

# Publicación de datos de investigación en abierto

Formadora: Cristina Azorín  
Moderadora: Marta Jordán



# Normas y recomendaciones

- Los micros están silenciados para facilitar la comprensión de las explicaciones de la formadora
- Está prohibido grabar (vídeo o audio), directa o indirectamente, total o parcialmente, la sesión
- Podéis deshabilitar vuestra cámara si lo creéis conveniente
- Podéis hacer preguntas a través del chat. Se contestaran al final de la sesión

# ¿De qué hablaremos?

## DATOS DE INVESTIGACIÓN

01

QUÉ SON Y TIPOLOGÍA

02

CONVOCATORIAS

Nacionales e internacionales

03

PRINCIPIOS FAIR

04

EL PLAN DE GESTIÓN DE DATOS

05

EL DDD

Dipòsit Digital de Documents de la UAB

06

PUBLICAR DATOS EN EL DDD  
Requisitos

07

OTROS REPOSITORIOS  
Para los datos de investigación

08

BUENAS PRÁCTICAS

En la publicación de los datos





# Ciencia Abierta: La investigación y los datos científicos accesibles y abiertos a todos los ciudadanos

## Open Science

**Open Repositories**  
Repositorios Abiertos



**Open Access**  
Acceso Abierto

Acceso sin trabas económicas, tecnológicas o jurídicas a las publicaciones científicas

**Open Access Journals**  
Revistas de Acceso Abierto



**Open Reproducible Research**  
Investigación Reproducible en Abierto

Acceso libre a los elementos experimentales para la reproducción de la investigación



**Open Research Data**  
Datos de Investigación Abiertos

**Open Source in Open Science**  
Código Abierto para la Ciencia Abierta

**Open Peer Review**  
Revisión por Pares Abierta



**Open Science Evaluation**  
Evaluación de la Ciencia en Abierto

Evaluación abierta de los resultados de investigación, ampliando la revisión tradicional con la contribución de la comunidad

**Open Metrics and Impact**  
Impacto y Métricas Abiertas



**Open Data**  
Datos Abiertos

Datos que están disponibles en línea de forma gratuita y que se pueden usar, reutilizar y distribuir



**Open Big Data**  
Datos Masivos Abiertos

**Open Government Data**  
Datos Gubernamentales Abiertos



# ¿Qué son los datos de investigación?

Según la Comisión Europea:

- hacen referencia a información, **factual** o **numérica**
- recogida para ser **examinada** y **considerada**
- sirve de base al **razonamiento**, la **discusión** o el **cálculo**

Ejemplos: estadísticas, resultados de experimentos, medidas, observaciones en trabajo de campo, encuestas, entrevistas e imágenes.

# Tipología de los datos de investigación

- Datos generados en el proceso de investigación o **datos primarios**: deben ordenarse, documentarse y se puede solicitar asesoramiento en caso de tratarse de datos personales o sensibles.
- Datos asociados a resultados publicados o **datos finales**: deben identificarse y facilitar el acceso y la reutilización.



# El ciclo de los datos científicos



Ayuda a planificar la investigación



El personal de tu Biblioteca te puede asesorar



# Beneficios de publicar los datos

- Refuerza la **ciencia abierta**: datos disponibles libremente
- Visibilidad y aumento del **impacto**
- Identificación con **DOI**; permite la citación y las métricas
- Permite la verificación y reproducción de los resultados: **transparencia de la investigación**
- Promociona la innovación a través de la **compartición** y la **reutilización** de los datos
- Evita la duplicidad en la obtención y recogida de datos: **optimización** de tiempo, costes y esfuerzos
- Facilita la **colaboración** y el **debate**: anima la diversidad de análisis y opiniones

# Convocatorias – Plan Estatal (2017-2020)

## GENERACIÓN DE CONOCIMIENTOS E INFRAESTRUCTURAS DE INVESTIGACIÓN

- Posicionamiento científico internacional (O1)
- Participación en H2020, grandes proyectos e infraestructuras de investigación internacionales (O2)
- Capacidad tractora de centros e instituciones sobre el conjunto del Sistema (O3)
- Valorización de la I+D+i pública (R1)
- Impacto científico y social (R2)
- Fortalecimiento y acceso a ICTs, interoperabilidad y modernización de infraestructuras y equipación (R3)
- Redes de colaboración científica e integración transdisciplinar (R4)
- Difusión de los resultados a la sociedad (R5)
- Acceso Abierto y Ciencia Abierta (R5)



