



# INTRODUCCIÓN A LA GESTIÓN DE DATOS DE INVESTIGACIÓN.

Valencia, 20 de junio de 2024

Ricardo Enríquez Malavé rienma@upvnet.upv.es bibliotecadatos@bib.upv.es







- 1. Contexto.
- Mandatos de los financiadores. Infraestructuras europeas. Situación legal.
- 3. Plan de Gestión de Datos.
- 4. Principios FAIR.
- 5. Publicación.
- 6. Depósito en repositorios.







**UPV** 

BIBLIOTECA Y DOCUMENTACIÓN CIENTÍFICA

Universitat Politècnica de València

www.upv.es · rienma@upvnet.upv.es

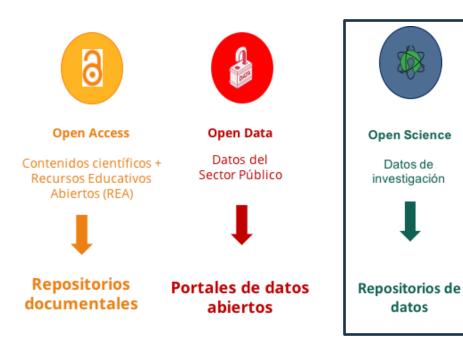




Componentes de la ciencia abierta

## Recomendación sobre la Ciencia Abierta, 41<sup>a</sup> reunión celebrada en París, 9-24 noviembre 2021.

https://unesdoc.unesco.org/ark:/48223/pf0000379949





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

Open Repositories Repositorios Abiertos

**Open Access** Journals Revistas de

Acceso Abierto

### **Open Access** Acceso Abierto

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

Open Science



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

RiuNet



**Open Peer** Review Revisión por Pares Abierta

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

comunidad

Evaluación de la Ciencia en Abierto

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



Open Data

**Datos Abiertos** 

Datos que están disponibles en línea de forma gratuita y que se pueden usar, reutilizar v **Open Big Data Datos Masivos** Abiertos

> Open Government Data

Datos Gubernamentales Abiertos

El personal de tu Biblioteca te puede asesorar



### Catálogo de datos de Open Data UPV

Bienvenido al catálogo de datos abiertos de Open Data UPV. Puede consultar los conjuntos de datos disponibles, y recuerde que puede solicitar un conjunto de datos pulsando en el enlace Solicitud de



**Buscar dato** Ei: ambiente Etiquetas populares Transparencia

**UPVTransparen** 

94

conjuntos de datos



# Amsterdam Call for Action on Open Science

### Eliminar las barreras a la ciencia abierta

- 1. Cambiar los sistemas de evaluación y recompensa en ciencia.
- 2. Facilitar la minería de texto y datos.
- 3. Mejorar la comprensión de Derechos de Propiedad Intelectual.
- 4. Crear transparencia de los costes y condiciones de la comunicación académica.

### Desarrollo de infraestructuras de investigación

- 5. Introducir los Principios FAIR y datos seguros.
- 6. Establecer infraestructuras digitales comunes.

### Fomentar y crear incentivos para la ciencia abierta

- 7. Adoptar principios de acceso abierto.
- 8. Estimular nuevos modelos de publicación para la transferencia de conocimiento.
- 9. Estimular la investigación basada en la evidencia sobre las innovaciones en ciencia abierta.

### Incorporación y promoción de políticas científicas abiertas

10. Implementar, monitorizar y perfeccionar las políticas de acceso abierto.

### Estimular e integrar la ciencia abierta en la investigación y la sociedad

- 11. Involucrar a investigadores y nuevos usuarios en ciencia abierta.
- 12. Alentar a las partes interesadas a compartir conocimientos e información sobre ciencia abierta.

https://www.ua-foundation.org/wp-content/uploads/2016/04/Amsterdam-Call-for-Action-on-Open-Science.pdf



### ACCESO ABIERTO A LAS PUBLICACIONES

## Tipos de versiones



Draft/Discussion Paper (Borrador)

Buena práctica de ciencia abierta



Preprint/Submitted Version (Versión enviada al editor)

