

Introducción a la Gestión de datos de investigación

Amparo Miranda

Charo Pérez

Sección de Apoyo a la Investigación

Biblioteca de la Universidad de Sevilla



Contenido

- Ciencia abierta y datos de investigación
- Datos de investigación:
 - Definición/clasificación
 - Principios FAIR
 - Aspectos éticos/Protección de datos
- Gestión de los datos
 - Definición
 - Antes del proyecto: planificación/ Plan de Gestión de Datos
 - Durante el proyecto: organización/procesamiento y análisis
 - Al finan del proyecto: preservación/depósito/compartición
- Depósito de datos en idUS

A principios de siglo surge un movimiento que propugna el **acceso abierto** a la literatura científica financiada con **fondos públicos** en Internet, **libre y gratuita** sin barreras.

¿Por qué surge?

- Investigación financiada con fondos públicos debe estar accesible sin restricciones económicas.
- Mercantilización de la ciencia y precios abusivos por parte de las editoriales.
- Resistencia de los autores a perder sus derechos de copyright.
- Facilidad de las nuevas tecnologías para intercambiar información: la e-Ciencia

Desde 2009 el concepto de acceso abierto empieza a evolucionar a uno más amplio, la Ciencia **Abierta**, que propugna la **apertura de todas las fases de la investigación** sobre la base de una ciencia colaborativa hecha por y para la sociedad (Anglada y Abadal, 2018)



El concepto de ciencia abierta empieza a arraigar en la Unión Europea que irá perfilando y adaptando sus políticas al nuevo modelo.

“Open Science aims at transforming science through ICT tools, networks and media, to make research more open, global, collaborative, creative and closer to society”. (Comisión Europea, 2015)

La Comisión Europea toma como objetivo que la Open Science sea una realidad cuanto antes:

- Infraestructuras digitales: EOSC, ORE
- Financiación: Programa H2020/Horizonte Europa



Creando infraestructuras digitales que faciliten la Ciencia Abierta



Research and Innovation

Open Research Europe

Search

Browse Gateways & Collections How to Publish About Resource Hub Blog

Rapid & Transparent Publishing

Fast publication and open peer review for research stemming from Horizon 2020, Horizon Europe and Euratom funding across all subject areas.

Submit your research

Open Research Europe



Powered by F1000Research



EUROPEAN OPEN SCIENCE CLOUD

La [Nube Europea de Ciencia Abierta \(EOSC\)](#) es un entorno para alojar y procesar datos de investigación para apoyar la ciencia de la UE.

Introduciendo requerimientos de Ciencia Abierta en sus programas de financiación



Ciencia abierta a lo largo del programa

Ciencia abierta

Mejor difusión y explotación de los resultados de investigación e innovación, así como apoyo a la participación activa de la sociedad

Acceso abierto obligatorio para las publicaciones: los beneficiarios se asegurarán de que ellos o los autores conservan los derechos de la propiedad intelectual necesarios para cumplir los requisitos de acceso abierto.

Garantizar el acceso abierto a los datos de investigación: de conformidad con el principio «tan abierto como sea posible y tan cerrado como sea necesario»; plan obligatorio de gestión de datos para datos FAIR (fáciles de encontrar, accesibles, interoperables y reutilizables) y datos de investigación abiertos.

- Apoyo a las habilidades de los investigadores en materia de ciencia abierta, así como sistemas de recompensa.
- Uso de la Nube Europea de la Ciencia Abierta.



Los proyectos financiados en Horizonte 2020 se deben depositar en abierto tanto los resultados de investigación como los datos – datasets- en un repositorio de acceso abierto:

- Antes de 6 meses desde su publicación en el caso de las Ciencias.
- El plazo será de 12 meses para las Ciencias Sociales y Humanidades



Y también en la financiación nacional

- *Plan de difusión e internacionalización, en su caso, de los resultados, incluyendo la **previsión sobre publicaciones y datos de acceso en abierto.***
- *Resumen del plan de gestión de datos previsto durante y al finalizar el proyecto.*

PEICTI

Plan Estatal de
Investigación Científica,
Técnica y de Innovación
2021-2023





Con el propósito de alcanzar el acceso abierto completo e inmediato a aquellas publicaciones que han sido financiadas con fondos públicos, surge en 2018 la iniciativa Plan S. Basado en [10 principios](#), ha sido impulsado por [COALition S](#), un grupo formado por organizaciones que financian la investigación (Wellcome Trust, Bill & Melinda Gates Foundation, y las agencias nacionales de financiación de la investigación de 14 países)



LEGISLACIÓN CONSOLIDADA

Ley 17/2022, de 5 de septiembre, por la que se modifica la Ley 14/2011, de 1 de junio, de la Ciencia, la Tecnología y la Innovación.

Jefatura del Estado
«BOE» núm. 214, de 06 de septiembre de 2022
Referencia: BOE-A-2022-14581

Ley de la Ciencia (2022)

Preámbulo

La ciencia, la tecnología y la innovación como ejes clave en la consecución de los Objetivos de Desarrollo Sostenible de la Agenda 2030.

“el compromiso con la difusión universal del conocimiento, mediante el posicionamiento a favor de las políticas de acceso abierto a la información científica”

Artículo 2. Objetivos generales

c) Impulsar la **ciencia abierta** al servicio de la sociedad y promover iniciativas orientadas a facilitar el **libre acceso a los datos, documentos y resultados generados por la investigación**, desarrollar infraestructuras y plataformas abiertas, y fomentar la participación abierta de la sociedad civil en los procesos científicos.

Ley de la Ciencia (2022)

Artículo 37. Ciencia abierta

1. ... que los resultados de la investigación, incluidas las **publicaciones científicas, datos, códigos y metodologías**, estén disponibles en acceso abierto...

2. El personal de investigación del sector público o cuya actividad investigadora esté financiada mayoritariamente con fondos públicos y que opte por diseminar sus resultados de investigación en publicaciones científicas, deberá depositar una **copia de la versión final aceptada** para publicación **y de los datos asociados** a las mismas en **repositorios institucionales o temáticos** de acceso abierto, de forma **simultánea a la fecha de publicación**.

LOSU (Ley Orgánica del Sistema Universitario)

Preámbulo

Necesitamos una Ciencia Abierta, que asuma ese conocimiento como un bien común, accesible y no mercantilizado, una Ciencia Ciudadana en la que se construya conocimiento de manera compartida, asumiendo la complejidad de la investigación de manera colectiva.

Artículo 12. Fomento de la Ciencia Abierta y Ciencia Ciudadana.

2. Depósito de una copia de la **versión final aceptada y los datos en repositorios**, de forma simultánea a la fecha de publicación.
3. La versión digital de las publicaciones académicas se depositará en los **repositorios institucionales**, sin perjuicio de otros repositorios de carácter temático o generalista.
5. Los datos deberán seguir los **principios FAIR** y, siempre que sea posible, difundirse en **acceso abierto**.
6. **Transparencia** en los **acuerdos** de suscripción con editoriales científicas.
8. Las agencias de calidad incluirán la **accesibilidad en abierto**.
10. Se fomentará la **Ciencia Ciudadana**.

LOSU (Ley Orgánica del Sistema Universitario)

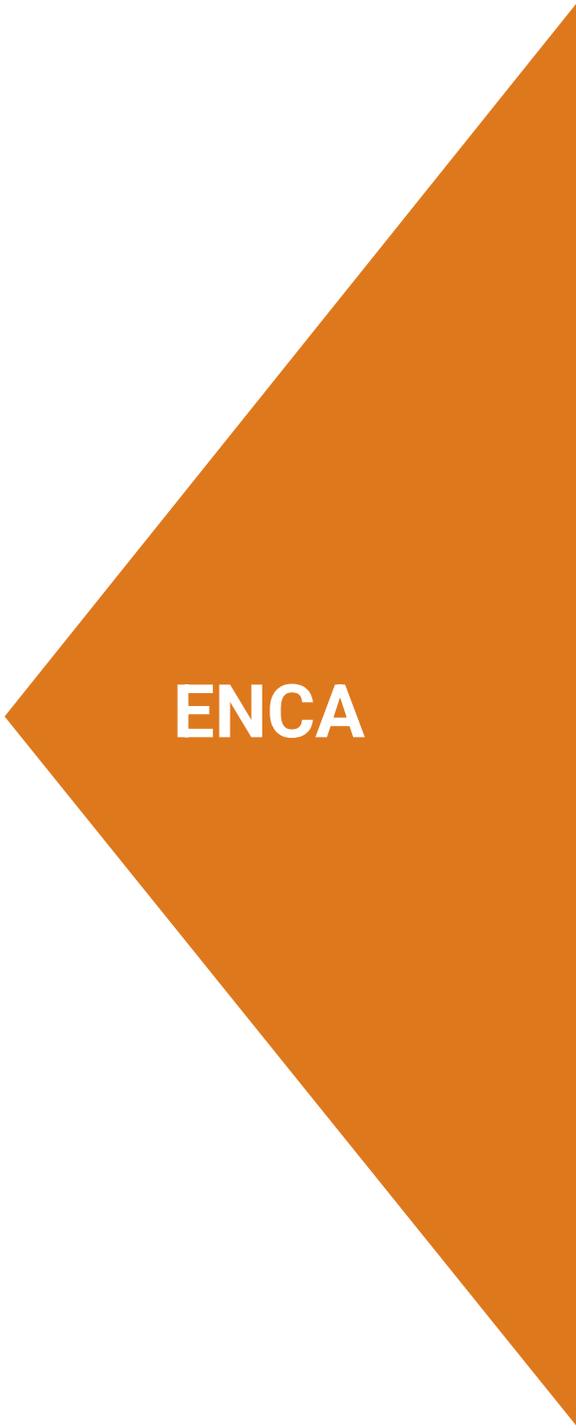
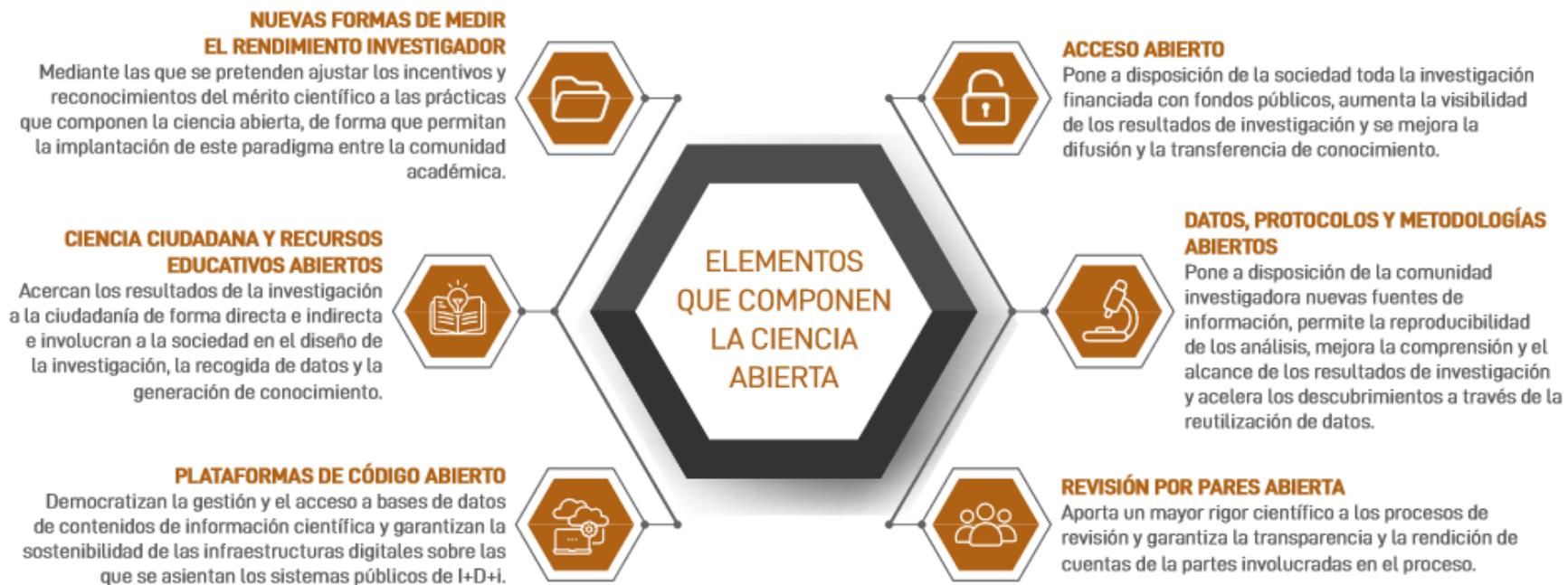
Artículo 56. Programación y sistema de financiación

