

Gestión de datos de investigación y Planes de Gestión de Datos para documentalistas

5-8 de junio de 2023

Horario: 12-14/ 11.30-14/ 11.30-14/ 11.30-14:30/

Tutoría 13 de junio a las 12 h

Remedios Melero. IATA-CSIC, Valencia, España

Tema 1. Acceso abierto y ciencia abierta (5 de junio, 12:00-14:00)

- El acceso abierto a las publicaciones y a los datos de investigación
- El ciclo de vida de los datos
- Políticas sobre los datos de investigación

Tema 2. Fundamentos de la gestión de datos de investigación (5 de junio, 12:00-14:00)

- Los datos de investigación: Definiciones y tipología
- Formatos
- Datos abiertos. Datos FAIR

Tema 3. Como citar y enriquecer los datos. Aspectos legales y éticos (6 junio, 11:30-14:00)

- Licencias para compartir datos
- Cómo citar datos. Cómo enriquecer los datos
- Normativa GDPR. Anonimización de datos

Tema 4. Depósito y localización de datasets (7 junio, 11:30-14:00)

- Depósito de datasets
- Repositorios de datos
- Localización de datos

Tema 5. El plan de gestión de datos (PGD) (8 junio, 11:30-14:30)

- Qué es un PGD
- Recursos y herramientas para generar PGD

Tema 3. Como citar y enriquecer los datos. Aspectos legales y éticos (6 junio, 11:30-14:00)

- Licencias para compartir datos
- Cómo citar datos. Cómo enriquecer los datos
- Normativa GDPR. Anonimización de datos

Cómo licenciar

Licencia

Documento por el cual el licenciador (de origen el autor) establece los términos de la explotación/uso de la obra (reproducción, distribución, comunicación pública, transformación).

El caso de las licencias Creative Commons.....



Tres “Capas” de las licencias



Creative Commons Attribution 4.0 International Public License

By exercising the Licensed Rights (defined below), You accept and agree to be bound by the terms and conditions of this Creative Commons Attribution 4.0 International Public License ("Public License"). To the extent this Public License may be interpreted as a contract, You are granted the Licensed Rights in consideration of Your acceptance of these terms and conditions, and the Licensor grants You such rights in consideration of benefits the Licensor receives from making the Licensed Material available under these terms and conditions.....

Attribution 4.0 International (CC BY 4.0)

This is a human-readable summary of (and not a substitute for) the [license](#). [Disclaimer](#).

You are free to:

- Share** — copy and redistribute the material in any medium or format
- Adapt** — remix, transform, and build upon the material for any purpose, even commercially.

The licensor cannot revoke these freedoms as long as you follow the license terms.

Under the following terms:

- Attribution** — You must give [appropriate credit](#), provide a link to the license, and [indicate if changes were made](#). You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.

1 Data



CC0

Creative Commons Zero is ideal for openly sharing data – it has no restrictions on reuse whatsoever. While CC0 contains no requirement for attribution, citing CC0 datasets is widely accepted and expected in research.

Other CC licenses

If a CC0 license is not suitable for your data, consider the following CC licenses - all of which require attribution and further restrictions:

CC-BY Prevents others from applying legal restrictions beyond the terms of the license to the licensed dataset.

CC BY-SA Requires outputs derived from licensed dataset to also be licensed as CC BY-SA.

CC BY-NC Prevents the licensed data from being used for commercial purpose.

CC BY-ND Prevents the licensed data from being modified.

CC BY-NC-ND Prevents the licensed data from being used for commercial purposes or modified.

CC BY-NC-SA Prevents the licensed data from being used for commercial purposes, and requires outputs derived from licensed dataset to also be licensed as CC BY-SA






Tres “Capas” de las licencias



Caution!

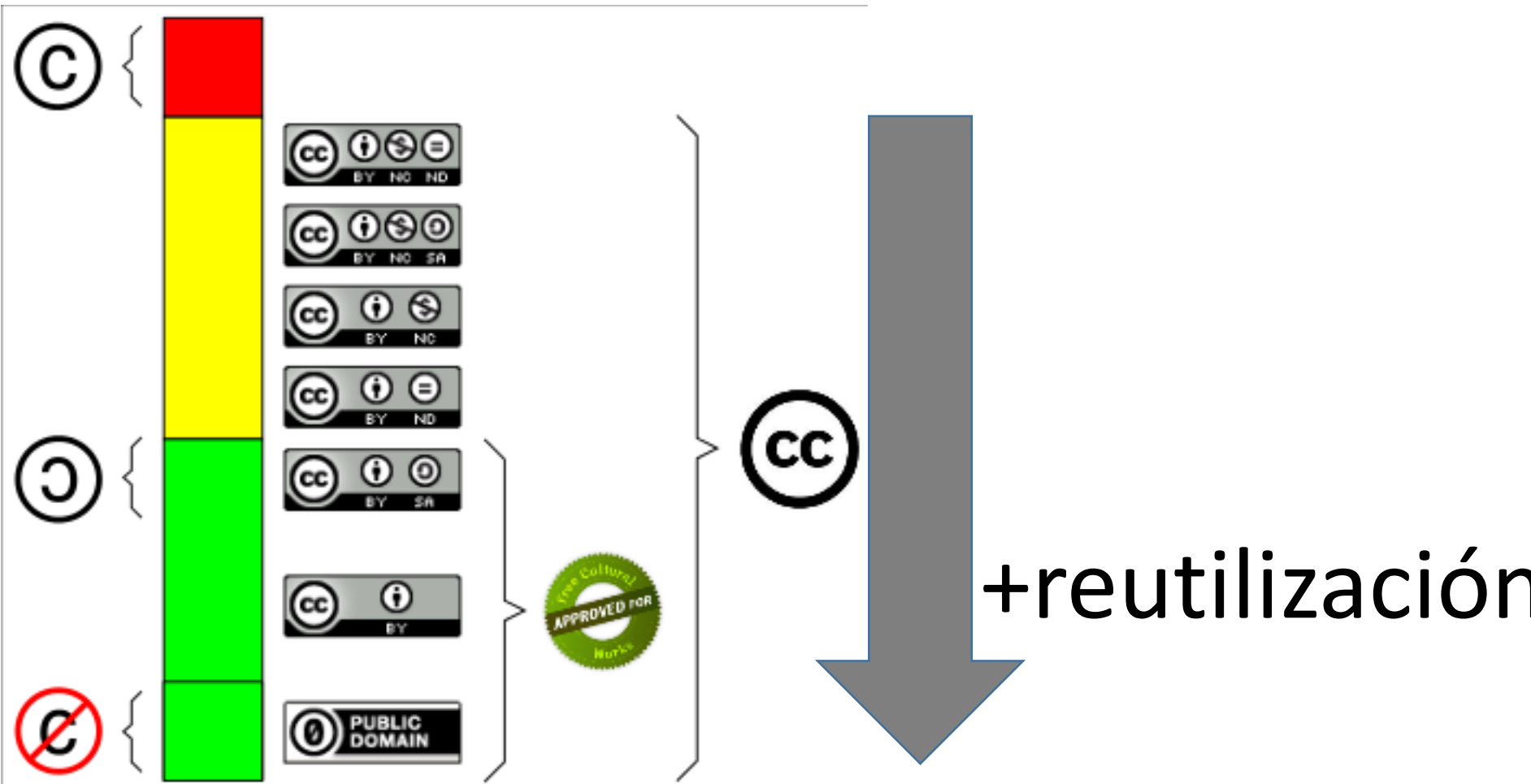
NC, ND and SA licenses have implications for reuse and interoperability. We suggest using a license that keeps your data as open as possible and as closed as necessary.