Mandato

Postprint/Accepted Manuscript (Versión final del autor con revisión)



Published Version (Versión final publicada)



### Vegetable ashes as Supplementary Cementitious Materials

Martirens F.1, Mouzó J.2

 Universidad Central de las Villos, Santa Clara, Cuba, <u>finantirens € east cu</u>.
 Instituto de Ciencia y Tecnologia del Hormigóa. Universidat Politécnica de Valencia, Spain: <u>impenzo@est.upv.es</u>

#### Abstract

Approximately 140 billion metric tons of biomass are produced every year in the world from agricultura. The sales resulting from firing agricultural waster such as rice land, super case and others can be used as Supplementary Ceracuthious Materials (SCM). They can be raixed with line alone or in ternary mintures with Portland cement and line. If fired at temperatures around 600-700 °C the agricultural askes exhibit good reactivity. Despite excensive research work carried out on the use of agricultural askes as source of SCMs, few success stories are reported an practical applications on an industrial scale. Details of the firing technology should be re-assessed, with special amphasis on the scale at which the technology begins to be contomically suitable. Further research is also needed to understand the mechanisms of structural transformation of anorphous silica during calcination, and the impact of the silves on cement bydration in blended systems.

#### 1 Introduction

Approximately 140 billion metric tons of biomass are produced every year in the world from agriculture [1]. Widely available, renewable, and virtually free, waste biomass is an important resource. This volume of biomass can be converted to an enormous amount of energy and may materials. Equivalent to approximately 50 billion tons of oil [2], agricultural biomass waste converted to energy could substantially offset fossil facts, reduce greenhouse gases emissions and provide sources for renewable energy.

Residues from exops take the form of residual stalks, strow, leaves, roots, basic, aut or seed shells, waste



## **ACCESO ABIERTO A LAS PUBLICACIONES**



Open Access pathways permitted by this journal's policy are listed below by article version. Click on a pathway for a more detailed view.