c) Se incluyen los programas de **Ciencia Abierta y Ciencia Ciudadana**, la **transferencia e intercambio del conocimiento**, la **innovación** en la financiación por objetivos para las Universidades.

Artículo 69. Acreditación de los cuerpos docentes universitarios

- La documentación estará **accesible, en modo abierto**, utilizando **los repositorios institucionales**.
- Evaluación tanto cualitativa como cuantitativa de los méritos, con una **amplia gama de indicadores de relevancia científica e impacto social**.
- Evaluación basada en la especificidad del área o ámbito de conocimiento, teniendo en cuenta, entre otros criterios, [...], la **relevancia local**, el **pluralismo lingüístico** y el **acceso abierto a datos y publicaciones científicas**.

ENCA (Estrategia Nacional de Ciencia Abierta)



ENCA

[Estrategia Nacional de Ciencia Abierta 2023-2027](#)

ENCA (Estrategia Nacional de Ciencia Abierta)

Ejes estratégicos



Eje A. Infraestructuras digitales para la Ciencia abierta.



Eje B. Gestión de datos de investigación siguiendo los principios FAIR.



Eje C. Acceso abierto a las publicaciones científicas.

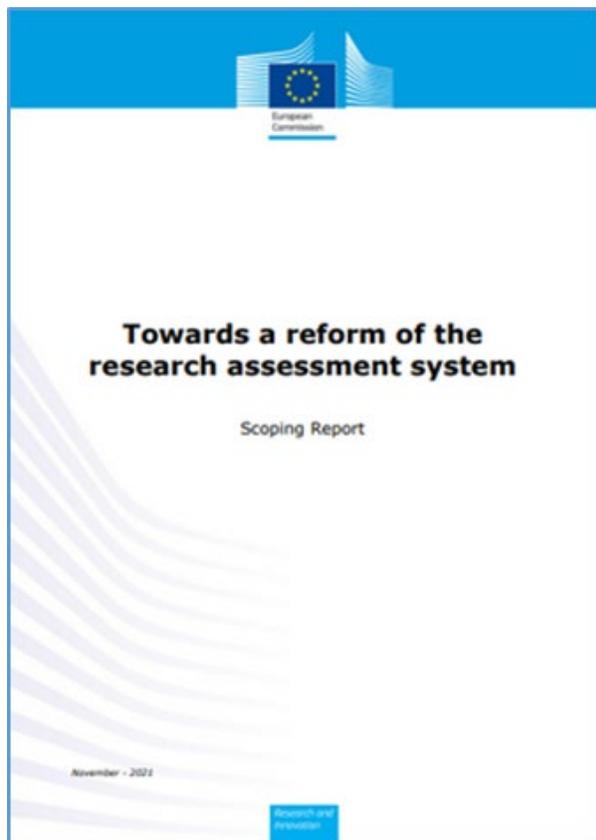


Eje D. Incentivos, reconocimientos y formación.

ENCA

[Estrategia Nacional de Ciencia Abierta 2023-2027](#)

La evaluación de la ciencia está siendo objeto de debate internacional desde hace tiempo y es un motor indispensable en el camino hacia la ciencia abierta



DORA 10

LEIDEN MANIFESTO FOR RESEARCH METRICS

El movimiento de reforma en Europa ha cristalizado en la creación de la [Coalition for Advancing Research Assessment \(CoARA\)](#), promovida por la Comisión Europea, Science Europe y la European University Association (EUA).

Ha elaborado un [acuerdo](#) para impulsar las **nuevas políticas de evaluación**:

- Diversidad de contribuciones.
- Evaluación cualitativa.
- Evitar el uso inapropiado de métricas de revistas.
- Financiar la reforma de la evaluación
- Impulsar la concienciación sobre la necesidad de la reforma.

Reforma de la evaluación en España

ANECA se adhiere al manifiesto de Leiden, a DORA y a COARA



1. Principios

En este primer bloque se establecen los principios básicos que guían el proceso de actualización de los criterios para la evaluación de la investigación. Los mismos se aplican de forma transversal a los programas de evaluación de personas de ANECA (PEP, ACADEMIA y SEXENIOS).

“El conocimiento es un bien común, que no puede ser mercantilizado”

P9. Incorporar elementos de ciencia abierta dentro de la estrategia europea y española.



Principios y directrices para la actualización de criterios de evaluación de la investigación de ANECA 2021

DIRECTRICES DE INVESTIGACIÓN – ESTÁNDARES E INDICADORES PARA LA EVALUACIÓN CUALITATIVA DE LA PRODUCCIÓN CIENTÍFICA

DI13. Se recomienda que tanto las aportaciones presentadas para obtener el sexenio de investigación como las 4 aportaciones relevantes que deben identificarse en ACADEMIA estén depositadas en un repositorio con *Green Open Access* como los repositorios institucionales de los que disponen las universidades. En el futuro será obligatorio, salvo impedimento legal o imposibilidad técnica, este depósito para todas las aportaciones mencionadas.

DIRECTRICES DE INVESTIGACIÓN – MÉRITOS DE ESPECIAL RELEVANCIA (PEP, SEXENIOS Y MÉRITOS ESPECÍFICOS DE CU Y COMPLEMENTARIOS DE ACADEMIA)

DI20. Se podrá valorar, dependiendo del campo, la apertura de datos vinculados a una publicación dentro de la revista.

Reforma de la evaluación en España

GOBIERNO DE ESPAÑA
MINISTERIO DE LA PRESIDENCIA, RELACIONES CON LAS CORTES Y MEMORIA DEMOCRÁTICA

Agencia Estatal Boletín Oficial del Estado

Castellano Buscar MI BOE Menú

Está Ud. en > Inicio > BOE > Calendario > 06/09/2023 > Documento BOE-A-2023-19027 < anterior siguiente >

Real Decreto 678/2023, de 18 de julio, por el que se regula la acreditación estatal para el acceso a los cuerpos docentes universitarios y el régimen de los concursos de acceso a plazas de dichos cuerpos.

[Ver texto consolidado](#)

Publicado en: «BOE» núm. 213, de 6 de septiembre de 2023, páginas 122924 a 122953 (30 págs.)
 Sección: I. Disposiciones generales
 Departamento: Ministerio de Universidades
 Referencia: BOE-A-2023-19027
 Permalink ELI: <https://www.boe.es/eli/es/rd/2023/07/18/678>

Otros formatos:

ENTREVISTA — Directora de la ANECA

Pilar Paneque: “La exigencia de publicar constantemente lleva a un sistema científico de cantidad, no de calidad”

La responsable de la acreditación y evaluación del profesorado universitario promete darle la vuelta a lo que tienen que hacer los científicos para hacer carrera y lograr así un cambio en el sistema: “El 'paper' no puede ser lo único que haga el investigador”

- Simplificación del procedimiento
- Uso repositorios institucionales
- Pluralidad de contribuciones
- Estancias 9 meses en instituciones distintas a la de obtención de título de doctor. (PTU)
- Carácter universal
- CV breve
- Sin certificación documental de todos los méritos....



Reforma de la evaluación en España

📅 20/10/2023

ANECA actualiza los principios y los criterios de evaluación de los sexenios de investigación

La dirección de ANECA trabaja en la revisión de los principios y los criterios de evaluación de los sexenios de investigación, que se publicarán por Resolución de la Comisión Nacional Evaluadora de la Actividad Investigadora (CNEAI) a finales de 2023 en el Boletín Oficial del Estado (BOE)



ANECA dará a conocer el borrador de la Resolución a principios de noviembre para recibir sugerencias y aportaciones para su mejora, antes de su revisión definitiva y publicación en el BOE.

Ajustar la convocatoria a la LOSU y al nuevo R.D. Acreditación:

- Mayor diversidad aportaciones
- Relevancia de la aportación no del medio de difusión.
- Promoción del acceso abierto y uso de los repositorios institucionales y temáticos.
- Valorar positivamente investigaciones multidisciplinares e interdisciplinares.

Borrador de criterios de evaluación de la Convocatoria 2023 de Sexenios de investigación

Así, en el caso de las publicaciones seriadas o periódicas, se recuerda a las personas solicitantes que, desde la entrada en vigor de la LCTI, el 2 de diciembre de 2011, el personal de investigación **cuya actividad investigadora esté financiada mayoritariamente con fondos públicos hará pública la versión final aceptada para publicación en repositorios institucionales o temáticos de acceso abierto** (artículo 37 de la LCTI, en su versión previa a la Ley 17/2022). Este precepto, reforzado ahora en los artículos 12.2 y 12.3 de la LOSU, será por tanto de aplicación para las publicaciones presentadas con fecha 2011 o posterior, excepto en el caso de otros formatos de publicación como las monografías.