Licencias de Creative Commons

			¿Usarlas de forma comercial?	¿Modificarlas?
Atribución	BY			
Compartir bajo la Misma Licencia	BY-SA	 		Solo bajo una licencia similar
Sin Obras Derivadas	BY-ND	 		
No comercial	BY-NC	 		Solo con propósitos no comerciales
No comercial bajo la Misma Licencia	BY-NC-SA	  		Solo con propósitos no comerciales
No comercial Sin Obras Derivadas	BY-NC-ND	  		

En todas es necesario darle el crédito al autor

norfipc.com





[Home](#) > [Chooser](#)

LICENSE CHOOSER

Follow the steps to select the appropriate license for your work. This site does not store any information.

1 Do you know which license you need?

- Yes. I know the license I need.
- No. I need help selecting a license.

NEXT

2 Attribution

3 Commercial Use

4 Derivative Works

5 Sharing Requirements

6 Confirm that CC licensing is appropriate

Cómo asignar estas licencias

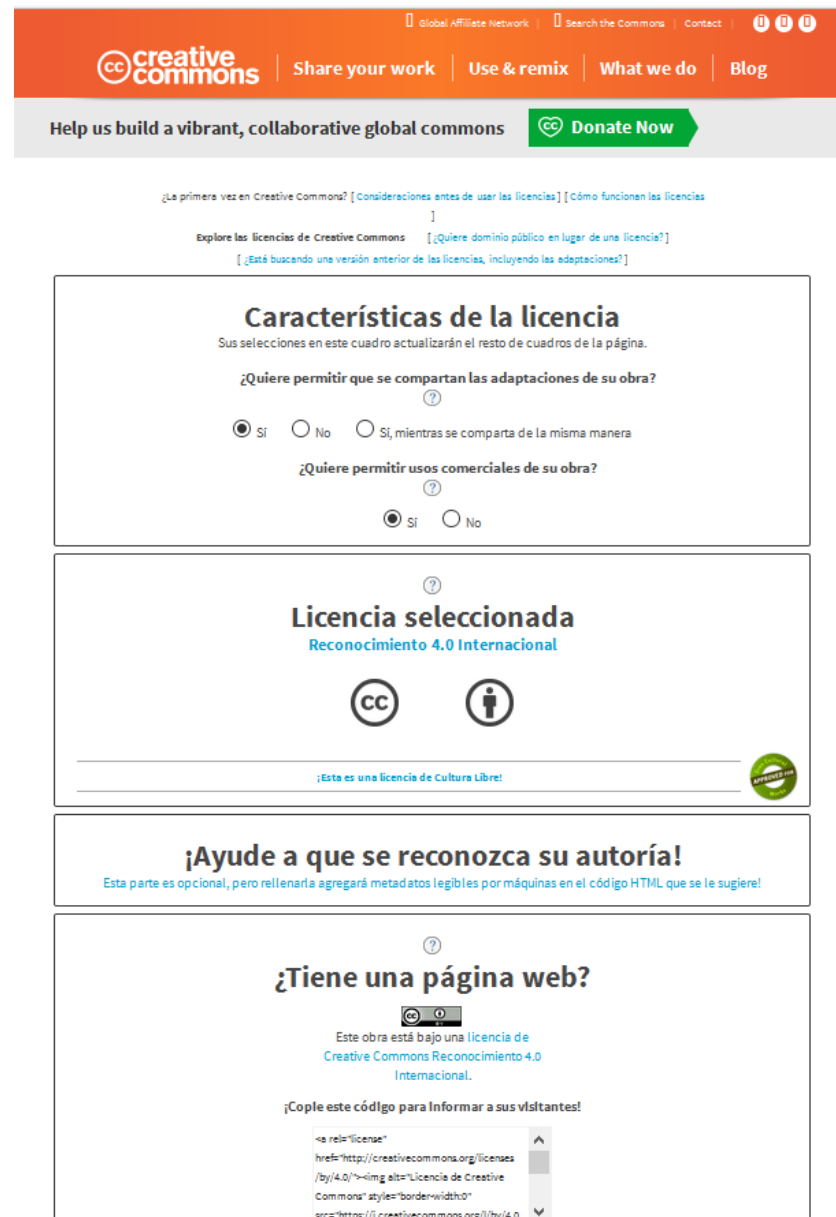
Public Domain Dedication and License (PDDL)

This {DATA(BASE)-NAME} is made available under the Public Domain Dedication and License version v1.0 whose full text can be found at <http://opendatacommons.org/licenses/pddl/>

NB: you may also wish to apply some complementary Community Norms Open Database License

This {DATA(BASE)-NAME} is made available under Open Database License whose full text can be found at <http://opendatacommons.org/licenses/odbl/>. Any rights in individual contents of the database are licensed under the Database Contents License whose text can be found <http://opendatacommons.org/licenses/dbcl/>

<https://opendatacommons.org/licenses/>



<https://creativecommons.org/choose/>

Licencias recomendadas por el OKF conformes a su definición de abierto (atribución y compartir igual) . Open Data Commons y Creative Commons

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)

<http://opendefinition.org/licenses/>

Choose a License

Answer the questions or use the search to find the license you want

[Start again](#) ← →


Is your data within the scope of copyright and related rights?

[Yes](#) [No](#)

Search for a license...



Public Domain Mark (PD)

The work identified as being free of known restrictions under copyright law, including all related and neighboring rights.

[Publicly Available](#) 



Public Domain Dedication (CC Zero)

CC Zero enables scientists, educators, artists and other creators and owners of copyright- or database-protected content to waive those interests in their works and thereby place them as completely as possible in the public domain, so that others may freely build upon, enhance and reuse the works for any purposes without restriction under copyright or database law.

[Publicly Available](#)   [OPEN DATA](#)

Creative Commons Attribution (CC-BY)

Modified description ...

[Publicly Available](#)   [OPEN DATA](#)

Creative Commons Attribution-ShareAlike (CC-BY-SA)

This creative commons license is very similar to the regular Attribution license, but requires you to release all

Cómo citar ficheros de datos

Por qué y cómo citar los datos (para autores)

¿Por qué?