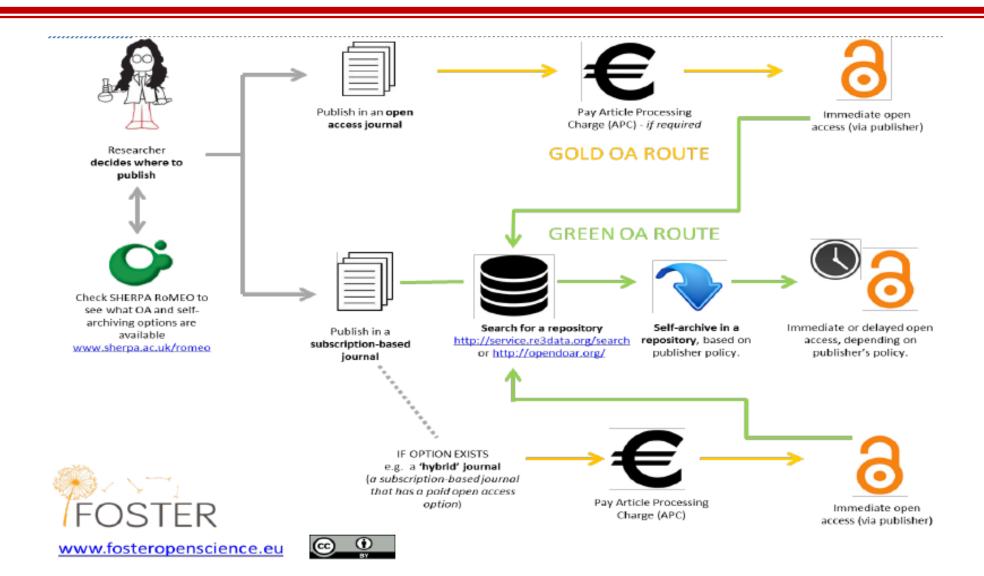
Sherpa Romeo

About Search TJ List Statistics

Published Version	£  ■  X None	Política institucional de acceso abierto de la Universitat Politècnica de València (julio 2011)	
Accepted Version [pathway a]	➤ None 🐔 🗁 Author's Homepage	<ul> <li>Asegurar la máxima visibilidad de los resultados de investigación, innovación, formación y transferencia de la UPV.</li> <li>La UPV solicita a sus miembros que</li> </ul>	
Accepted Version [pathway b]	<ul><li>Institutional Repository, Funder Designated Location</li></ul>	depositen en RiuNet sus publicaciones académicas y científicas.  • La UPV recomienda a sus autores que	
Submitted Version	■ None <b>¥</b> ≡ Preprint Repository, Author's Homepage	publiquen los resultados de sus investigaciones en revistas con políticas de acceso abierto que permitan el depósito de los trabajos en repositorios abiertos.	



### **ACCESO ABIERTO A LAS PUBLICACIONES**





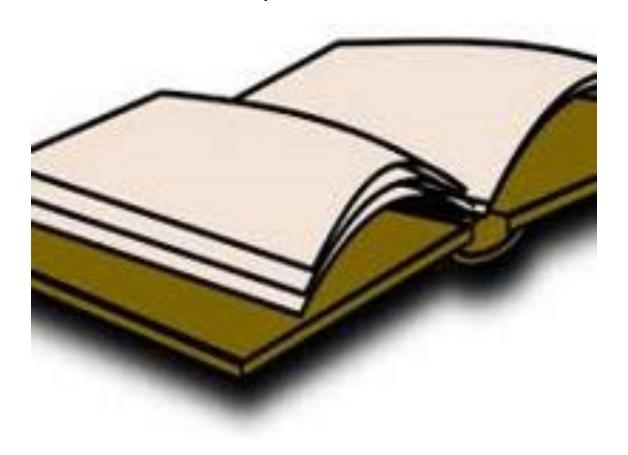
## LAS DOS VÍAS DEL ACCESO ABIERTO

Acceso Abierto Verde	Repositorios	Depósito gratuito	(c) Editor Raramente CC
Acceso Abierto Dorado	Revistas de Acceso Abierto	Pago de APC Publicación gratuita	(c) Autores CC BY-xxx
Acceso Apierto Dorado	Opción Híbrida	Pago de APC	(c) Autores CC BY-xxx



## ¿DATOS?

Bajo el paraguas conceptual de la ciencia abierta se está promoviendo *el mismo horizonte* para datos y publicaciones.





El valor de los datos reside en su uso.



## Datos de investigación

- Material factual registrado.
- Aceptado en la comunidad científica.
- Necesario para documentar, respaldar y validar la investigación.

### Características:

- Fuente única.
- Difíciles de obtener de nuevo.
- Datos finales.

✓ Investigación cuantitativa y cualitativa.



## QUÉ ES UN DATASET.

Proceso de obtención

Experimentales

Simulaciones

Observaciones



Objetivo de su recogida

Específico

De alcance medio

De interés general

Tratamiento recibido

Raw data

Final research data (procesados, derivados).

Los más valiosos

Los que piden





## **QUÉ INCLUYE**

Sí: cuestionarios, cintas de audio, vídeos, desarrollo de modelos, fotografías, diseños y muestras, software, mediciones tomadas por máquinas, colecciones documentales, entrevistas, encuestas, observaciones de campo, videograbaciones, simulaciones computacionales,

= investigación cuantitativa y cualitativa.

NO: notas de laboratorio, conjuntos parciales de datos, análisis preliminares, borradores, planes, reportes, objetos, informes en Pdf,

#### Observational

- Captured in situ (real-time)
- Cannot be recaptured, recreated, or replaced (unique data).
- Examples include sensor readings, telemetry, survey results, images, and human observation.

#### Experimental

- Data from lab equipment and under controlled conditions
- Often reproducible, but can be expensive to do so
- Examples include gene sequences, chromatograms, magnetic field readings, and spectroscopy.

#### Simulation

- Data generated from test models studying actual or theoretical systems
- Models and metadata where the input more important than the output data
- Examples include climate models, economic models, and systems engineering.

#### Derived or compiled

- . The results of data analysis, or aggregated from multiple sources
- Reproducible (but very expensive)
- Examples include text and data mining, compiled database, systems engineering, and 3D models

#### Reference or canonical

- · Fixed or organic collection datasets, usually peer-reviewed, and often published and curated
- Examples include gene sequence databanks, census data, chemical structures.