Además, **el vigente artículo 37** de la LCTI, tras su modificación por la Ley 17/2022), **añade que se deberá depositar una copia de la versión final aceptada para publicación y los datos asociados a las mismas en repositorios institucionales o temáticos de acceso abierto, de forma simultánea a la fecha de publicación, siendo este precepto de aplicación para las publicaciones presentadas con fecha 2022 o posterior.**

Asimismo, de acuerdo con el artículo 5 37 de la LCTI y el artículo 12.5 de la LOSU, **los conjuntos de datos que se sometan a evaluación, con fecha 2022 o posterior, deberán cumplir con los principios FAIR (fáciles de encontrar, accesibles, interoperables y reutilizables) y, siempre que sea posible, se difundirán en acceso abierto en repositorios o infraestructuras de datos de confianza.**

Mandatos y políticas: Horizonte Europa

Horizonte Europa: programa marco 2021-2027

- Principios FAIR
- Elaboración de un Plan de Gestión de Datos
- Depósito de los datos en un repositorio de confianza
- Garantizar el acceso abierto a través del repositorio
- «Tan abierto como sea posible según sea necesario»
- Licencia Creative Commons CC BY o CC 0
- Proporcionar cualquier herramienta o instrumento necesario para reutilizar o validar los datos.

[Guía del participante](#) pag. 82



Mandatos y políticas: Plan Nacional

➤ [Proyectos de Generación del conocimiento. Preguntas frecuentes](#) (pag.44)

88.7. Datos de Investigación - ¿Debo incluir en la solicitud de mi proyecto el Plan de Gestión de Datos (PGD)?

No.

En la memoria científico-técnica de la propuesta solamente se deberá incorporar una descripción inicial que contenga qué datos se van a recoger o generar en el marco del proyecto (tipologías y formatos), cómo será el acceso a los mismos (quién, cómo y cuándo se podrá acceder a ellos), de quién son los datos y en qué repositorio está previsto su depósito, difusión y preservación. Asimismo, se recogerán, en su caso, las condiciones éticas o legales específicas que los regulen (ej. privacidad de los datos y su reglamentación; datos protegidos o protegibles por propiedad intelectual o industrial, etc.) que condicionen su disponibilidad, uso y/o reutilización. Esta declaración inicial en el momento de la solicitud no condicionará la creación de un Plan de Gestión de Datos formal en el caso de que el proyecto resulte finalmente financiado.

88.12. Datos de Investigación - ¿Qué repositorios se podrán utilizar para depositar los datos?

Se deberán depositar los datos de investigación en repositorios de confianza (TDR, Trusted Digital Repositories), institucionales, multidisciplinares (ej. Zenodo); o bien en infraestructuras colaborativas de datos, en disciplinas que utilizan datos de forma masiva (DID, Data Intensive Domain), bien dentro del dominio o con carácter general o multidisciplinar (ej. EUDAT).

88.11. Datos de Investigación - ¿Qué datos de investigación se deben depositar?

Se han de **depositar** en repositorios de acceso abierto todos los datos que subyacen a la investigación, esto es, los datos brutos generados o producidos en el transcurso de la investigación. Así mismo, se deberán **publicar** junto a los artículos científicos los datos finales que sean necesarios para garantizar la verificación y reproducibilidad de los resultados presentados. Según las disciplinas y los flujos de producción, análisis y



Políticas de la revistas científicas

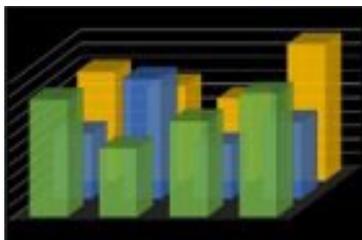
La mayoría de los editores científicos exigen que los datos en los que se basan los resultados de investigación estén disponibles en abierto en repositorios de datos con el objeto de facilitar el proceso de revisión por pares, facilitar la reproducibilidad, aumentar la confianza en la ciencia, etc.

Políticas de acceso a los datos de investigación:

- [Elsevier](#)
- [Springer Nature](#)
- [Wiley](#)
- [Sage](#)
- [Taylor and Francis](#)
- [Oxford University Press](#)
- [Cambridge University Press](#)

Datos de investigación

Se consideran datos de investigación a todo aquel material que ha sido generado, recolectado, observado o registrado durante el ciclo de vida de un proyecto de investigación y que se utilizan como evidencia de un proceso de investigación, están reconocidos por la comunidad científica y sirven para validar los resultados de la investigación y garantizar su reproducibilidad.



NO se consideran datos: notas y ejemplares de laboratorio, sets de datos parciales, análisis preliminares, borradores de trabajos, planes para investigaciones futuras, informes que han tenido un proceso de revisión por pares, comunicaciones con colegas.

Ejemplos de los datos de investigación



Subject

Social Sciences (185)
 Medicine, Health and Life Sciences (173)
 Arts and Humanities (166)
 Other (120)
 Earth and Environmental Sciences (94)
 Engineering (78)
 Computer and Information Science (71)
 Agricultural Sciences (46)
 Chemistry (38)
 Physics (34)
 Business and Management (12)
 Mathematical Sciences (10)
 Law (2)
 Astronomy and Astrophysics (1)

[Less...](#)

<https://dataverse.csuc.cat/>



Materia

Ciencias sociales (433)
Artes y humanidades (425)
Ciencias de la tierra y el medioambiente (80)
Ingeniería (79)
Física (35)
Derecho (24)
Química (20)
Ciencias médicas, de la salud y de la vida (20)
Otra (19)
Ciencias de la información y computación (17)
Ciencias matemáticas (10)
Ciencias agrarias (8)
Astronomía y astrofísica (8)
Administración y empresas (3)

[Menos...](#)

<https://edatos.consorciomadrono.es/>

Clasificación de los datos de investigación

- **Según su naturaleza:** cualitativos o cuantitativos.
- **Según su nivel de procesamiento:** en estado bruto (datos primarios), procesados o analizados.
- **Según su apertura:**
 - **Abiertos:** aquellos a los que cualquier persona puede acceder, usar y compartir.
 - **Restringidos:** datos confidenciales, datos personales, o datos comerciales confidenciales.

Clasificación de los datos de investigación

- Según su formato:

- Textuales (Word, PDF, RTF, etc.)
- Numéricos (Excel, CSV, etc.)
- Multimedia (JPEG, MPEG, WAV, etc.)
- Estructurados (XML, MySQL, etc.)
- Código de software (Java, C, etc.)
- Específicos de un software (Mesh, 3D CAD, modelo estadístico, etc.),
- Específicos de una disciplina o instrumento



Clasificación de los datos de investigación

- **Según la fuente:**
 - **Observacionales:** datos capturados en tiempo real. Por ejemplo, neuroimágenes, datos de muestras, datos de sensores, datos de encuesta.
 - **Experimentales:** datos capturados en equipos de laboratorio. Por ejemplo, secuencias de genes, cromatogramas, datos de campos magnéticos.
 - **Simulación:** datos generados a partir de modelos de prueba. Por ejemplo, climatológicos, matemáticos o modelos económicos.
 - **Derivados o compilados:** datos reproducibles, pero de difícil reproducción. Por ejemplo, de texto y minería de datos, modelos 3D, bases de datos compilados.
 - **De referencia:** conglomerado o conjunto de datos. Por ejemplo, bancos de datos de secuencias de genes, estructuras químicas o portales de datos espaciales.

Importancia de compartir los datos

Del 1-1-2019 al 31-12-2022

- ✓ Retractados 1400 artículos (175 de Elsevier, 140 Taylor and Francis, 104 Wiley....)

<http://retractiondatabase.org>



Causas:

- Datos no fiables
- Datos inventados
- Datos manipulados
- No reproducibles
- Dudas del editor
- Errores en los resultados

Fuente: Gestión de datos de investigación y Planes de Gestión de Datos (PGD) para documentalistas /remedios Melero <http://hdl.handle.net/20.500.12105/16205>

Por qué compartir los datos

- ✓ Fomenta el debate y las colaboraciones con otros investigadores.
- ✓ Maximiza la transparencia de las investigaciones.
- ✓ Permite validar y replicar los resultados obtenidos.
- ✓ Incremento del impacto y visibilidad de la investigación.
- ✓ Reduce la duplicidad de recogida de datos y los costes que conlleva.
- ✓ Fomento de la innovación con otros usos potenciales.
- ✓ Requisito de los organismos financiadores.
- ✓ Exigencia de las revistas científicas.

¿Dónde encontrar datos?

Repositorios de conjuntos de datos (datasets) generalistas y especializados

- [dataMED](#): buscador de datasets y repositorios en el ámbito biomédico.
- [DataONE](#): buscador de datasets en repositorios sobre la vida en la tierra y medio ambiente.
- [DataCite Search](#): buscador generalista de datasets abiertos con DOI.
- [Dryad](#): buscador de datasets asociados a publicaciones de ciencias y medicina.
- [e-cienciaDatos](#): buscador de datasets del Consorcio Madroño (Universidades de Madrid y UNED)
- [Eudat B2FIND](#): buscador generalista de datasets mantenido por la organización EUDAT Collaborative Data Infrastructure.
- [Figshare](#): buscador internacional en el que se pueden recuperar datos por categorías temáticas.
- [Google Dataset Search](#): buscador de datasets en repositorios de datos científicos, bases de datos de gobiernos locales y nacionales, sitios web de editores y autores y otras fuentes.
- [Eurostat](#): Estadísticas y datos de alta calidad sobre Europa.
- [Zenodo](#): buscador de datos de investigación de la Unión Europea. Desarrollado por CERN Data Center e Invenio.

¿Dónde encontrar datos?