1. Facilita la reproducibilidad y la validación de los resultados, permite la reutilización de los datos.
2. Proporciona crédito a los generadores de datos.
3. Las publicaciones vinculadas a datos disponibles públicamente se han asociado a un aumento de las citas.
4. Mejora la conectividad y el seguimiento de la procedencia de los datos descritos en las publicaciones.
5. Las agencias de financiación (ej Comisión Europea), los editores (Plos) y las instituciones (p.e. CSIC) exigen cada vez más que se compartan los datos.

¿Cómo?

1. En el caso de los datos primarios: determine un repositorio apropiado a largo plazo para archivar los datos. Su editor debe proporcionar acceso a una lista de repositorios de archivo aceptables.
2. Deposite sus datos y obtenga del repositorio un número de acceso o un DOI del conjunto de datos.
3. Para los datos secundarios: cite lo que utiliza. Cuando utilice los datos de otros, cite tanto la literatura revisada por pares como los conjuntos de datos utilizados
4. Incluya citas de datos "formales" siempre que sea posible: Cuando los conjuntos de datos tienen identificadores formales y estables o números de acceso, deben incluirse en la lista de referencias principal.
5. Sea lo más completo posible, pero no se invente los metadatos: Si un registro de datos no tiene un autor/creador o un título claros, no se invente uno.
6. Consulte la Guía para autores de su editor para dar formato a la referencia de su conjunto de datos.

Cousijn, H., Kenall, A., Ganley, E. *et al.* A data citation roadmap for scientific publishers. *Sci Data* **5**, 180259 (2018). <https://doi.org/10.1038/sdata.2018.259>



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:

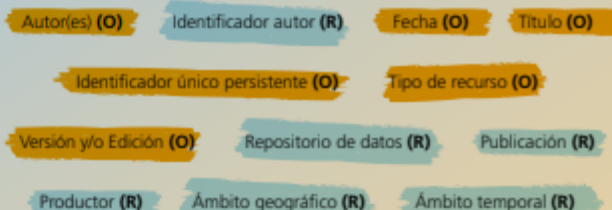
- Se debe facilitar la identificación, localización y el acceso a los datos mediante un identificador único y persistente (DOI, Handle, etc.).
- Cada conjunto y subconjunto de datos (dataset) debe citarse de forma independiente.
- Las citas de los datos utilizados han de aparecer en la sección de referencias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



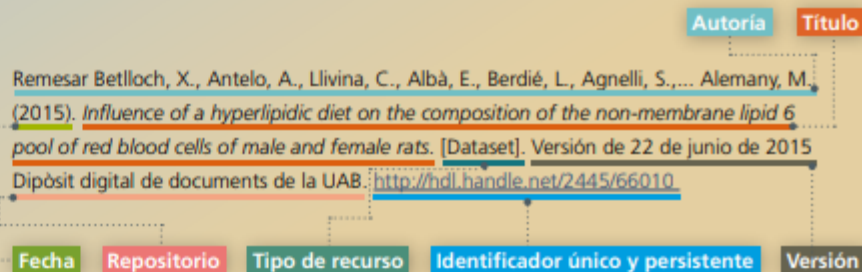
Enlaza los datos con los documentos resultado de investigación y viceversa, y crea las referencias bibliográficas de los mismos.

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.).



Ejemplo de cita estilo APA



El personal de tu Biblioteca te puede asesorar



Ejemplo de formato de cita al usar el DOI Citation formatter

<http://citation.crosscite.org/>

DOI Citation Formatter

Paste your DOI:

10.5061/dryad.7dj7t

For example 10.1145/2783446.2783605

Select Formatting Style:

apa

Begin typing (e.g. Chicago or IEEE.) or use the drop down menu.

Select Language and Country:

en-US

Begin typing (e.g. en-GB for English, Great Britain) or use the drop down menu.

Format

Wang, M., Gu, D., Du, M., Xu, Z., Zhang, S., Zhu, L., ... Chen, J. (2016). Data from: Common genetic variation in ETV6 is associated with colorectal cancer susceptibility [Data set]. Dryad Digital Repository. <https://doi.org/10.5061/dryad.7dj7t>

Copy to clipboard

Do you want to integrate this service? Check the [Documentation](#)

DOI Registration Agencies



<https://data.research.cornell.edu/content/data-citation>

Cómo “enriquecer” nuestros datos



MeSH on Demand identifies MeSH® terms in your submitted text (abstract or manuscript). MeSH on Demand also lists PubMed similar articles relevant to your submitted text.

Search Reset Help/FAQ Features

Enter text to be processed here - then click Search

<https://meshb-prev.nlm.nih.gov/MeSHonDemand>



MeSH on Demand identifies MeSH® terms in your submitted text (abstract or manuscript). MeSH on Demand also lists PubMed similar articles relevant to your submitted text.

Search Reset Help/FAQ Features

[Chagas disease](#) also known as [American trypanosomiasis](#) is a potentially life-threatening illness caused by the [protozoan parasite](#) [Trypanosoma cruzi](#). About 6.7 million people worldwide are estimated to be infected with [T. cruzi](#). The disease is found mainly in endemic areas of 21 continental Latin American countries, 1 where it has been mostly transmitted to [humans](#) and other [mammals](#) by contact with [faeces](#) or urine of triatomine bugs vector-borne known as [kissing bugs](#) among many other popular names depending on the geographical area.

Start PubMed Search

Export Data

MeSH Terms

- [i Humans](#)
- [i Animals](#)
- [i Trypanosoma cruzi](#)
- [i Parasites](#)
- [i Chagas Disease](#)
- [i Triatoma](#)
- [i Mammals](#)
- [i Feces](#)



Tesoro de la UNESCO

Lengua del contenido español

Alfabéticamente

Jerarquía

Grupos

A Á B C D E É F G H I J K L M
N O Ó P Q R S T U V W X Y Z

[Abadía](#) → [Edificio religioso](#)
[Abandono de menores](#) → [Niño abandonado](#)
[Abandono escolar](#) → [Deserción escolar](#)
[Abandono infantil](#) → [Niño abandonado](#)
[Abastecimiento alimenticio](#) → [Suministro de alimentos](#)
[Abastecimiento de agua](#)
[Abastecimiento de energía](#)
[Abastecimiento de víveres](#) → [Suministro de alimentos](#)
[Abecedario](#) → [Alfabeto](#)
[Abogacía](#) → [Profesión jurídica](#)
[Abogado](#) → [Profesión jurídica](#)
[Abono](#) → [Fertilizante](#)
[Aborto](#)
[Aborígenes](#) → [Población indígena](#)
[Abreviatura](#)
[Absentismo](#) → [Permiso](#)
[Absolutismo](#) → [Totalitarismo](#)
[Absentismo escolar](#) → [Ausencia injustificada](#)
[Abuso de autoridad](#) → [Opresión](#)
[Abuso de drogas](#) → [Toxicomanía](#)
[Abuso de los derechos humanos](#) → [Violación de los derechos humanos](#)
[Abuso de menores](#)
[Abuso de poder](#) → [Opresión](#)
[Abuso sexual](#)
[Academia de ciencias](#)
[Acatamiento de la ley](#) → [Aplicación de la ley](#)
[Accesibilidad a la información](#) → [Acceso a la información](#)
[Acceso a la alimentación](#) → [Derecho a la alimentación](#)
[Acceso a la cultura](#) → [Derechos culturales](#)