## ☐ Publication (1733700) +

☐ Image (793267) +

Type

- Dataset (185993)
- ☐ Software (98437)
- ☐ Presentation (38385)
- Other (18518)
- ☐ Poster (16004)
- ☐ Video (7704)
- ☐ Lesson (4527)
- ☐ Physicalobject (1008)



### **EN SENTIDO AMPLIO**

"documentos en formato digital, distintos de las publicaciones científicas, recopilados o elaborados en el transcurso de actividades de investigación científica y utilizados como prueba en el proceso de investigación, o comúnmente aceptados en la comunidad investigadora como necesarios para validar las conclusiones y los resultados de la investigación."

Datos procesados de interés general.

DIRECTIVA (UE) 2019/1024 DEL PARLAMENTO EUROPEO Y DEL CONSEJO de 20 de junio de 2019 relativa a los datos abiertos y la reutilización de la información del sector público . Art. 2.9.



### Datos ABIERTOS.

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

"Tan abierto como sea posible, tan cerrado como sea necesario." (<u>H2020 Programme</u> Guidelines on FAIR Data Management in Horizon 2020)



No se difundirán en abierto los datos de investigación en determinadas situaciones: existencia de cláusulas de confidencialidad, posible explotación comercial o industrial, datos personales sin tratar o sin consentimiento, etc.



### **Resumen Datos**

Contenida o presentada de diversas maneras

Cifras, tablas, imágenes (fijas o en movimiento), grabaciones de audio o visuales.

Recopilados, generados u obtenidos en el transcurso o como resultado de una investigación

Realización de experimentos de campo o de laboratorio, de ensayos, de encuestas, entrevistas, grupos de discusión o análisis de datos.

Que el investigador utiliza posteriormente como base de datos para...

hacer cálculos o sacar conclusiones para desarrollar, apoyar o revisar teorías, prácticas y resultados.

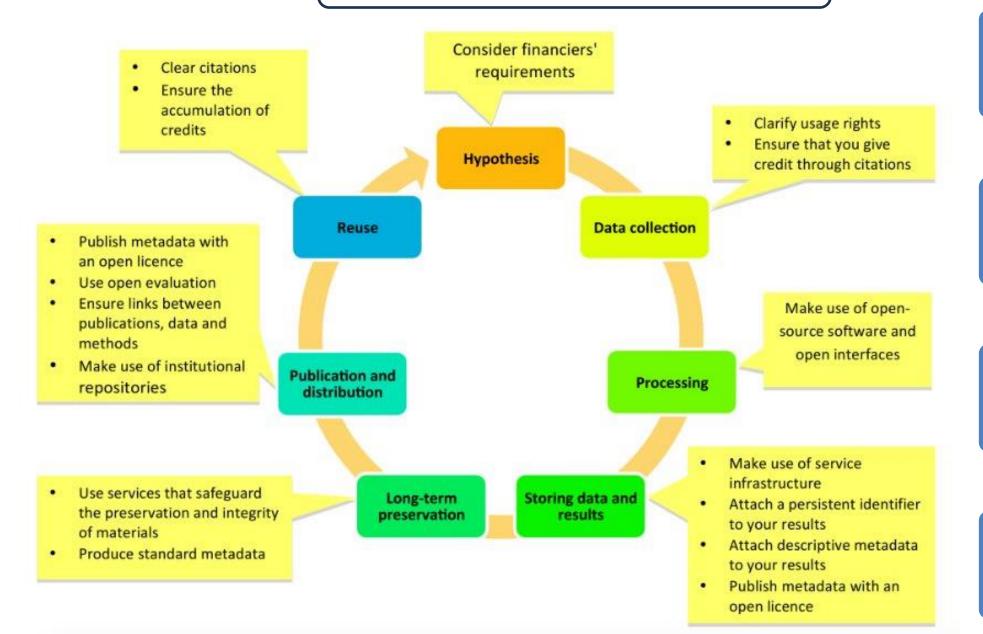
"Repositorio/s"

Que la Universidad posee.

https://library.soton.ac.uk/researchdata/phd



### CÓMO FUNCIONA LA INVESTIGACIÓN



Conseguir los fondos: MANDATOS, LEGISLACIÓN.