Los proyectos pueden incluir, optativamente, un Plan de Gestión de datos de investigación y el depósito de estos en repositorios institucionales



# Cómo cumplir con los mandatos sobre gestión y publicación de datos en Horizonte 2020

Programa Horizonte 2020 (art. 29.3)



## A quién afecta

**Investigadores** con proyectos subvencionados por Horizonte 2020



Por razones de confidencialidad, seguridad, explotación industrial... **puede no publicar sus datos**



## Qué obliga a depositar

**Los datos**, incluidos sus metadatos, necesarios para validar los resultados presentados en las publicaciones científicas

**Otros datos**, incluidos sus metadatos, especificados en los planes de gestión de datos de los proyectos de investigación



## Requisitos

**Desarrollar y mantener** un Plan de Gestión de Datos

**Depositar** los datos en un repositorio de datos de investigación

**Indicar** qué herramientas se requieren para usar los datos

**Permitir** el acceso, explotación y disseminación de datos

## Ventajas

- Permiten **validar los resultados** presentados en publicaciones científicas y otras fuentes de información
  - Permiten basarse en los **resultados de investigaciones previas**
  - **Fomenta la colaboración** y evita la duplicación de esfuerzos
    - **Acelera la innovación**
- Mejora la **transparencia del proceso científico**



Recomendado el uso de licencias Creative Commons



# Convocatorias – Horizon Europe (2021-2027)



## WHAT'S NEW?



**Open Science** will become the modus operandi of Horizon Europe. It will go beyond the open access policy of Horizon 2020 and **require open access to publications, data, and to research data management plans.**



# Principios FAIR



## Findable:

- ✓ Identificador persistente: DOI (preferentemente) i/o URI
- ✓ Publicación en repositorios institucionales o temáticos que describan los datos siguiendo estándares internacionales
- ✓ Los repositorios han de facilitar que todos los registros se puedan encontrar fácilmente y permitir su recuperación

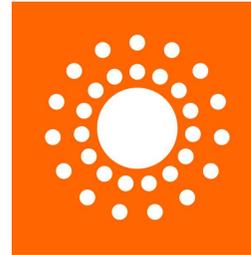
# Principios FAIR

## Accessible:

- ✓ Los datos y los metadatos serán accesibles porque los repositorios utilizan protocolos de comunicación estándares que permiten el intercambio de información entre plataformas



# Principios FAIR



## Interoperable:

✓ Los metadatos han de cumplir unos estándares internacionales (Dublin Core, DataCite...), en cuanto a códigos y nomenclaturas



# Principios FAIR



## Reusable:

- ✓ Los datos deben ser descritos exhaustivamente y en detalle: origen, fecha de recogida, circunstancias ambientales, observaciones...
- ✓ Para la descripción se utilizarán los estándares de cada área temática. Si ello no es posible se debe utilizar un estándar general reconocido
- ✓ Debe haber una mención de derechos para cada dataset (licencias CC o en dominio público)

# Plan de Gestión de Datos (PGD)

El Data Management Plan (DMP) describe el ciclo de vida de gestión de los datos que se producen durante la investigación.

El plan debería incluir información sobre:

- la gestión durante y después del final del proyecto
- tipos de datos que se recopilarán y procesarán
- metodología y estándares aplicados
- qué datos estarán disponibles en acceso abierto
- preservación de los datos durante y después del final del proyecto



# 10 pasos para elaborar un Plan de Gestión de Datos

Un **Plan de Gestión de Datos** (PGD) o Data Management Plan (DMP) es un **documento formal, que debe presentarse al inicio de la investigación, en el que se describe qué**

**vas a hacer con tus datos durante y después de finalizar tu investigación** y que puede modificarse si se producen cambios en el proceso de la misma.

## ¿Por qué?

Es una **buena práctica**, es un **elemento clave de Open Science** y es **obligatorio** en los nuevos proyectos H2020.