Información del vocabulario

TÍTULO	Tesoro de la UNESCO
DESCRIPCIÓN	El Tesoro de la UNESCO es una lista controlada y estructurada de términos para el análisis temático y la búsqueda de documentos y publicaciones en los campos de la educación, cultura, ciencias naturales, ciencias sociales y humanas, comunicación e información. Continuamente ampliada y actualizada, su terminología multidisciplinaria refleja la evolución de los programas y actividades de la UNESCO.
DC:IDENTIFIER	http://vocabularies.unesco.org/thesaurus
EDITOR	UNESCO
DC:RIGHTSHOLDER	UNESCO
DERECHOS	CC-BY-SA
LICENCIA	http://creativecommons.org/licenses/by-sa/3.0/igo/
CREADO	sábado, 1 de enero de 1977 00:00:00
ÚLTIMA MODIFICACIÓN	viernes, 7 de abril de 2023 10:16:17
MATERIA	http://id.loc.gov/authorities/subjects/sh85029027 http://id.loc.gov/authorities/subjects/sh85062913 Social sciences http://id.loc.gov/authorities/subjects/sh85034755

<https://obofoundry.org/>

Open Biological and Biomedical Ontology Foundry

Community development of interoperable ontologies for the biological sciences

Learn about OBO best practices and community resources

- [OBO Foundry principles](#)
- [OBO tutorial](#)
- [Ontology browsers, tutorials, and tools](#)

Participate

- [Code of Conduct](#)
- [Join the OBO mailing list and the OBO Community Slack workspace](#)
- [OBO Foundry Operations and Working Groups](#)
- [Submit bug reports or suggestions for improvement via GitHub](#)
- [Submit your ontology to be considered for inclusion in the OBO Foundry](#)

OBO Library: find, use, and contribute to community ontologies

Download table as: [[YAML](#) | [JSON-LD](#) | [RDF/Turtle](#)]

Search Table

Ontology Domains:

Group By Domain

Hide Inactive

Hide Obsolete

Upper

ID ^	Title ^	Description	Quick Access	Re-Use ^	Social
------	---------	-------------	--------------	----------	--------



ONTOLOGY SEARCH

Home | Ontologies | Documentation | About

List of all ontologies in OLS

Show 10 entries

Search:



Ontology Name	Short name	Description	Loaded	Action
Agronomy Ontology	AGRO	AgrO is an ontology for representing agronomic practices, techniques, variables and related entities	Thu Mar 03 00:31:45 GMT 2022	Search Terms Properties Individuals Download
Allotrope Merged Ontology Suite	AFO	Allotrope Merged Ontology Suite	Tue May 28 16:48:55 BST 2019	Search Terms Properties Individuals Download
Alzheimer's Disease Ontology (ADO)	ADO	Alzheimer's Disease Ontology is a knowledge-based ontology that encompasses varieties of concepts related to Alzheimer'S Disease, fundamentally structured by upper level Basic Formal Ontology(BFO). This Ontology is enriched by the interrelational entities that demonstrate the nextwork of the understanding on Alzheimer's disease and can be readily applied for text mining.	Tue Jun 28 17:23:14 BST 2022	Search Terms Properties Individuals Download
Amphioxus Development and Anatomy Ontology (AMPHX)	AMPHX	An ontology for the development and anatomy of Amphioxus (Branchiostoma lanceolatum).	Tue Nov 09 05:14:38	Search Terms Properties

Ejemplo de ontología en árbol


OLS / Anatomical Entity Ontology **AEO** / **CARO:0000000**  Copy





anatomical entity

 http://purl.obolibrary.org/obo/CARO_0000000  Copy

Biological entity that is either an individual member of a biological species or constitutes the structural organization of an individual member of a biological species. [http://purl.obolibrary.org/obo/CARO_MAH]

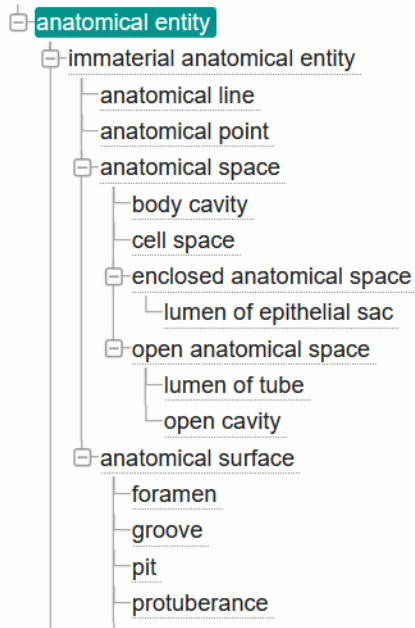
 Tree view

 Term mappings

 Graph view

Reset tree

Show all siblings



Term information

has alternative id

- EHDAA2:0003000
- UBERON:0001062

has obo namespace

- <http://www.xspan.org/obo.owl#>

id

- CARO:0000000

Term relations

Subclass of:

- [Thing](#)

RESEARCH DATA MANAGEMENT SERVICE GROUP

Comprehensive Data Management Planning & Services

Guide to writing "readme" style metadata

A readme file provides information about a data file and is intended to help ensure that the data can be correctly interpreted, by yourself at a later date or by others when sharing or publishing data. [Standards-based metadata](#) is generally preferable, but where no appropriate standard exists, for internal use, writing "readme" style metadata is an appropriate strategy.

 Want a template? **Download one** and adapt it for your own data!

- [Best practices](#)
- [Recommended content](#)
 - [General information](#)
 - [Data and file overview](#)
 - [Sharing and access information](#)
 - [Methodological information](#)
 - [Data-specific information](#)
- [References](#)
- [Related information](#)

```
[This DATSETNAMEREADME.TXT file was generated on YYYY-MM-DD by NAME  
<help text is included in angle brackets, and can be deleted before saving>
```

GENERAL INFORMATION

1. Title of Dataset:
2. Author Information
 - A. Principal Investigator Contact Information
 - Name:
 - Institution:
 - Address:
 - Email:
 - B. Associate or Co-investigator Contact Information
 - Name:
 - Institution:
 - Address:
 - Email:
 - C. Alternate Contact Information
 - Name:
 - Institution:
 - Address:
 - Email:
3. Date of data collection (single date, range, approximate date) <suggested format YYYY-MM-DD>:
4. Geographic location of data collection <latitude, longitude, or city/region, State, Country, as appropriate>:
5. Information about funding sources that supported the collection of the data:



Plantilla readme.txt

Readme.txt ([acceso](#) [descarga](#)) es una plantilla para que rellenen los investigadores que quieran depositar un dataset en e-cienciaDatos.