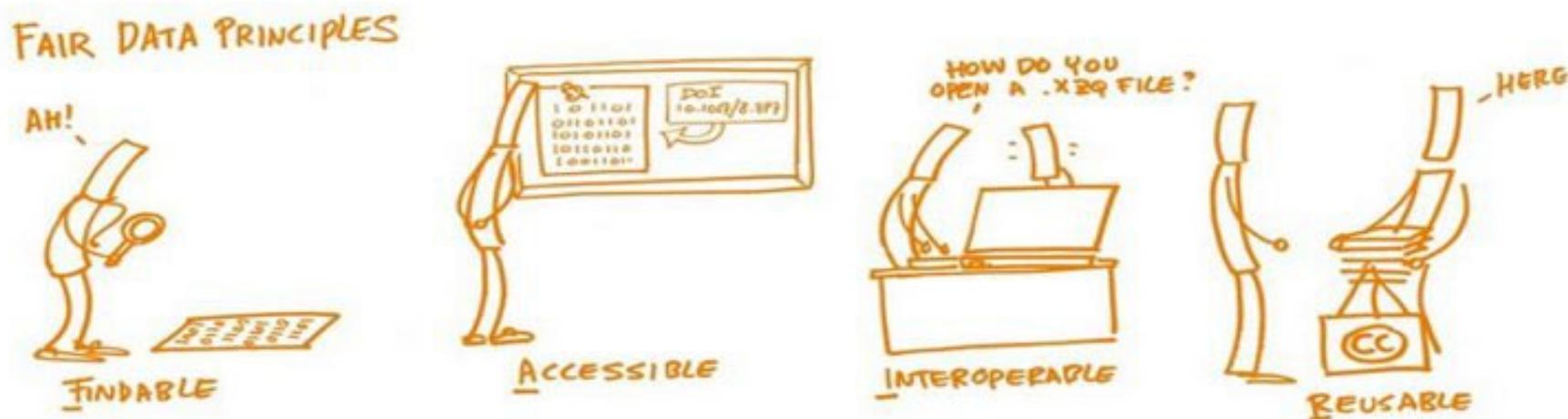
Directorios o buscadores de repositorios

- [Re3data](#)
- [OAD Data Repositories](#)
- [DataMED](#)
- [Repository Finde](#)
- [Recolecta](#)

Principios FAIR

Son un conjunto de principios rectores que fueron publicados en la revista Scientific Data de Nature (2016) en el artículo: [“Principios FAIR para el manejo y administración de datos científicos”](#) e incorporados inicialmente al programa Horizonte 2020 y después al Programa Horizonte Europa para hacer que los datos de investigación fueran **fáciles de encontrar**, **accesibles**, **interoperables** y **reutilizables**.

<https://force11.org/info/the-fair-data-principles/>



Fuente: <https://book.fosteropenscience.eu/es/02OpenScienceBasics/02OpenResearchDataAndMaterials.html>

Principios FAIR



Findables -Localizables

- Los datos han de estar indexados en algún sistema que disponga de motor de búsqueda (repositorio, agregador, etc.).
- Requieren de estándares y metadatos enriquecidos: permite que los datos puedan ser comprendidos, idealmente conforme a los estándares específicos de su disciplina científica.
- Es imprescindible el uso de identificadores únicos y persistentes, que pueden darse a nivel de objeto (DOI, handle, etc.), Los identificadores persistentes permiten localizar y citar los datos de forma unívoca.



Cómo se cumple:

- Identificador persistente: DOI, Handle...
- Descritos con metadatos enriquecidos.
- Los metadatos están en línea en un recurso de búsqueda, están recolectados por motores de búsqueda.
- El registro de metadatos especifica el identificador persistente.

Accesible - Accesibles

- Recuperables mediante su identificador: protocolo de comunicación abierto y estandarizado. Este protocolo ha de ser abierto, libre y de implementación universal.
- Los metadatos han de ser accesibles, incluso si los datos dejan de estar disponibles, a través de un archivo que vele también por la conservación de los metadatos sobre los datos.



Cómo se cumple:

- El ID persistente enlaza con los datos o metadatos asociados para que estos puedan ser recuperados.
- El protocolo para recuperar los datos sigue estándares reconocidos, por ejemplo, <http://>.
- Acceso con autenticación y autorización en caso de que sea necesario.
- Los metadatos serán accesibles incluso cuando los datos no estén ya disponibles.

Interoperable- Interoperables

- Los datos y metadatos deben usar un lenguaje formal, accesible, compartido y de amplia aplicación para la representación del conocimiento.
- Los datos deben poder usarse y combinarse con otros datos o herramientas.
- Se deben usar vocabularios que sigan también los principios FAIR, y establecer referencias cruzadas y enlaces a otros datos o metadatos.

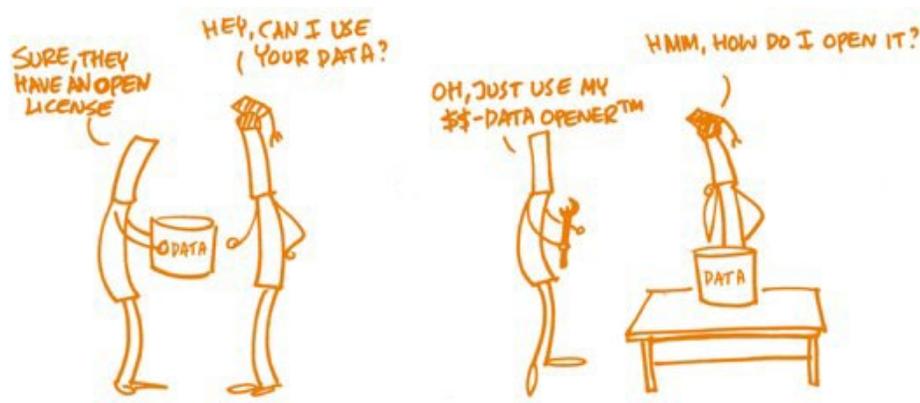


Cómo se cumple:

- Formatos abiertos
- Metadatos siguen estándares reconocidos.
- Vocabularios controlados, palabras clave, tesauros u ontologías.
- Referencias y enlaces a otros datos relacionados.

Reusable- Reutilizables

- Para ello los metadatos han de complementarse con atributos precisos y relevantes.
- Es necesario indicar licencias de uso de datos claras y accesibles.
- Los datos y sus metadatos han de estar asociados con su procedencia y han de cumplir con los estándares relevantes para el dominio.



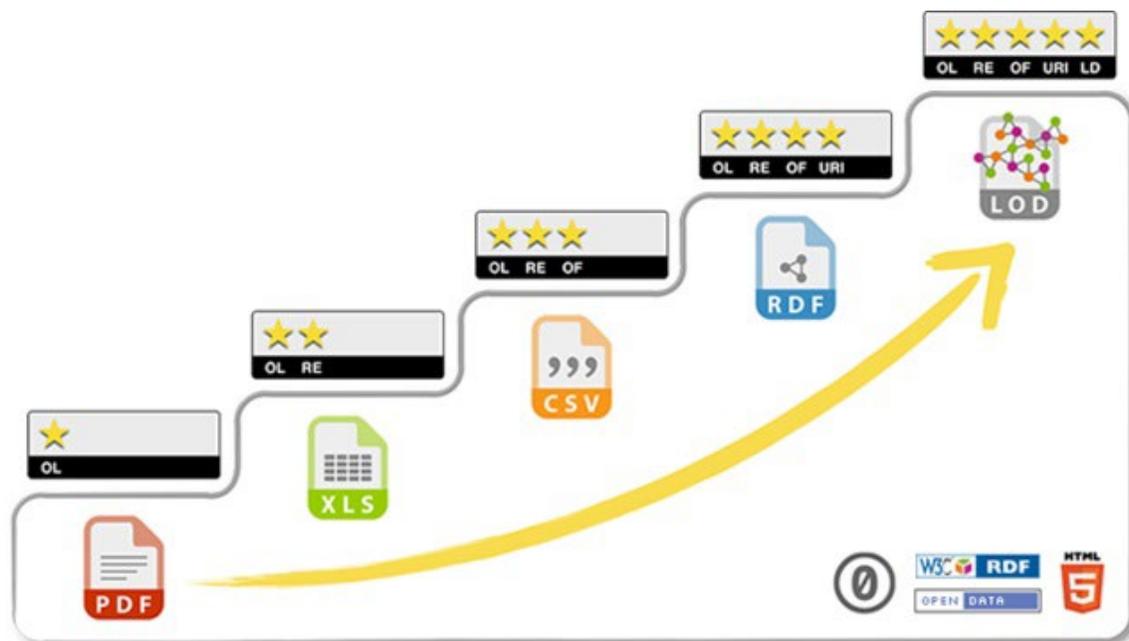
Cómo se cumple:

- Los datos son precisos y están bien descritos con todos los atributos relevantes.
- Los datos tienen una licencia de uso de datos clara y accesible.
- Queda claro cómo, por qué y quién ha creado y procesado los datos.
- Los datos y sus metadatos cumplen con los estándares que usa la comunidad de esa disciplina.

Herramientas de verificación

- Checklist de verificación para comprobar el grado FAIRness de los datos de investigación
Jones, Sarah & Grootveld, Marjan (2021). ¿Cómo comprobar si sus datos son FAIR?. Traducido por Alicia Fátima Gómez
<https://doi.org/10.5281/zenodo.5508099>
- Australian Research Data commons
<https://ardc.edu.au/resources/aboutdata/fair-data/fair-self-assessment-tool/>
- FAIRsFAIR
<https://fairaware.dans.knaw.nl/>

5 estrellas de datos abiertos



- ★ publica tus datos en la Web (con cualquier formato) y bajo una licencia abierta¹
- ★★ publícalos como datos estructurados (ej: Excel en vez de una imagen de una tabla escaneada)²
- ★★★ usa formatos no propietarios (ej: CSV en vez de Excel)³
- ★★★★ usa URIs para denotar cosas, así la gente puede apuntar a estas⁴
- ★★★★★ enlaza tus datos a otros datos para proveer contexto⁵

Ejemplo datos FAIR

- Agregador nacional de repositorios científicos de acceso abierto. <https://recolecta.fecyt.es/>

científicos de acceso abierto

The screenshot shows the RECOLECTA website interface. At the top, there are logos for U23 FECYT and RECOLECTA. Below the logos is a navigation menu with items: ¿QUÉ ES?, SERVICIOS, ACCESO ABIERTO, ALIANZAS FECYT, and DOCUMENTOS. A central dashboard displays five statistics: 182 Repositorios, 115 Repositorios institucionales, 3.248 K Resultados de investigación, 1.609 K Artículos, and 32.474 Set de datos. Below this is a search bar containing 'trastorno mental' and a search button. Underneath the search bar is a 'Buscador avanzado' section with a checked 'ACCESO ABIERTO' filter. There are three input fields for 'AUTOR / NÚMERO ORCID', 'REPOSITORIO', and 'IDENTIFICADOR (DOI / HANDLE)'. Below these are 'INCLUDE' checkboxes for 'PUBLICACIONES', 'DATOS DE INVESTIGACIÓN' (checked), and 'PROYECTOS'. At the bottom, there are three filter buttons: 'Repositorio', 'Autor', and 'Idioma', each with a plus sign. To the right of these filters, there is a blue button labeled 'Datos de investigación' and the text 'Dataset. 2023'. Below this, the title of the dataset is displayed: 'RAS, CORE-OM (PROBLEMS AND RISK SUBSCALES) AND WHOQOLBREF QUESTIONNAIRES APPLIED IN A SAMPLE OF PEOPLE WITH SERIOUS MENTAL ILLNESS [DATASET], CUESTIONARIOS RAS, CORE-OM (SUBESCALAS PROBLEMAS Y RIESGO) Y WHOQOLBREF ADMINISTRADOS EN UNA MUESTRA DE PERSONAS CON TRASTORNO MENTAL GRAVE [DATASET]'.