Escribir el PLAN GENERAL DE DATOS.



Recoger los datos de la investigación: FORMATOS, ALMACENAMIENTO, PRINCIPIOS FAIR, etc..



Finalizar la investigación: publicar los resultados, depositar los datos: REVISTAS, REPOSITORIO, LICENCIAS.



### CICLO DE LOS DATOS.



ANTES de la investigación: 1º versión del PGD: qué tipo de datos vamos a recopilar y cómo.

DURANTE la investigación: recopilación, descripción, almacenamiento y publicación.

DESPUÉS de la investigación: preservación y difusión.

https://zenodo.org/records/4057867



## NIH issues a seismic mandate: share data publicly

The data-sharing policy could set a global standard for biomedical research, scientists say, but they have questions about logistics and equity.

By Max Kazlov









The US National Institutes of Health, located in Bethesda, Maryland, is the world's largest funder of biomedical research. Credit: Grandbrothers/Alamy

In January 2023, the US National Institutes of Health (NIH) will begin requiring most of the 300,000 researchers and 2,500 institutions it funds annually to include a data-management plan in their grant applications – and to eventually make their data publicly available.

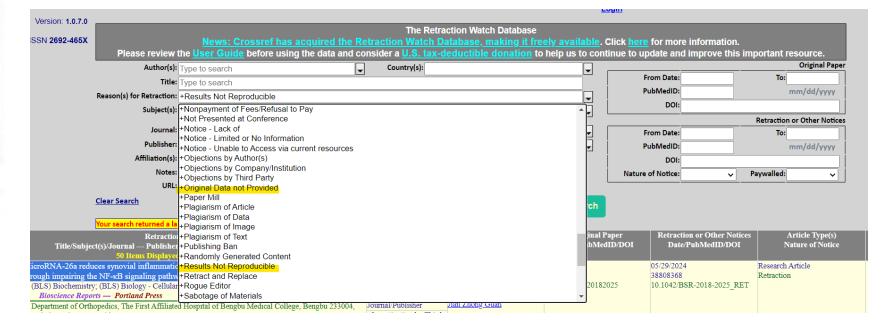
Researchers who spoke to Nature largely applaud the open-science principles underlying the policy—and the global example it sets. But some have concerns about the logistical challenges that researchers and their institutions will face in complying with it. Namely, they worry that the policy might exacerbate existing inequities in the science-funding landscape and could be a burden for early-career scientists, who do the lion's share of data collection and are already stretched thin.

https://www.nature.com/articles/d41586-022-00402-1

### **EL PROBLEMA**

Crisis de reproducibilidad de la investigación científica.

Investigación irreproducible en Estados Unidos: entre 10.000 y 50.000 millones en estudios que utilizan métodos deficiente s.





## Reproducibilidad de una publicación científica

Reutilización: para futuras investigaciones.

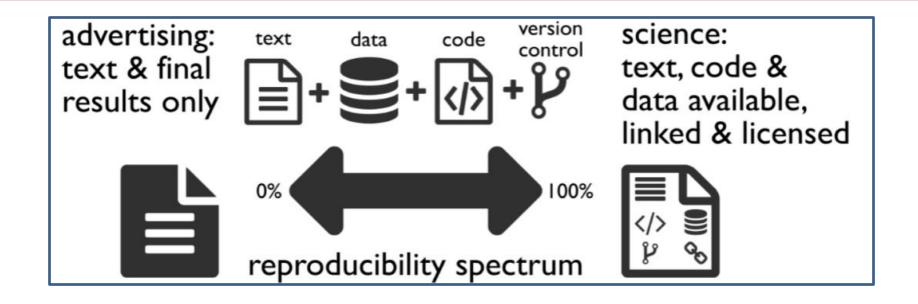
Reproducibilidad: para verificar la investigación.