Todos los conjuntos de datos han de estar documentados para ser comprensibles y reutilizables: autoría, título, descripción, metodología, proyectos financiadores, cobertura temporal y geográfica, derechos de uso y privacidad, etc.

Los gestores del repositorio e-cienciaDatos utilizarán la información proporcionada en la plantilla readme.txt para:

- cumplimentar los metadatos estandarizados que describen el dataset y que son recolectados por diversos servidores científicos.
- elaborar un fichero readme.txt que se incluirá junto al resto de ficheros del dataset.

Formulario "readme" en e-cienciaDatos

```
-----  
INFORMACIÓN GENERAL  
-----  
1. Título del dataset  
  
2. Contacto  
  
   Investigador/a de contacto  
   Nombre:  
   Filiación:  
   Correo electrónico:  
   ORCID:  
  
3. Descripción del proyecto  
  
4. Descripción del dataset  
  
5. Notas  
  
6. Fecha de depósito de los ficheros  
  
7. Fecha de creación de los ficheros  
  
8. Idioma
```

Ejemplos de datasets y de sus “leame” (readme) o de su usencia

<https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/HQLI7W>

<https://dataverse.harvard.edu/file.xhtml?persistentId=doi:10.7910/DVN/K2KDZU/FZEL3N>

Tratamiento de datos personales.

Anonimización de datos

DECLARACIÓN DE HELSINKI DE LA AMM – PRINCIPIOS ÉTICOS PARA LAS INVESTIGACIONES MÉDICAS EN SERES HUMANOS



Adoptada por la
18ª Asamblea Médica Mundial, Helsinki, Finlandia, junio 1964
y enmendada por la
29ª Asamblea Médica Mundial, Tokio, Japón, octubre 1975
35ª Asamblea Médica Mundial, Venecia, Italia, octubre 1983
41ª Asamblea Médica Mundial, Hong Kong, septiembre 1989
48ª Asamblea General Somerset West, Sudáfrica, octubre 1996

21st marzo 2017

Tipos de políticas

Declaración

nombre

<https://www.wma.net/es/policies-post/declaracion-de-helsinki-de-la-amm-principios-eticos-para-las-investigaciones-medicas-en-seres-humanos/>

El término **datos personales** se refiere a cualquier información que pueda usarse para identificar individuos vivos (y, a veces, incluso fallecidos). Por ejemplo, su nombre o fecha de nacimiento. Algunos datos personales, como el origen, la opinión política, las creencias religiosas, la salud, la afiliación sindical o la orientación sexual de una persona, se clasifican como **datos personales confidenciales**. Si se manejan datos personales o confidenciales como parte de su investigación, debemos asegurarnos de que están protegidos.



<https://www.cessda.eu/Training/Training-Resources/Library/Data-Management-Expert-Guide/5.-Protect/Ethics-and-data-protection>

Cuestiones previas a la recopilación de datos personales

¿Qué datos personales recopilará?

Identifique qué datos personales se recopilarán y utilizarán para su investigación. Por ejemplo, ¿Reunirá los nombres y direcciones de los participantes de la investigación?

¿Cómo se cubrirán los costes de anonimización?

Discuta si realmente necesita recopilar datos personales para llevar a cabo su investigación. Tenga en cuenta que cualquier información personal innecesaria recopilada y agregada a su conjunto de datos probablemente deberá eliminarse o anonimizarse más tarde si planea compartirla. Esto puede ser costoso. En algunos casos, la limpieza de datos puede costar más que la recolección de datos en sí.

¿Realmente necesita recopilar datos personales?

Tómese siempre el tiempo necesario para volver a considerar si hay formas en que puede recopilar datos de forma anónima. Esto podría ahorrarle muchos dolores de cabeza más tarde. Así que solo se compromete a recopilar datos personales si es esencial para su investigación.

Al recopilar y / o manejar datos personales, los investigadores deben seguir una serie de principios que incluyen:

Transparencia: procesamiento de datos personales "de manera legal, justa y transparente"

Minimización de datos: el uso de los datos se limitará al propósito de la investigación respectiva.

Precisión: los datos inexactos deben ser "borrados o rectificados sin demora".
Integridad y confidencialidad: los datos deben estar protegidos por medidas de seguridad adecuadas (técnicas y organizativas).

Con el nuevo Reglamento General de Protección de Datos (GDPR), la Unión Europea proporciona un marco legal para la protección de datos dentro de la UE, así como para la exportación de estos datos fuera de sus fronteras.

En cuanto a la recopilación de datos personales, el GDPR incluye una exención para la investigación:

- Si se trata de "aspectos de interés público, investigación científica o histórica o fines estadísticos" (Art. 5.1 2016/679 / UE)

<https://www.privacy-regulation.eu/es/89.htm>

<https://www.privacy-regulation.eu/es/5.htm>

- Si “el interesado ha dado su consentimiento para el procesamiento de sus datos personales para uno o más propósitos específicos” (Art. 6.1 2016/679 / UE).

Reglamento General de Protección de Datos (RGPD) de 28 de mayo de 2018 [Ley Orgánica 3/2018](#), de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales.

LOS DERECHOS QUE TIENES PARA PROTEGER TUS DATOS PERSONALES

EL 25 DE MAYO DE 2018 SE APLICA EL REGLAMENTO EUROPEO DE PROTECCIÓN DE DATOS Y ES IMPORTANTE QUE CONOZCAS CUÁLES SON TUS DERECHOS



En resumen, a recolección de datos de investigación que incluyan datos sensibles con tratamientos de poca complejidad

Requieren:

- **Consentimiento inequívoco (no vale genérico)**
- **Base legal**
- **Finalidad (¿para qué?)**
- **Análisis de riesgos**
- **Transparencia en la información a los interesados**
- **Medidas de seguridad y preservación (en caso necesario)**

El **consentimiento informado** es el procedimiento mediante el cual se garantiza que el sujeto ha expresado voluntariamente su intención de participar en la investigación, después de haber comprendido la información que se le ha dado, acerca de los objetivos del estudio, los beneficios, las molestias, los posibles riesgos y las alternativas, sus derechos y responsabilidades

(https://es.m.wikipedia.org/wiki/Consentimiento_informado)



Consentimiento Informado:
¿Lo firmas o lo lees?