Cuestiones a tener en cuenta

- Aspectos legales. Propiedad intelectual/Licencias
- Aspectos éticos. Buenas prácticas en la investigación/Comités éticos
- Protección de datos personales. Reglamento General de Protección de datos (GDPR)



Aspectos legales

Derechos de autor y propiedad intelectual:

- Afecta a todo tipo de datos tanto si reutilizan de otra persona como si se archivan los propios.
- Si reutilizamos datos, ¿quién es propietario? ¿Permite reutilizarlos? ¿Permite archivarlos y depositarlos en un repositorio?
- Si los datos proceden de archivos privados o de empresas comerciales, ¿se pueden reutilizar? ¿en qué condiciones? ¿se pueden depositar?
- Datos creados por varios investigadores ¿quién tiene los derechos de autor?
- Proyectos en colaboración con otros países ¿Qué leyes hay que tener en cuenta?
- [Ley 2/2019, de 1 de marzo](#)

Aspectos éticos

Toda investigación debe cumplir con los principios éticos y la legislación vigente a nivel nacional, europeo.

- [Código de buenas prácticas en Investigación](#) de la US.

Si el proyecto incluye datos personales, biomédicos o experimentación animal debe ser evaluado y aprobado por:

- [Comités Éticos](#) de la Universidad de Sevilla.
- [Comité de Ética de Investigación de la Universidad de Sevilla](#)
- [Comité Ético de Experimentación animal](#)

Aspectos éticos

- ✓ **Proyectos en los que participan personas** es obligatorio que firmen un consentimiento informado
- ✓ Normalmente este consentimiento informado se otorga **antes** del inicio de la investigación.
- ✓ **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.

Qué debe hacer el investigador:

- Informar a los participantes sobre el propósito de la investigación;
- Discutir qué pasará con su contribución (incluido el futuro archivado e intercambio de sus datos);
- Indicar las medidas que se tomarán para salvaguardar su anonimato y confidencialidad;
- Describir el derecho a retirarse de la investigación.
- Informar a los participantes sobre los beneficios potenciales del intercambio de datos.

Aspectos éticos

FICHAS INFORMATIVAS

Ayudan a proporcionar a los participantes la información básica necesaria para tomar una decisión informada sobre si participar o no en el proyecto de investigación.

Contenido:

- El propósito de la investigación;
- Qué implica participar en la investigación;
- Los beneficios y riesgos de participar en la investigación;
- Detalles de la investigación, por ejemplo, fuente de financiación, institución patrocinadora, nombre del proyecto, datos de contacto de los investigadores y cómo presentar una queja;
- Los procedimientos para retirarse del proyecto de investigación;
- El uso planificado de los datos durante la investigación, difusión, almacenamiento, publicación y archivo de los datos;
- Las estrategias para asegurar el uso ético de los datos;
- Los procedimientos para salvaguardar la información personal, mantener la confidencialidad y anonimizar los datos, particularmente en relación con el archivo, el intercambio y la reutilización de datos.

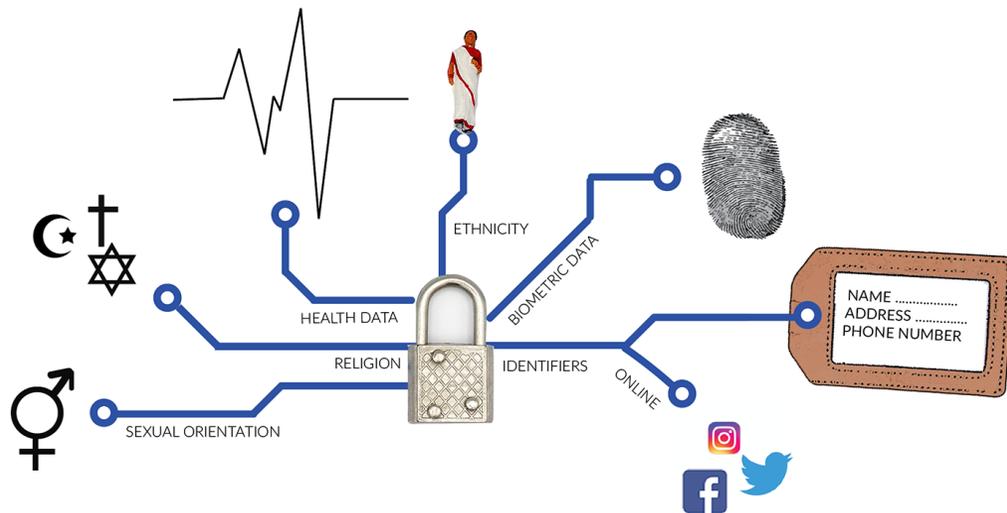
MODELOS DE CONSENTIMIENTOS INFORMADOS

- [Archivo de datos del Reino Unido](#)
- [Unidad de Ciencias del Cerebro y Cognición del MRC - Universidad de Cambridge](#)
- [FORS \(Centro Suizo de Expertos en Ciencias Sociales\)](#)

Protección de datos

- [Reglamento General de Protección de Datos](#) (GDPR; Unión Europea, 2016a) se aplica a cualquier investigador de la UE o del Espacio Económico Europeo (EEE) que recopile datos personales sobre un ciudadano de cualquier país, en cualquier parte del mundo, así como cualquier investigador a nivel mundial que recopile datos personales sobre ciudadanos de la UE.
 - Los datos personales son cualquier información relativa a **una persona física viva identificada o identificable**.
 - Las **distintas informaciones**, que recopiladas pueden llevar a la identificación de una determinada persona, también constituyen datos de carácter personal.

Protección de datos



Datos afectados por el Reglamento Protección de Datos (RGPD) (UE) 2016/679:

Datos personales relacionados con una persona física que se puede identificar por atributos como un nombre, un número de identificación, edad, ubicación, etc.

Datos personales sensibles referidos a información relacionada con el origen racial o étnico de una persona, opiniones políticas, religión, datos de salud, vida sexual, genéticos, biométricos o delitos penales.

En cuanto a la recopilación de datos personales, el RGPD 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)
- 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).

Fuente: <https://dmeg.CESSDA.eu/Data-Management-Expert-Guide/1.-Plan/Data-in-the-social-sciences>

Protección de datos

Legislación

Reglamento Protección de Datos (RGPD) (UE) 2016/679, Reglamento europeo relativo a la protección de las personas físicas en lo que respecta al tratamiento de sus datos personales y a la libre circulación de estos datos
<https://www.boe.es/doue/2016/119/L00001-00088.pdf>

Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales (LOPD-GDD) <https://www.boe.es/buscar/doc.php?id=BOE-A-2018-16673>

En la Universidad de Sevilla:

- [Tratamiento de datos personales en actividades de investigación](#)
- [Protección de datos personales Universidad de Sevilla](#)
- [Procedimiento de Creación de un Tratamiento de DP](#)



Las dudas o consultas Delegada de Protección de Datos de la US: dpd@us.es

Protección de datos

- **Datos seudonimizados** que siguen siendo datos personales pero que no permiten identificar a un sujeto, ya que la información adicional que permitiría dicha identificación se mantiene por separado (ejemplo de seudonimizaciones la sustitución de un nombre por un código o por un seudónimo).
- Los datos personales que hayan sido anonimizados, cifrados o presentados con un seudónimo, pero que **puedan utilizarse para volver a identificar a una persona**, siguen siendo datos personales y se inscriben en el ámbito de aplicación del RGPD.

Protección de datos

- **Datos anonimizados** referidos a información que no puede asociarse a una persona identificable o datos personales anónimos de tal manera que el interesado no es o ya no es identificable. La anonimización pueden ser por generalización, por aleatorización y por eliminación
- Los datos personales que hayan sido anonimizados, de forma que la persona **no sea identificable o deje de serlo**, dejarán de considerarse datos personales. Para que los datos se consideren verdaderamente anónimos, la anonimización debe ser irreversible.

Protección de datos

- Herramientas para la anonimización de datos personales:



- Amnesia de OpenAire



- ARX Data Anonymization Tool

<https://www.aepd.es/documento/guia-orientaciones-procedimientos-anonimizacion.pdf>

Gestión de datos de investigación

Por gestión de datos de investigación (**RDM**, Research Data Management) entendemos el conjunto de tareas que permiten la recopilación, organización, documentación, almacenamiento y preservación de los datos utilizados o generados durante un proyecto de investigación. Gestionar los datos de investigación es importante para asegurar la reproducibilidad de la investigación, así como para demostrar la veracidad de la información y la metodología de la misma.

Por qué gestionar los datos de investigación

1. Cumplir con los requerimientos de los organismos financiadores.
2. Mayor transparencia para la validación de los resultados de la investigación.
3. Mejorar la protección de datos y minimizar el riesgo de pérdida de los mismos.
4. Asegurar que los datos sean FAIR (**F**indable, **A**ccessible, **I**nteroperable and **R**eusable).
5. Ahorrar tiempo evitando duplicaciones y haciendo un uso eficiente de los recursos disponibles.
6. Mejorar el perfil del investigador, el impacto y la visibilidad de los proyectos.