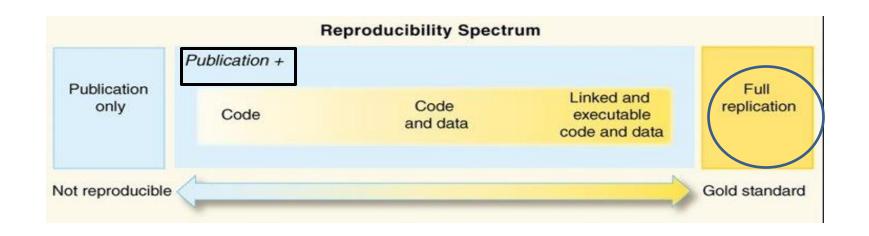


¿Duplicar esfuerzos? ¿Resultados sin datos? ¿Grandes proyectos sin datos?



## Reproducibilidad de una publicación científica





See this research in an executable code view.

Research Article

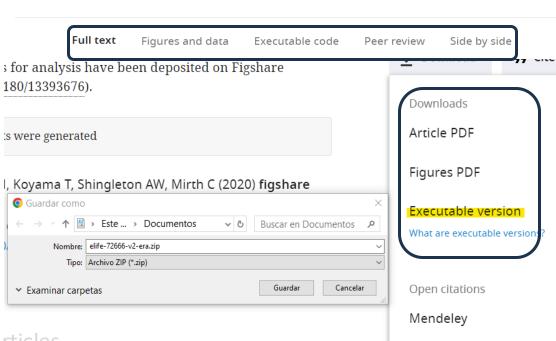
Developmental Biology, Evolutionary Biology

### Ecdysone coordinates plastic growth with robust pattern in the developing wing

André Noqueira Alves, Marisa Mateus Oliveira, Takashi Koyama, Alexander Shingleton , Christen Kerry Mirth <sup>™</sup>

Instituto Gulbenkian de Ciência, Portugal; School of Biological Sciences, Monash University, Australia; Department of Biology, University of Copenhagen, Denmark; Department of Biological Sciences, University of Illinois at Chicago, United States

Mar 9, 2022 · https://doi.org/10.7554/eLife.72666 8 @





### Steps to enrich your published article with code and data



https://elifesciences.org/collections/d72819a9/executable-research-articles

tabular data files

to Stencila Hub

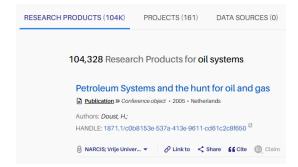


## ¿Qué podemos hacer?



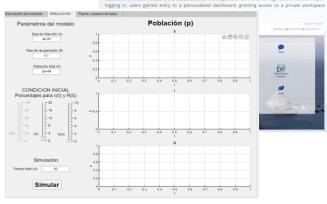
Publicar los nuevos resultados

Descubrir RESULTADOS en un agregador



### VIRTUAL RESEARCH ENVIRONMENT

Starting in 2020, the SeaDataNet infrastructure integrated a Virtual Research Environment (VRE) as part of the SeaDataCloud project. This initiative pioneered a pilot application that provided diverse services and research support utilising marine data and SeaDataNet tools such as WebDDV and DMVs within the cloud environment. Access to the SeaDataCloud VRE required a Marine-ID for login authentication. Upon



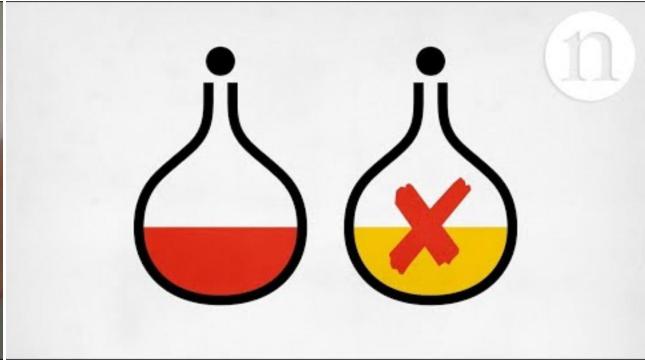
♠ > Software → > VRE

Reproducir los resultados en un entorno virtual (VRE) Identificar DATOS asociados a los artículos en repositorios FAIR así como SOFTWARE









La replicabilidad de resultados

<u>Is there a reproducibility crisis in science?</u>



## PRÁCTICA: TEST DE REPRODUCIBILIDAD

# Reproducibility Self Assessment

Reproducibility is the cornerstone of scientific advancement.