Formulario de consentimiento

El formulario de consentimiento debe estar escrito en un lenguaje sencillo, sin jerga, y debe permitir que el participante responda claramente a puntos como:

- El participante ha leído y entendido la información sobre el proyecto.
- El participante ha tenido la oportunidad de hacer preguntas.
- El participante acepta voluntariamente participar en el proyecto.
- El participante entiende que puede retirarse en cualquier momento sin dar razones y sin penalización.
- Cómo se protegerá la confidencialidad, p. si se usarán nombres reales o seudónimos, cómo se anonimizarán los datos, etc.
- Términos de consentimiento separados para datos que pueden tener diferentes riesgos de divulgación, p. transcripciones de entrevistas anónimas, grabaciones de audio, videos, fotos
- Cómo se utilizarán los datos en las publicaciones
- Si los participantes están de acuerdo con el archivo y la reutilización de datos
- Firmas y fechas de firma para el participante y el investigador.
- El participante debe recibir una copia del formulario y el investigador debe conservar el original firmado.

<https://ukdataservice.ac.uk/learning-hub/research-data-management/ethical-issues/consent-for-data-sharing/>



Consent Form for [name of project]

Please tick the appropriate boxes

Yes No

Taking Part

I have read and understood the project information sheet dated DD/MM/YYYY.

I have been given the opportunity to ask questions about the project.

I agree to take part in the project. Taking part in the project will include being interviewed and recorded (audio or video).¹

I understand that my taking part is voluntary; I can withdraw from the study at any time and I do not have to give any reasons for why I no longer want to take part.

Use of the information I provide for this project only

I understand my personal details such as phone number and address will not be revealed to people outside the project.

I understand that my words may be quoted in publications, reports, web pages, and other research outputs.

Please choose **one** of the following two options:

I would like my real name used in the above

I would **not** like my real name to be used in the above.

Use of the information I provide beyond this project

I agree for the data I provide to be archived at the UK Data Archive.²

I understand that other authenticated researchers will have access to this data only if they agree to preserve the confidentiality of the information as requested in this form.

I understand that other authenticated researchers may use my words in publications, reports, web pages, and other research outputs, only if they agree to preserve the confidentiality of the information as requested in this form.

So we can use the information you provide legally

I agree to assign the copyright I hold in any materials related to this project to [name of researcher].

Name of participant [printed] Signature Date

Researcher [printed] Signature Date

Project contact details for further information: Names, phone, email addresses, etc.

Notes:

1. Other forms of participation can be listed.
2. More detail can be provided here so that decisions can be made separately about audio, video, transcripts, etc.

CONSENTIMIENTO INFORMADO

El proyecto de investigación para el cual le pedimos su participación se titula:

.....

Para que usted pueda participar en este estudio es necesario contar con su consentimiento, y que conozca la información básica necesaria para que dicho consentimiento pueda considerarse verdaderamente informado. Por ello, le ruego que lea detenidamente la siguiente información. Si tuviera alguna duda exprésela antes de firmar este documento.

1 INFORMACIÓN BÁSICA QUE DEBE CONOCER

- a) **Objetivo del estudio:**
- b) **Metodología a utilizar para el estudio, tipo de colaboración que se espera de usted y duración de dicha colaboración:** Al colaborar con esta investigación, le pedimos que participe en un entrevista online de y conteste a las preguntas que le realizará el investigador relacionadas con..... La duración prevista de la reunión es, de alrededor de una hora.
- c) **Beneficios que se esperan obtener con la investigación:** Se espera que
- d) **Posibilidad de no participación, retirada en cualquier momento y consecuencias:** La participación en este estudio es completamente voluntaria. Usted puede retirarse del estudio en cualquier momento firmando la revocación del consentimiento que se incluye al final del documento. Tanto si decide no participar como si decide retirarse una vez iniciada la participación, ello no tendrá ninguna consecuencia negativa para usted, y será aceptada sin problemas por el investigador.
- e) **Financiación:**
- f) **¿Qué institución lo realiza?:**
- g) **Gratuidad por la participación:** Los participantes en estas entrevistas no obtendrán ninguna compensación económica por la participación en este estudio.
- h) **Previsión de uso posterior de los resultados:** Los resultados se utilizarán con fines de investigación y publicación científica.

2 COMPROMISO DE CONFIDENCIALIDAD

- a) **Medidas para asegurar el respeto a la vida privada y a la confidencialidad de los datos personales:**
- En materia de protección de datos, se dará cumplimiento a las disposiciones contempladas en el REGLAMENTO (UE) 2016/679 DEL PARLAMENTO EUROPEO Y DEL CONSEJO de 27 de abril de 2016 relativo a la protección de las personas físicas en lo que respecta al tratamiento de datos personales, traspuestas en la Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales.
- b) **Medidas para acceder a la información relevante para usted que surjan de la investigación o de los resultados totales:** Sepa que tiene derecho a acceder a la información generada sobre usted en el estudio. Además, se le puede facilitar el informe de resultados una vez finalizado. En ambos casos para acceder a esa información póngase en contacto con el investigador principal cuyos datos de contacto se incluyen en este documento.

Finalmente, el informante se compromete igualmente a respetar la confidencialidad de las cuestiones debatidas en el grupo de discusión y a no divulgar la identidad del resto de los participantes.

3 CONSENTIMIENTO

Don/Dña _____
mayor de edad, titular del DNI: _____, por el presente documento manifiesto que he sido informado/a de las características del Proyecto de Investigación titulado: He leído tanto el apartado 1 del presente documento titulado "información básica que debe conocer", como el apartado 2 titulado "Compromiso de confidencialidad", y he podido formular las dudas que me han surgido al respecto. Considero que he entendido dicha información. Estoy informado/a de la posibilidad de retirarme en cualquier momento del estudio.
En virtud de tales condiciones, consiento participar en este estudio.
Y en prueba de conformidad, firmo el presente documento en el lugar y fecha que se indican a continuación.

En _____, _____ de _____ de 2020.

Nombre y apellidos del / de la participante: _____ Nombre y apellidos del investigador: _____

Firma: _____

Firma: _____



Tools and templates

Data lifecycle

Plan to share

Legal and ethical

Rights

Document your data

Format your data

Store your data

Collaborative research

Training

● Tools and templates

Handbook

"Providing researchers with access to the tools they need"



SHARE ↩

This is our selection of tools and templates that researchers may find useful for various data management tasks in social sciences research:

- [Model consent form](#) (doc) that takes into account consent for data sharing and future data reuse
- [Sample survey consent statement](#) (doc) that considers consent for data sharing and future data reuse



EUROPEAN COMMISSION
Directorate-General for Research & Innovation

Horizon 2020 Programme

Guidance

How to complete your ethics self-assessment

Version 6.1
4 February 2019



https://ec.europa.eu/research/participants/data/ef/h2020/grants_manual/hi/ethics/h2020_hi_ethi_cs-self-assess_en.pdf