### Herramientas gratuitas para elaborar un PGD



PGDonline  
(Consortio Madroño)  
<http://dmp.consortiomadrono.es/>



DMPonline (Digital Curation  
Centre, UK)  
<https://dmponline.dcc.ac.uk/>



Revisa los **requerimientos** de la entidad financiadora (H2020).



**Identifica los datos:** tipología, procedencia, volumen, formatos y ficheros.



**Define cómo se organizarán y gestionarán los datos:** nombre de los ficheros, control de versiones, software necesario...



**Explica cómo se documentarán los datos:** identifica la información a procesar, consulta si hay estándares o esquemas de metadatos, identifica herramientas que permitan gestionarlos.



Describe los procesos que aseguran una **buena calidad de los datos**.



**Prepara una estrategia de almacenamiento** (durante el proceso) y de preservación de datos (repositorio).



**Define las políticas de datos del proyecto:** cuestiones sobre propiedad intelectual y cómo se tratarán los datos sensibles y personales.



**Describe cómo se difundirán los datos:** dónde, cuáles, cuándo se van a difundir. Si publicarás los datos en un repositorio, como información suplementaria del artículo o como un "data paper".



**Asigna roles y responsabilidades** para las personas y organizaciones participantes en el proyecto.



**Prepara un presupuesto realista:** la gestión de datos cuesta tiempo y dinero en términos de software, hardware, servicios y personal.



# PGD para la tesis doctoral

Consorci de Serveis Universitaris de Catalunya (CSUC)

**Guia per elaborar un pla de gestió de dades per a doctorands**

[ddd.uab.cat/record/218680](http://ddd.uab.cat/record/218680)

**Portal DMP** – plantilla que facilita la elaboració del PGD

[dmp.csuc.cat](http://dmp.csuc.cat)

**eiNa DMP**  
Pla de Gestió de Dades de Recerca

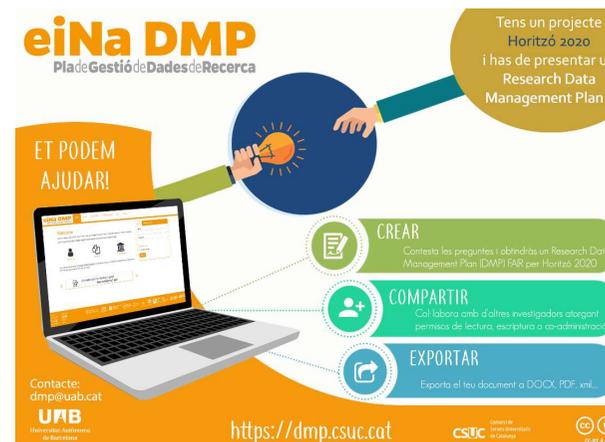
# PGD – Material de apoyo



- Videotutorial: [youtu.be/51FHTJN\\_0WQ](https://youtu.be/51FHTJN_0WQ)



- Infografía: [ddd.uab.cat/record/181194](http://ddd.uab.cat/record/181194)



# El Dipòsit Digital de Documents UAB



## Datasets

- [ddd.uab.cat/collection/datasets](http://ddd.uab.cat/collection/datasets)

## Planes de Gestión de Datos

- [ddd.uab.cat/collection/plagesdad](http://ddd.uab.cat/collection/plagesdad)

# El DDD es...

<b>F</b>	Asigna identificadores (DOI/URI)	
	Utiliza estándares internacionales	
	Facilita la búsqueda y recuperación	
<b>A</b>	Datos y metadatos accesibles a través de protocolos de comunicación estándares	
<b>I</b>	Metadatos normalizados	
<b>R</b>	Descripción exhaustiva de los datos	
	Mención de derechos (licencias CC o dominio público)	
	Descripción según estándares temáticos o generales ampliamente reconocidos	

# Publicar los datos en el DDD

- ✓ identificador ORCID
- ✓ resumen del contenido (puede ser en varios idiomas)
- ✓ palabras clave (en varios idiomas)
- ✓ licencia de reutilización ([recomendaciones de la UAB](#))
- ✓ indicar el código del proyecto (si fuese el caso)
- ✓ indicar la obra relacionada (artículo, tesis, libro...)
- ✓ 5 Gb gratuitos para cada dataset (si necesitas más espacio contacta con la Biblioteca)



