un dataset. Los metadatos permiten a los datos ser localizados, accesibles y reutilizados.”

(DataONE. Metadata)

<https://www.dataone.org/education-modules>

Tipos de metadatos	Definición	Elementos
DESCRIPTIVOS	Información requerida para localizar, identificar, reutilizar, citar o enlazar un recurso digital.	Autor, título, fechas, resumen, cobertura temporal, cobertura geográfica, etc.
ESTRUCTURALES	Información sobre la estructura del recurso digital de forma que puedan unirse archivos relacionados.	Enlace a la publicación relacionada, parte de, etc.
TÉCNICOS	Información sobre aspectos técnicos del objeto digital	Formato de los datos, hardware/software necesario, versión, acceso, etc.
ADMINISTRATIVOS	Información para la gestión del recurso: cuándo fue creado, derechos de autor, preservación.	Fecha de creación, identificación del creador, <u>copy</u> , licencias, embargo, etc.

(Research Data Netherlands.
RDNL course essentials Data 4 support)

Los datos de investigación: Metadatos II

El valor de los metadatos

Para los creadores

- Evitar duplicados
- Compartir información
- Promover contribuciones al campo

Para los usuarios

- Más fácil encontrar datos
- Más fácil evaluar el contenido de los datos y su aplicación
- Proporcionar información sobre como adquirir, procesar y usar datos

Para la organización

- Asegura la inversión en datos permitiendo su posterior reutilización para otros fines
- Trasciende a la gente y el tiempo para ofrecer permanencia de datos y crear una memoria institucional
- Anuncia la investigación

Los metadatos aseguran la autenticidad, la fiabilidad, la disponibilidad y la integridad de los objetos de información a lo largo del tiempo y posibilitan su gestión y comprensión.

También es necesario gestionarlos de forma precisa.

Los metadatos permiten:

- Localizar el recurso
- Organizar los recursos
- La interoperabilidad entre sistemas

(DataONE. Metadata)

<https://www.dataone.org/education-modules>

Los datos de investigación: Metadatos III

Un **esquema de metadatos** es un conjunto de elementos individuales que se usan para describir los datos. A cada metadato-elemento en el esquema se le asigna un nombre y un significado.

(Research Data Netherlands.
RDNL course essentials Data 4 support)

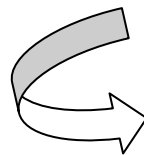
Es recomendable utilizar esquemas genéricos para repositorios multidisciplinares. Los más relevantes son:

- **Dublin Core Metadata Terms:** Es un esquema flexible y aplicable a distintos tipos de documentos.
- **Data Documentation Initiative:** Es modular y extensible. Es el utilizado por Dataverse, el software del repositorio de e-cienciaDatos.
- **DataCite Metadata Schema:** Es jerárquico, flexible y extensible. Es el utilizado por OpenAIRE para recolectar metadatos de repositorios.

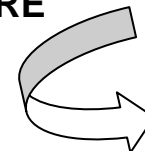
(Rasero, V. "Taller sobre gestión de datos para bibliotecarios: Los metadatos". 2016)

Directrices OpenAIRE para repositorios de datos. 2014 v2.0

Interoperabilidad
entre repositorios



Metadatos compatibles con la
infraestructura de OpenAIRE



Perfil de aplicación basado en el esquema
DataCite 3.1

La Gestión de Datos de Investigación

Marisa Pérez Aliende

Universidad Autónoma de Madrid

mp.aliende@uam.es

Sevilla, 13 de Junio 2017