Fuente: Pérez Allende, M.L. (2015) Portal PAGODA. Plan de Gestión de Datos. Universidad Carlos III de Madrid

Ciclo de vida de los datos de investigación

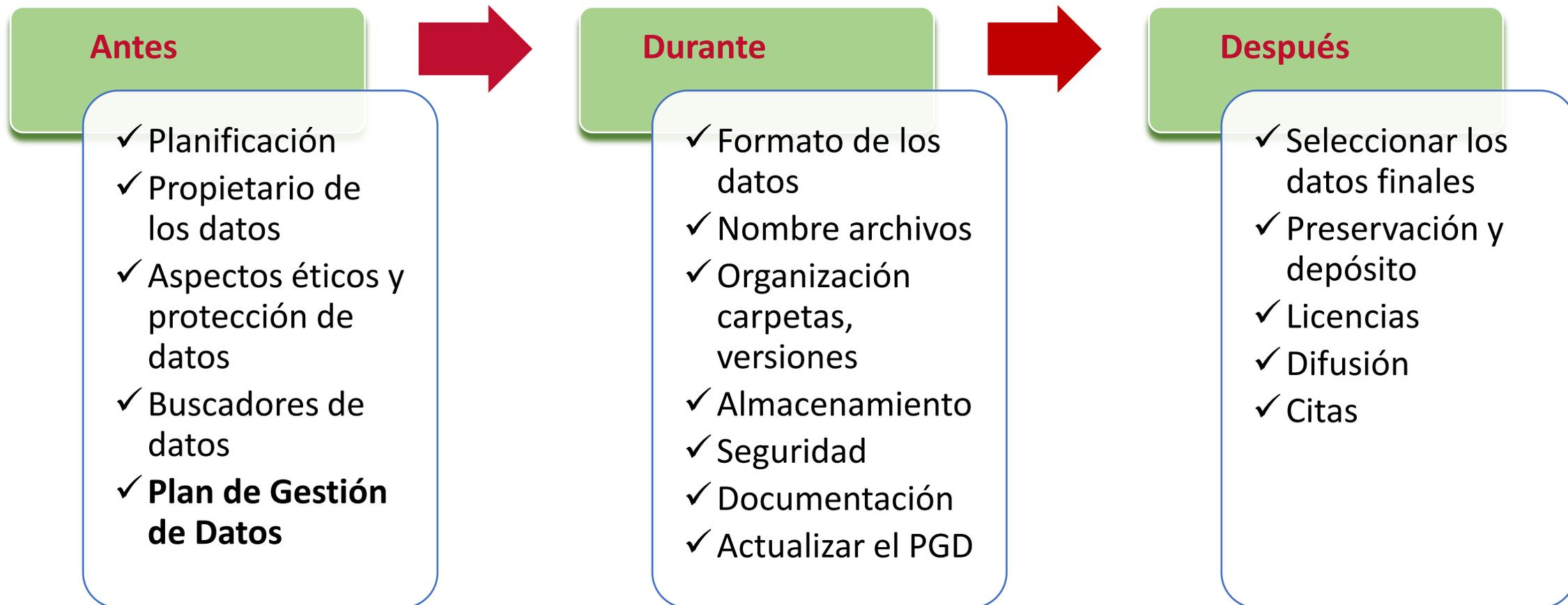


Ciclo de vida de los datos de investigación



https://rdmkit.elixir-europe.org/data_life_cycle

Ciclo de vida de los datos y las fases del proyecto



Antes del proyecto: planificación

- Organización del proyecto
- ¿Qué instituciones colaboran?
- ¿Quiénes forman parte del proyecto?
- Aspectos éticos y protección de datos
- Propiedad de los datos
- ¿Qué tipo de datos, quién los recopilará y procesará?
- ¿Qué software se utilizará?
- ¿Qué formatos tendrán los datos?
- ¿Dónde se almacenarán durante la investigación?
- ¿Quién accederá a esos datos?
- ¿Qué datos se pondrán al final en acceso abierto?
- ¿Dónde se depositarán?



PLAN DE GESTIÓN DE DATOS

Plan de Gestión de Datos (PGD): definición

Un Plan de Gestión de Datos (DMP Data Management Plans) es un documento que describe como se gestionarán los datos generados y procesados en un proyecto de investigación, tanto durante el proyecto como una vez que éste ha finalizado y al finalizar el proyecto de investigación.

Es obligatorio para los proyectos financiados con fondos públicos.

Plan de Gestión de Datos (PGD): definición

“El PGD define la metodología de extracción de la información, la unidad de medida, los procesos de creación, limpieza y análisis, el formato, el software de lectura y procesamiento, las licencias y permisos de acceso y (re)utilización, su sistema de control de versiones, su autoría, su difusión, y su lugar de almacenamiento y preservación a largo plazo. Este plan se creará al comienzo del proyecto de investigación financiado y, según la naturaleza y evolución del proyecto, puede que no sea un documento definitivo, sino que variará, se completará y/o modificará a lo largo del ciclo de vida de la investigación”.

[Preguntas frecuentes Proyectos de generación del conocimiento. Convocatoria 2022](#)

Antes del proyecto: Plan de Gestión de Datos (PGD)

Ventajas de hacer un Plan de Gestión de Datos



1. [Herramienta útil para pensar en el futuro](#)
2. [Permite una fácil gestión de proyectos](#)
3. [Aclara el presupuesto necesario](#)
4. [Hace que los datos sean FAIR](#)
5. [Muestra responsabilidad](#)

Plan de Gestión de Datos (PGD): contenido

1. Resumen de los datos: se debe describir de forma general los datos que se utilizarán en el proyecto. Incluye el tipo y el formato, la finalidad, el tamaño y el origen de los datos.

2. Datos FAIR: se debe asegurar que los datos sean:

- Localizables
- Accesibles
- Interoperables
- Reutilizables

3. Otros resultados de la investigación: se consideran qué cuestiones relativas a los datos FAIR pueden aplicarse a la gestión de otros productos de la investigación generados o reutilizados en el proyecto como por ejemplo: software, flujos de trabajo, protocolos, nuevos materiales, muestras, etc.

Plan de Gestión de Datos (PGD): contenido

4. Asignación de recursos: se indican cuáles serán los costes para hacer que los datos sean FAIR (por ejemplo, costes directos e indirectos relacionados con el almacenamiento, el archivo, la reutilización, la seguridad, etc.) y quién será responsable de la gestión de los datos.

5. Seguridad de los datos: se garantiza que los datos se almacenan de forma segura en repositorios de confianza para su preservación y conservación a largo plazo.

6. Ética: se presentan todas las cuestiones éticas o legales que puedan tener un impacto en el intercambio de datos. Además, cuando la investigación utilice datos personales, deberá hacerse referencia a aspectos como el consentimiento informado y la conservación a largo plazo.

7. Otras cuestiones: se describe si se van a utilizar, otros procedimientos (nacionales, de proveedores, sectoriales, departamentales) para la gestión de datos.

Plan de Gestión de Datos (PGD): plantillas

- [Previsión del PGD](#) propuesta del Plan para presentar junto con la memoria técnica cuando se solicita la financiación de un proyecto de investigación si los datos se van a depositar en idUS.
- [Instrucciones para la elaboración del PDG](#) de la Agencia Estatal de investigación
- [PGD Horizonte Europa \(Español\)](#) proyectos financiados y que utilicen idUS, el repositorio de la Universidad de Sevilla para depositar los datos.
- [PGD Horizon Europe \(English\)](#) para proyectos financiados del Horizonte Europa.
- [Plantilla básica de un PGD](#) elaborada por las Bibliotecas del grupo G9 de Universidades.
- https://dmp.csuc.cat/public_templates
- [Guía de Ayuda para el PGD](#) guía para elaborar un PGD.

Plan de Gestión de Datos (PGD): herramientas



PGDonline

[PGDonline](#)

Herramienta del consorcio de universidades Madroño, adaptada y traducida al español de la herramienta Argos (OpenAIRE).



[Argos](#)

Herramienta desarrollada por OpenAIRE para la creación, gestión y difusión de un PGD



[DMP Online](#)

Herramienta desarrollada por el [Digital Curation Center](#) . Muy recomendable su [Check List for a Data Management Plan](#).



[eiNa DMP](#)

Herramienta en línea, desarrollada por el CSUc, basada en el modelo creado por el DCC. Incluye ejemplos reales en el cuestionario para ayudar en la elaboración del Plan de Gestión de Datos.

Plan de Gestión de Datos (PGD): ejemplos

- <https://www.dcc.ac.uk/resources/data-management-plans/guidance-examples>
- https://dmptool.org/public_plans
- <https://zenodo.org/search?q=data%20management%20plan&l=list&p=1&s=10&sort=bestmatch>
- https://dmponline.dcc.ac.uk/public_plans
- https://dmp.csuc.cat/public_plans

Durante el proyecto: buenas prácticas para la gestión

1. Formatos
2. Nombre de los archivos
3. Organización de los datos
4. Control de versiones
5. Almacenamiento
6. Seguridad
7. Documentación de los datos



1. Usar formatos abiertos

- Formatos estándar y abiertos para asegurar el acceso a largo plazo.
- Guardar los archivos en formatos abiertos y propietarios (por ejemplo, data.xlsx y data.csv) para mejorar la usabilidad.
- Mantener el mismo nombre de archivo para el mismo archivo en diferentes formatos (por ejemplo, data.doc y data.txt).
- Comprimir los datos para facilitar el intercambio y la descarga de archivos grandes.
- Incluir la documentación de los datos, las transformaciones y las directrices del software para acceder al formato propietario.
- Comprobar si hay errores u omisiones al convertir el archivo a un formato diferente.

Un formato abierto es aquel donde las especificaciones del software están disponibles para cualquier persona, de forma gratuita.

Herramientas para la conversión de formatos

- [OpenRefine](#)
- [Mr. Data Converter](#)
- [Format Factory](#)
- [Conversor a RDF](#)

1. Usar formatos abiertos

Tipo de datos	Formatos recomendados	Formatos aceptados
Datos textuales	rtf, txt, xml	html, doc / docx
Datos tabulares con metadatos mínimos	csv, tab	txt, xls / xlsx, mdb / accdb, dBase dbf, ods
Datos tabulares con metadatos extensos	SPSS por, (SPSS, Stata, SAS, XML DDI	SPSS sav, Stata dta, MS Access mdb / accdb
Datos geoespaciales	shp, shx, dbf, prj, sbx, sbn, tif, tfw, dwg, gml	mdb, mif, kml, ai, dxf o svg
Datos de imagen	tif	jpeg, jpg, jp2, gif, tif, tiff, raw, psd, bmp, png, pdf
Datos de audio	flac	mp3, aif, wav
Datos de video	mp4, ogv, ogg, mj2	avchd
Documentación y guiones	rtf, PDF / UA, PDF / A o PDF pdf, XHTML o HTML xhtml, htm, odt	txt, doc / docx, MS Excel xls / xlsx, xml

Más información:

- [UK Data Service](#)
- <https://dans.knaw.nl/en/file-formats/>
- [Cómo elegir el formato correcto para los datos abiertos"](#) Curso [European Data Portal](#).
- [Formatos soportados por idUS](#)

2. Nombrar los archivos

- Nombrar los archivos con nombres cortos y relevantes.
- No utilizar caracteres especiales:
~ ! @ # \$ % ^ & * () ` ; < > ¿ ? , [] { } ' " |
- Usar el guion bajo, mejor que el espacio en blanco.
- Identificar la actividad o proyecto en el nombre del archivo.
- Ser consistente con la nomenclatura que se escoge, mayúsculas, minúsculas, forma de las fechas, AAAA-MM-DD o AAAA-MM (ISO 8601 aplicada a la gestión de datos de investigación).
- Conservar las extensiones de 3 letras específicas a códigos de la aplicación (p. ej., .doc, .xls, .mov, .tif).