# Otros repositorios

A parte del DDD, hay otros repositorios que permiten incluir datasets



Encontraréis las características de cada un en:  
[ddd.uab.cat/record/150829](http://ddd.uab.cat/record/150829)

# Otros repositorios – criterios de selección

- ✓ ¿Existe algún repositorio consolidado en tu ámbito?
- ✓ ¿Permite publicar los datos en el formato que utilizas?
- ✓ ¿El tamaño de ficheros que puedes depositar se ajusta a tus necesidades?
- ✓ ¿Se permiten tipologías de acceso (abierto, restringido, embargado...)?
- ✓ ¿Los derechos y licencias que te ofrecen se ajustan a tus intereses?
- ✓ ¿El repositorio genera un identificador permanente?
- ✓ ¿Las condiciones de depósito y retirada se ajustan a tus necesidades?
- ✓ ¿Hay costes asociados?
- ✓ ¿Se pueden enlazar a las publicaciones asociadas?
- ✓ ¿Existe una política de preservación (copias, plazos de conservación...)?
- ✓ ¿Cumple los requisitos de posibles becas o ayudas?
- ✓ ¿Permite depositar varias versiones?

# Buenas prácticas



- ✓ añadir un archivo tipo README con los detalles del dataset (contenidos de cada fichero, unidades de medida empleadas, vía y condiciones de obtención de los datos, etc.)
- ✓ nombres de ficheros identificables (mejor emplear alguna convención)
- ✓ ficheros en formatos abiertos (txt, csv, pdf, LaTeX, jpg, MP3...)
- ✓ un PGD que acompañe a los datos... sube nota!
- ✓ citad los datos, los vuestros y los que utilizéis de terceros

**Recordad!**

**Cualquier persona que consulte vuestros datos,  
debe entenderlos para poderlos reutilizar**

# Ejemplos

- ✓ Relación entre el documento y los datos en el DDD

<https://ddd.uab.cat/record/182292>

**Document:** dataset ; recerca ; publishedVersion  
**Matèria:** Google news ; Agregadores de noticias ; Medios noticiosos ; Noticias ; Infomediación  
**Obra relacionada:** Tesis doctoral: "Medios de comunicación iberoamericanos y agregadores de noticias: Portugal".

**Enllaç a la tesi doctoral:** <https://ddd.uab.cat/record/188096>  
~~DOI: 10.5565/ddd.uab.cat/182292~~

- ✓ Disponibilidad de los datos desde la publicación

<https://doi.org/10.1007/s10021-020-00500-z>

## Data Availability

Data supporting the findings of this study (DBD, water content, grains size distribution, C and N contents,  $\delta^{13}\text{C}$  and  $\delta^{15}\text{N}$ ,  $^{210}\text{Pb}$  and DOC) are available at <https://ddd.uab.cat/record/216456> with the identifier <https://doi.org/10.5565/ddd.uab.cat/216456>



# 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:

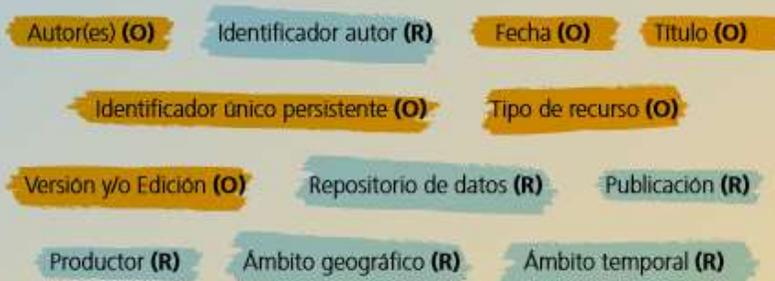
- Se debe facilitar la identificación, localización y el acceso a los datos mediante un identificador único y persistente (DOI, Handle, etc.).
- Cada conjunto y subconjunto de datos (dataset) debe citarse de forma independiente.
- Las citas de los datos utilizados han de aparecer en la sección de referencias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