4.1 Ethics issues checklist

Section 4: PROTECTION OF PERSONAL DATA	YES/NO	Page	Information to be provided	Documents to be provided/kept on file
Does your research involve processing of personal data?	<input type="checkbox"/>	<input type="checkbox"/>	<p>1) Details of the technical and organisational measures to safeguard the rights of the research participants. For instance: For organisations that must appoint a DPO under the GDPR: Involvement of the data protection officer (DPO) and disclosure of the contact details to the research participants. For all other organisations: Details of the data protection policy for the project (i.e. project-specific, not general).</p> <p>2) Details of the informed consent procedures.</p> <p>3) Details of the security measures to prevent unauthorised access to personal data.</p> <p>4) How is all of the processed data relevant and limited to the purposes of the project ('data minimisation' principle)? Explain.</p> <p>5) Details of the anonymisation /pseudonymisation techniques.</p> <p>6) Justification of why research data will not be anonymised/ pseudonymised (if relevant).</p> <p>7) Details of the data transfers (type of data transferred and country to which it is transferred – for both EU and non-EU countries).</p>	1) Informed Consent Forms + Information Sheets used (if relevant).
If YES: - Does it involve the processing of special categories of personal data (e.g. genetic, health, sexual lifestyle,	<input type="checkbox"/>	<input type="checkbox"/>	<p>1) Justification for the processing of special categories of personal data.</p> <p>2) Why can the research objectives not be reached by processing anonymised/ pseudonymised data (if applicable)?</p>	

17

EU Grants: Horizon 2020 Guidance — How to complete your ethics self-assessment: V6.1 – 04.02.2019

ethnicity, political opinion, religious or philosophical conviction,)?				
- Does it involve processing of genetic, biometric or health data?	<input type="checkbox"/>	<input type="checkbox"/>		1) Declaration confirming compliance with the laws of the country where the data was collected.
- Does it involve profiling, systematic monitoring of individuals or processing of large scale of special	<input type="checkbox"/>	<input type="checkbox"/>	<p>1) Details of the methods used for tracking, surveillance or observation of participants.</p> <p>2) Details of the methods used for profiling.</p> <p>3) Risk assessment for the data processing activities.</p>	1) Opinion of the data controller on the need for a data protection impact assessment (art.35 GDPR) (if relevant).



Propuesta de

REGLAMENTO DEL PARLAMENTO EUROPEO Y DEL CONSEJO

sobre el Espacio Europeo de Datos Sanitarios

(Texto pertinente a efectos del EEE)

{SEC(2022) 196 final} - {SWD(2022) 130 final} - {SWD(2022) 131 final} - {SWD(2022) 132 final}

EXPOSICIÓN DE MOTIVOS

CONTEXTO DE LA PROPUESTA

• Razones y objetivos de la propuesta

La Estrategia Europea de Datos [1](#) proponía la creación de espacios comunes europeos de datos en ámbitos específicos. **El Espacio Europeo de Datos Sanitarios (EEDS) es la primera propuesta de estos espacios comunes europeos de datos en ámbitos específicos. El EEDS, que abordará los retos específicos en el ámbito de la salud para el acceso a los datos sanitarios y para su intercambio de forma electrónica, es una de las prioridades de la Comisión Europea en el ámbito de la salud [2](#) y formará parte integrante de la construcción de una Unión Europea de la Salud. El EEDS creará un espacio común en el que las personas físicas podrán controlar fácilmente sus datos sanitarios electrónicos. Los investigadores, los innovadores y los responsables políticos podrán utilizar estos datos sanitarios electrónicos de una manera fiable y segura que respete la privacidad.**

**APROXIMACIÓN A LOS
ESPACIOS DE DATOS
DESDE LA PERSPECTIVA
DEL RGPD**

11 mayo de 2020

Un Espacio de Datos se puede definir como una infraestructura federada y abierta para permitir el acceso soberano de datos, basada en una gobernanza, políticas, reglas y estándares que definen un marco de confianza para todos los intervinientes. Las iniciativas de Espacios de Datos europeas y nacionales plantean modelos de tratamientos de gran complejidad organizativa y tecnológica, así como de una gran escala en el número de sujetos afectados, en la diversidad de categorías de datos tratados, en los estamentos sociales involucrados, en la amplitud geográfica, en los periodos de conservación, en el número de intervinientes y otros. Estas iniciativas no se plantean como una reducción o un compromiso de los derechos y libertades de las personas físicas con relación a la protección de sus datos personales, sino que abren un horizonte de posibilidades que, para garantizar la sostenibilidad con relación al modelo europeo de derechos y libertades, obligan a realizar un análisis objetivo y crítico de su implementación desde el diseño que sea acorde al impacto del tratamiento

Anonimización de datos

Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales. BOE Jueves 6 de diciembre de 2018 Sec. I. Pág. 119788

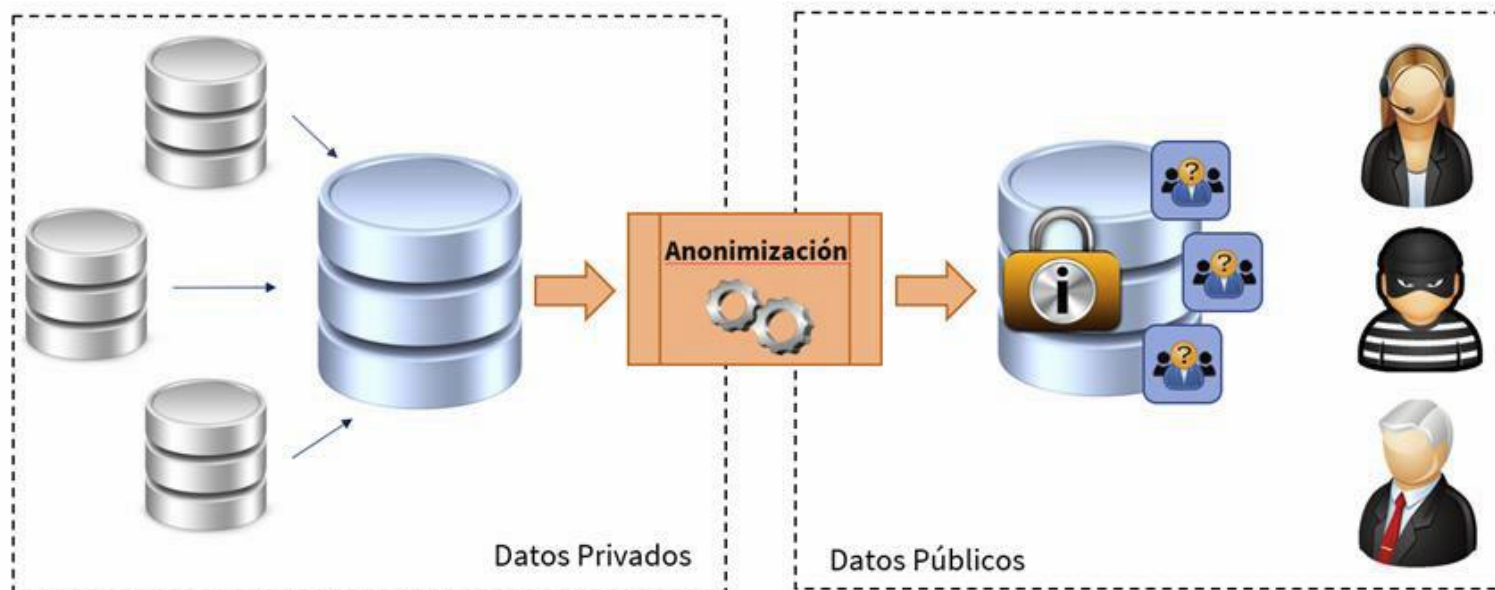
5) **«seudonimización»**: *el tratamiento de datos personales de manera tal que ya no puedan atribuirse a un interesado sin utilizar información adicional, siempre que dicha información adicional figure por separado y esté sujeta a medidas técnicas y organizativas destinadas a garantizar que los datos personales no se atribuyan a una persona física identificada o identificable;*