La nomenclatura ha de ser significativa, coherente y descriptiva, de forma que facilite la gestión y localización de archivos

Herramientas para renombrar ficheros:

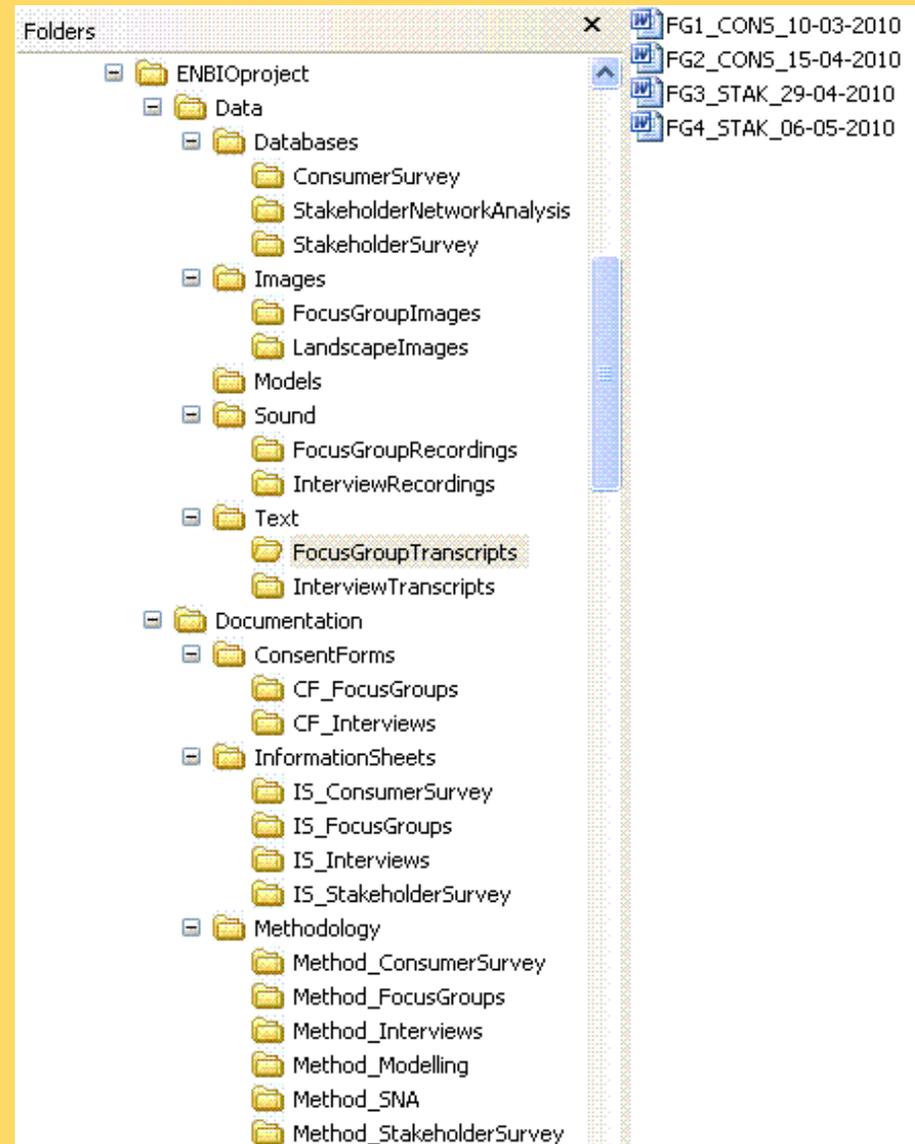
- [Bulk Rename Utility](#)
- [Renamer 6](#)

3. Organizar los archivos

El nombre, la estructura de las carpetas y el control de versiones de los ficheros deben facilitar la búsqueda, localización y comprensión de los datos.

- Pensar la mejor jerarquía para los archivos: profunda o superficial.
- Organización sistemática de las carpetas y ficheros.
- Restringir el nivel de carpetas a tres o cuatro.
- Separar los trabajos finalizados de los que los que están en curso.

<https://ukdataservice.ac.uk/learning-hub/research-data-management/format-your-data/organising/>



4. Control de versiones

- Si hay varias versiones nombrar por número (v01, v02, etc.).
- Para la versión final se puede nombrar como FINAL.
- Decidir cuántas versiones de un archivo y qué versiones se conservarán y durante cuánto tiempo.
- Registrar los cambios realizados en un archivo cuando se crea una nueva versión.
- Realizar un seguimiento de la ubicación de los archivos si están almacenados en distintas ubicaciones.
- Elegir una sola ubicación para las versiones importantes o definitivas.

Herramientas para gestionar versiones:

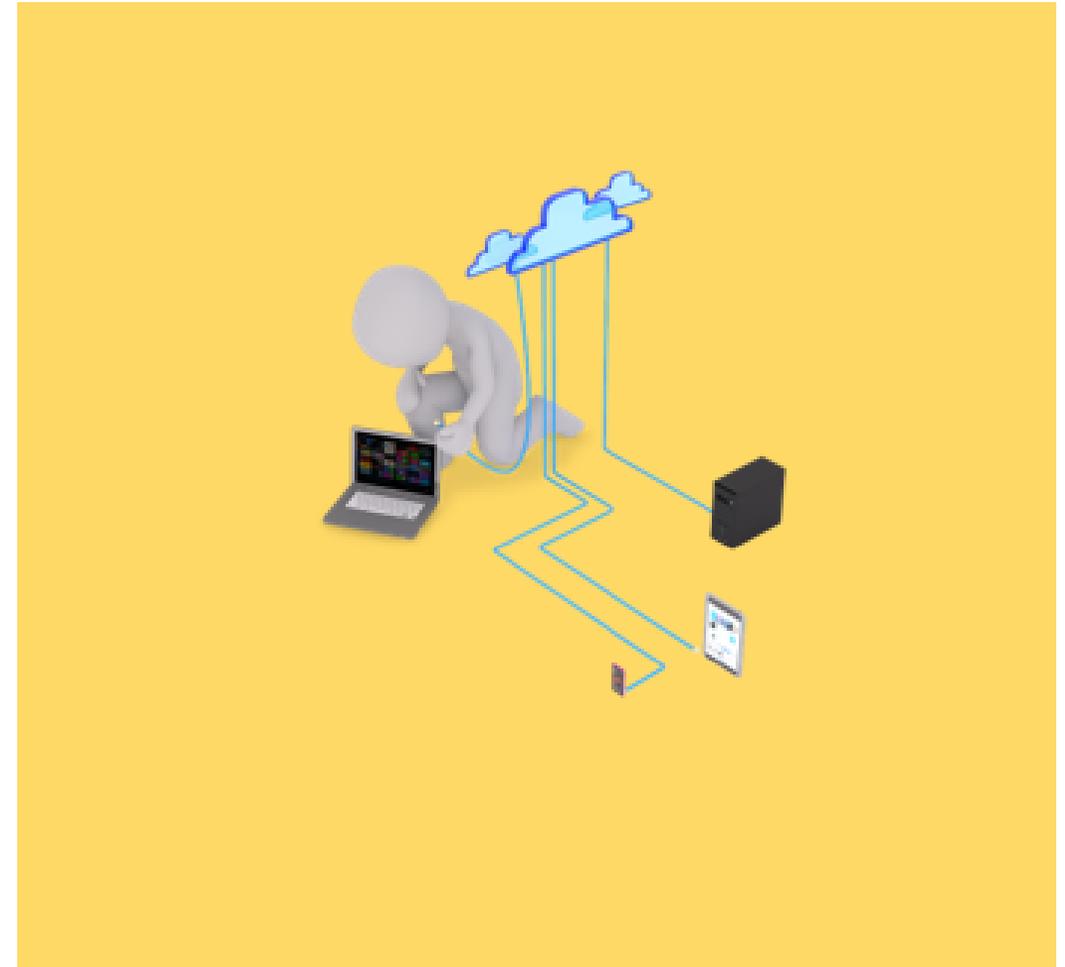
- [Apache Subversion](#)
- [Open Science Framework \(OSF\)](#)

5. Almacenamiento

Qué sistemas o herramientas se utilizarán para almacenar los datos durante y al final del proyecto, qué personas accederán, quien controlará el acceso a los datos y cómo se controlará que el acceso sea seguro y cómo se harán las copias de seguridad.

Los datos se pueden almacenar en:

- ✓ PC/Ordenador portátil.
- ✓ Almacenamiento óptico (CDs, DVDs).
- ✓ Unidades externas (USB, discos duros).
- ✓ Servicios de almacenamiento institucional.
- ✓ Servicios en la nube (OneDrive, Google drive...).



5. Almacenamiento

Regla 3-2-1

3 copias

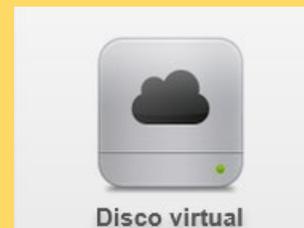
2 medios de almacenamiento

1 copia en una ubicación distinta

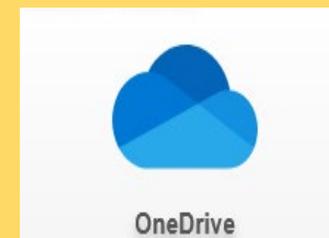


Utilizar los servicios de almacenamiento institucionales.
En el caso de la US

Disco Virtual



OneDrive



En algunas disciplinas hay también servicios de almacenamiento:

- [BlueBridge](#) para Ciencias del Mar
- [Elixir](#) para Ciencias de la vida
- [Software Heritage](#) para software en formato de código fuente

6. Seguridad

- Utilizar contraseñas robustas y cambiarlas cada cierto tiempo para resistir ataques informáticos y encriptación de datos altamente sensibles.
- No enviar datos por correo electrónico.
- Antivirus.