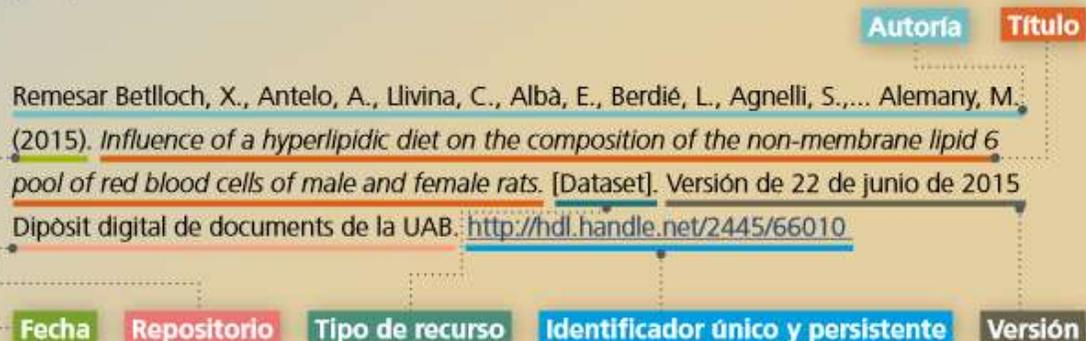
Enlaza los datos con los documentos resultado de investigación y viceversa, y crea las referencias bibliográficas de los mismos.

## 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





# Bibliografía

- Consorci de Serveis Universitaris de Catalunya. Recomanacions per seleccionar un repositori per al dipòsit de dades de recerca [Internet]. 2019 [Consulta 27 gener 2020]. Disponible a: <https://ddd.uab.cat/record/163562>
- España, Ministerio de Economía, Industria y Competitividad. Plan Estatal de Investigación Científica y Técnica y de Innovación 2017-2020 [Internet]. [Madrid]: Gobierno de España, [2017] [Consulta 15 gener 2020]. Disponible a: <http://www.ciencia.gob.es/stfls/MICINN/Prensa/FICHEROS/2018/PlanEstatalIDI.pdf>
- European Commission. EU funding for research and innovation 2021-2027 [Internet]. [Brussels]: European Commission, 2018 [Consulta 14 gener 2020]. Disponible a: [https://ec.europa.eu/commission/sites/beta-political/files/budget-may2018-research-innovation\\_en.pdf](https://ec.europa.eu/commission/sites/beta-political/files/budget-may2018-research-innovation_en.pdf)
- European Commission. H2020 Programme: Guidelines on FAIR Data Management in Horizon 2020 [Internet]. [Brussels]: European Commission, 2016 [Consulta 21 gener 2020]. Disponible a: [https://ec.europa.eu/research/participants/data/ref/h2020/grants\\_manual/hi/oa\\_pilot/h2020-hi-oa-data-mgt\\_en.pdf](https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf)

- European Commission. H2020 Programme: Guidelines to the rules on Open Access to scientific publications and Open Access to research data in Horizon 2020 [Internet]. [Brussels]: European Commission, 2017 [Consulta 30 gener 2020]. Disponible a: [https://ec.europa.eu/research/participants/data/ref/h2020/grants\\_manual/hi/oa\\_pilot/h2020-hi-oa-pilot-guide\\_en.pdf](https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf)
- Red de Bibliotecas Universitarias Españolas (REBIUN). Datos de investigación y acceso abierto [Internet]. [Consulta 21 gener 2020]. Disponible a: <https://ddd.uab.cat/record/165475>
- Zenodo. FAIR principles [Internet]. Genève: CERN, [201?] [Consulta 27 gener 2020]. Disponible a: <http://about.zenodo.org/principles/>

**Podéis consultar esta presentación desde:**  
<http://ddd.uab.cat/record/222106>

# Para acabar

- ✓ A través del correo electrónico recibiréis:
  - El enlace a la presentación para su consulta y descarga
  - El enlace a una encuesta que nos ayudará a mejorar
  - El certificado de asistencia
  - Preguntas y respuestas del chat

- ✓ Asesoramiento personalizado a: Pregunt@  
<https://www.bib.uab.cat/pregunta/pregunta.php>





# Muchas gracias!

[#bibliotequesUAB](#)