La AEPD publica recomendaciones para aquellos que realicen procesos de anonimización

- El documento analiza cuáles son los límites en la efectividad de los procesos de anonimización, hasta qué punto la información está realmente anonimizada y cómo se puede cuantificar el riesgo de reidentificaciónLista 1
- Se analiza la K-anonimidad, una técnica que, entre otros aspectos, permite estudiar el grado de identificación que podría existir en un conjunto de datos supuestamente anónimo

(Madrid, 14 de junio de 2019). La Agencia Española de Protección de Datos (AEPD) ha publicado una nota técnica llamada [La K-Anonimidad como medida de la privacidad](https://www.aepd.es/media/notas-tecnicas/nota-tecnica-kanonimidad.pdf), un documento orientado a organizaciones que aborden

<https://www.aepd.es/media/notas-tecnicas/nota-tecnica-kanonimidad.pdf>



- Las fuentes de datos empleadas para dichos tratamientos contienen datos personales que se catalogan como ***“identificadores”*** ya que, por sí solos, están asociados de forma unívoca a un sujeto, como son el DNI, el nombre completo, el pasaporte o el número de la seguridad social.
- El proceso básico de anonimización consiste en disociar de los identificadores el resto de los datos más genéricos asociados a un sujeto como la fecha de nacimiento, el municipio de residencia, el género, etc.
- Al conjunto de datos que no son “identificadores” pero podrían llegar a señalar de forma unívoca a un individuo se le denomina ***“pseudo-identificadores”***, ***“cuasi-identificadores”*** o **identificadores indirectos**.

La ***K-anonimidad*** es una propiedad de los datos anonimizados que permite cuantificar hasta qué punto se preserva la anonimidad de los sujetos presentes en un conjunto de datos en el que se han eliminado los identificadores. Dicho de otro modo, es una medida del riesgo de que agentes externos puedan obtener información de carácter personal a partir de datos anonimizados.

- **Atributos clave o identificadores:** son campos que identifican unívocamente a los sujetos de los datos (nombre, DNI, nº de pasaporte, teléfono, ...). Este tipo de datos deben eliminarse de los registros anonimizados.
- **Cuasi-identificadores:** son campos que, si bien por si mismos y de forma aislada no identifican a un individuo, agrupados con otros atributos *cuasi-identificadores* pueden señalar de forma unívoca a un sujeto. Las técnicas de anonimización trabajan sobre estos datos, eliminando campos que no son necesarios para el tratamiento (en aplicación del principio de minimización), agregándolos o generalizándolos.
- **Atributos sensibles:** son los campos que contienen datos que podrían tener un mayor impacto en la privacidad de un individuo concreto, entre ellos las categorías especiales de datos, y que no deben ser vinculados con el sujeto de datos al que pertenecen (enfermedades, tratamientos médicos, nivel de renta, ...). Esta información puede ser de gran interés en el objeto del tratamiento de datos, pero a menos que exista una legitimación para ello, debe mantenerse dissociada de un sujeto concreto.

Existen dos métodos ampliamente utilizados para implementar la K -anonimización y que no introducen perturbación en los datos:

- **La generalización**
- **La eliminación**

La generalización consiste en hacer que el valor de los atributos cuasi-identificadores sea menos precisos, transformándolos o generalizándolos dentro de un conjunto o intervalo que comparte los mismos valores, bien mediante la creación de rangos en el caso de atributos numéricos o el establecimiento de jerarquías para los atributos nominales. De este modo, el número de registros que poseen los mismos valores para un conjunto de atributos cuasi-identificadores se puede incrementar con el objeto de satisfacer los requisitos de privacidad a la vez que sigue siendo posible cumplir con la finalidad del tratamiento.

“añadir ruido sin perder información”

<https://amnesia.openaire.eu/>

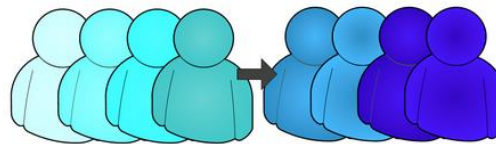


Anonimización de datos

[Home](#) [Get Amnesia!](#) [What is Amnesia?](#) [Documentation](#) [On-line version](#)
[About](#)

Amnesia

Amnesia is a data anonymization tool, that allows to remove identifying information from data. Amnesia not only removes direct identifiers like names, SSNs etc but also transforms secondary identifiers like birth date and zip code so that individuals cannot be identified in the data. Amnesia supports k -anonymity and k^m -anonymity.



AS OPEN AS POSSIBLE, AS CLOSED AS NECESSARY

Grantees have the right to **opt-out**, but need to say **why**

Top three reasons for opt-out:

- privacy
- intellectual property rights
- might jeopardise project's main objective

The approach has been tested during a Horizon 2020 pilot action

Year	Projects	Opted to share data
2015	of 431 signed projects	65.4%
from 2017	the current Open Research Data Pilot	expands to cover all areas of Horizon 2020, with the same rules

Ejemplo anonimización

ID	Age	Zipcode	Diagnosis
1	28	13053	Heart Disease
2	29	13068	Heart Disease
3	21	13068	Viral Infection
4	23	13053	Viral Infection
5	50	14853	Cancer
6	55	14853	Heart Disease
7	47	14850	Viral Infection
8	49	14850	Viral Infection
9	31	13053	Cancer
10	37	13053	Cancer
11	36	13222	Cancer
12	35	13068	Cancer

k-anonymization



ID	Age	Zipcode	Diagnosis
1	[20-30]	130**	Heart Disease
2	[20-30]	130**	Heart Disease
3	[20-30]	130**	Viral Infection
4	[20-30]	130**	Viral Infection
5	[40-60]	148**	Cancer
6	[40-60]	148**	Heart Disease
7	[40-60]	148**	Viral Infection
8	[40-60]	148**	Viral Infection
9	[30-40]	13***	Cancer
10	[30-40]	13***	Cancer
11	[30-40]	13***	Cancer
12	[30-40]	13***	Cancer