7. Documentación de los datos

- Ayuda a interpretar los datos permitiendo reutilizarlos de forma rápida y ágil.
- Creación del fichero **Readme.txt** para adjuntar a los datasets. Fichero en que se describe la información necesaria para que los datos sean comprensibles y reutilizables. Se debe de publicar en inglés.
- Diccionario de datos

Plantillas archivo readme.txt

idUS

<https://guiasbus.us.es/c.php?g=164595&p=5003623>

Consortio Madroño <https://edatos.consociomadrono.es/readme.xhtml>

Cornell University <https://data.research.cornell.edu/content/readme#fileoverview>

CSUC <https://confluence.csuc.cat/display/RDM/Plantilla+Fitxer+Readme.txt>

7. Documentación de los datos

Estructura de un archivo Readme:

- Título
- Identificador (DOI de los datos, o identificador del proyecto)
- Ubicación del archivo con los datos
- Licencia/Condiciones de uso

(Disclaimer: Los creadores de los datos y de su archivo no son responsables de ningún resultado o interpretación que pueda surgir de su reuso).

- Creador/a y afiliación
- Descripción de los datos
- Temas o palabras clave de los datos
- Formato de los datos
- Lugar de obtención de los datos
- Método de obtención de los datos
- Fecha de obtención de los datos

GENERAL INFORMATION

1. Dataset title:
[Obligatorio]

2. Authorship:
[Obligatorio | Ponga la información de todos los autores siguiendo el siguiente forma

Name:
Institution:
Email:
ORCID:

DESCRIPTION

1. Dataset language:
[Obligatorio]

2. Abstract:
[Obligatorio | Resumen del conjunto de datos, no de los resultados de la investigación diferentes idiomas.]

3. Keywords:
[Obligatorio | Mínimo 3 palabras claves. Puede añadirlas en diferentes idiomas.]

4. Date of data collection (fecha única o rango de fechas):
[Obligatorio | Formato DD-MM-YYYY]

5. Publication Date:
[Obligatorio si es aplicable. Fecha de depósito en el repositorio | Formato DD-MM-YYYY]

6. Grant information:
[Obligatorio si es aplicable | Financiación recibida. Repita el esquema para cada organis

Grant Agency [Organismo financiador]:
Grant Number [Código del proyecto]:

Al final del proyecto: depósito/licencias/citas

- ¿Qué datos vamos a depositar?
- ¿Dónde depositar los datos? Página web personal/ data journal/repositorio temático o institucional
- ¿Qué licencia le vamos a asignar a los datos?: Creative Commons /Open Data Commons
- ¿Cómo se citan los datos?

Al final del proyecto: depósito de los datos

Depositar los datos en un repositorio temático o institucional:

- Que cumpla con los criterios FAIR.
- Que permita almacenar datos abiertos, con embargo, restringidos o cerrados, en especial para datos sensibles.
- Que sirva para materiales diversos de investigación (distintos tipos, formatos, etc.).
- Que proporcione visibilidad, para facilitar la citación y aumentar el impacto.
- Que se integre con otras infraestructuras.
- Que sea transparente y fiable.



Directrices para elegir repositorio

- [Practical Guide to the International Alignment of Research Data](#)
- [Recomendaciones para seleccionar un repositorio donde depositar datos de investigación](#) del Consorci de Serveis Universitaris de Catalunya (CSUC).

Al final del proyecto: depósito en idUS



idUS

- Identificador persistente (DOI/handle).
- Estándar de metadatos general Dublin Core.
- Recolectado por motores de búsqueda y directorio de datos.
- Licencia de reutilización (Creative Commons).
- Estadísticas de visitas y descargas.
- 2 Gb gratuitos para cada dataset (si necesitas más espacio contacta con la Biblioteca).
- Garantiza la preservación a largo plazo.



Al final del proyecto: depósito en idUS

idUS / Investigación / Ciencias Sociales y Jurídicas / Geografía Humana / Datos de Investigación (Geografía Humana) / Ver ítem

Dataset

Entrevistas vinculadas al capítulo "Mapping urban integration of archaeological heritage. The city centre of Malaga, Andalusia, Spain" [Dataset]

Título alternativo	Interviews related to the book chapter "Mapping urban integration of archaeological heritage. The city centre of Malaga, Andalusia, Spain" [Dataset]
Autor/es	Quetglas-Llull, Llorenç Hernández-Escampa, Marco Barrera Fernández, Daniel
Compilador de datos	Barrera Fernández, Daniel
Gestor de datos	Barrera Fernández, Daniel
Departamento	Universidad de Sevilla. Departamento de Geografía Humana
Idioma (ISO)	español (spa)
Fecha de difusión	2022
Fecha de creación	2022
Resumen	Entrevistas realizadas para la publicación de un capítulo de libro actualmente en proceso editorial. La información de citación actual es la siguiente: Autores: Llorenç Quetglas-Llull, Marco Hernández-Escampa y Daniel ...
Tipo de dataset	Estadísticas

Mostrar el registro completo del ítem

Documentación necesaria:

- [Plantilla descripción de los datos](#)
- [Licencia no exclusiva de distribución](#)
- Fichero de datos
- Readme.txt

https://guiasbus.us.es/idus/datos_investigacion

<https://idus.us.es/handle/11441/137336>

Al final del proyecto: asignar licencia de uso

Una **licencia** es un contrato legal que especifica restricciones y permisos estándar para un trabajo, que luego puede ser utilizado por cualquier persona a quien se le otorgue la licencia, aunque los derechos de propiedad no se transfieren.

Para los datos abiertos se recomienda asignar la licencia **CC-BY**

Existen diferentes tipos de licencias:

- [Open Data Commons](#)
- [Creative Commons \(CC\)](#)
- [Elegir una licencia CC](#)

Para saber más sobre Creative Commons, puedes visualizar [este vídeo](#).

LAS COMBINACIONES, DE UN VISTAZO

	SI VEO ESTA LICENCIA...	YO PUEDO...	SI...
MÁS PERMISIVA	Open Data Commons Derechos Públicos (CC0) European, Eprints, Open Goldberg V.	• Compartir • Copiar • Remezclar • Ganar dinero	• Menciona al autor (en algunas jurisdicciones)
	Reconocimiento (BY) BY, FLOJ, Saylor.org.	• Compartir • Remezclar • Ganar dinero	• Menciona al autor
	Reconocimiento - Compartir Igual (BY-SA) BY-SA, Wikipedia, Wikimedia, Arduino, POPU	• Compartir • Remezclar • Ganar dinero	• Menciona al autor • Mantiene la misma licencia (BY-SA)
	Reconocimiento - Sin Derivadas (BY-NC) BY-NC, Drupal, Bohemia, GNU Free Software Foundation.	• Compartir • Ganar dinero	• Menciona al autor • No hago remezclas
	Reconocimiento - No Comercial (BY-NC) BY-NC, Brooklyn Museum, Wired.com Photography	• Compartir • Remezclar	• Menciona al autor • No gana dinero
	Reconocimiento - No Comercial - Compartir Igual (BY-NC-SA) BY-NC-SA, MIT Open CourseWare	• Compartir • Remezclar	• Menciona al autor • No gana dinero • Mantiene la misma licencia (BY-NC-SA)
MÁS RESTRICTIVA	Reconocimiento - No Comercial - Sin Derivadas (BY-NC-ND) BY-NC-ND, Videos TED Talks, Propublica	• Compartir	• Menciona al autor • No hago remezclas • No gana dinero • Mantiene la misma licencia (BY-NC-ND)

crea Cultura

Al final del proyecto: citar los datos

Los datos deben citarse siguiendo un formato de citas específico.

[DOI Citation Formatter](#) es un servicio ofrecido por [DataCite](#) que construye automáticamente las citas conforme al estilo que seleccionemos a partir del DOI asignado a los datos.

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.).

Elabora 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.).

Elementos de la cita:

- Autor(es) (O)
- Identificador autor (R)
- Fecha (O)
- Título (O)
- Identificador único persistente (O)
- Tipo de recurso (O)
- Versión y Edición (O)
- Repositorio de datos (R)
- Publicación (R)
- Productor (R)
- Ámbito geográfico (R)
- Ámbito temporal (R)

Ejemplo de cita estilo APA

Autoría: Remesar Betloch, X., Antelo, A., Uvina, C., Albà, E., Berdié, L., Agnelli, S.,... Alemany, M.
 Título: (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.
 Diposit digital de documents de la UAB: <http://hdl.handle.net/2445/66010>.

Elementos de la cita en el ejemplo:

- 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

CC BY

A modo de resumen

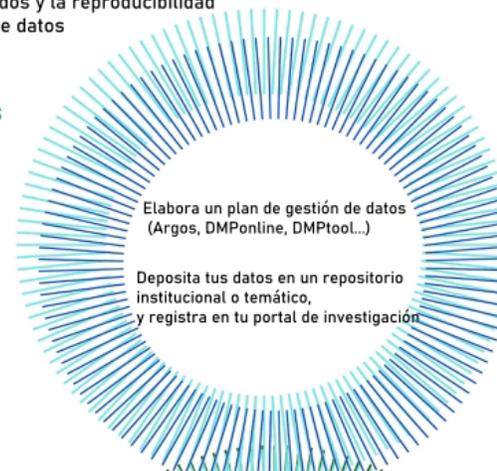
[Infografía sobre datos abiertos \(elaborada por el grupo de trabajo de Rebiun Línea 2. Subgrupo Acceso abierto. Datos de investigación\)](#)

Los datos de investigación importan: gestiona, comparte, publica y reutiliza

Compartir tus datos

- Promueve la innovación
- Facilita la colaboración
- Maximiza la transparencia y la fiabilidad
- Permite la verificación de los resultados y la reproducibilidad
- Reduce costes al evitar duplicación de datos

Planifica la gestión de tus datos

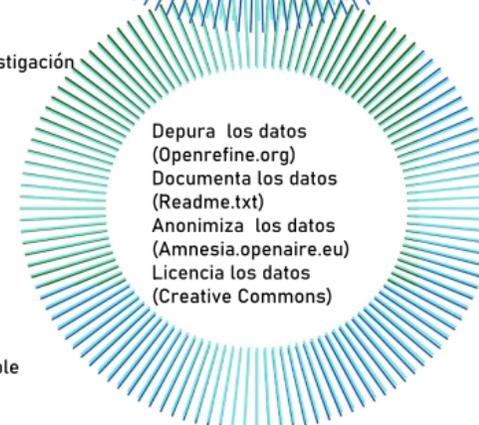


Políticas institucionales y legislativas, agencias de financiación, requieren abrir y compartir los datos de investigación

Haz tus datos



Localizable Accesible Interoperable Reutilizable



¡Acude a tu biblioteca! te apoya y aconseja

Muchas gracias

bibinvestigacion@us.es

Sección de Apoyo a la Investigación

Biblioteca de la Universidad de Sevilla