Take our quick self assessment to see how your lab measures up, and learn how to apply best practices to ensure your work remains replicable, both for you and your readers.

START ASSESSMENT

Based on the resources developed by protocols io and Reproducibility for Everyone.





- 1. Entrad en <a href="https://plos.org/reproducibility-assessment/">https://plos.org/reproducibility-assessment/</a>
- 2. Cumplimentad los datos con vuestro último proyecto (o simuladlo)

Tiempo: 10 minutos.

https://plos.org/reproducibility-assessment/



## ¿CÓMO LO HACEMOS?

# GESTIÓN DE DATOS DE INVESTIGACIÓN= actividades asociadas con los datos que no incluyan su uso directo:

organización, copias de seguridad, archivo de datos, distribución, publicación, confidencialidad, seguridad, sincronización, versionado, accesos, licencias, permisos, confidencialidad, tratamiento de datos personales ...





## ¿Para qué?

Integridad, reproducibilidad, recuperación, acceso...

- Mantiene la integridad de la investigación.
- Permite que otros puedan hacer uso de los datos (reutilización)
- Reduce el riesgo de pérdida de datos (+seguridad y preservación).
- Asegurar el acceso continuo a los datos, aumentando el impacto y la visibilidad de nuestras investigaciones.
- Área emergente: la comunidad necesita infraestructura.
- Necesidad de grandes cantidades de datos.
- Algunos datos son únicos e irremplazables.







## ¿CÓMO hay que desarrollarla?

## 1. INFRAESTRUCTURA institucional.













2. ROL de datos en el grupo de investigación (*EJE B1 en la ENCA*).

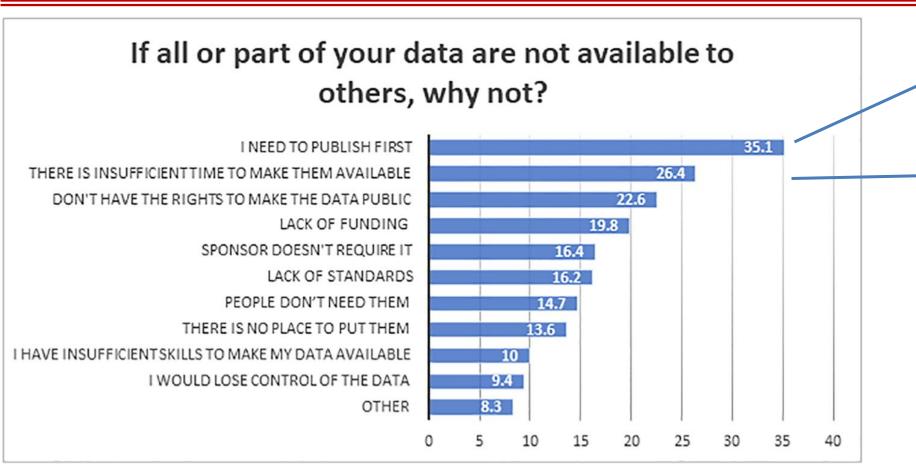
Tiene que haber dentro del grupo de investigación una persona encargada de la Gestión de datos de investigación.

- 1. Desarrollar un PLAN GENERAL DE DATOS.
- 2. Proveer acceso abierto a los datos de investigación SI ES POSIBLE.
- 3. Hacer los datos FAIR.





### ¿PROBLEMAS?



**Fuente**: Tenopir, C. et al. (2020). Data sharing, management, use, and reuse: Practices and perceptions of scientists worldwide. PLOS ONE, 15(3).

https://doi.org/10.1371/journal.pone.0229003

Esto es lo que va a cambiar: se publicará el paper y los datos.

Esta es la tarea del encargado de la gestión de los datos y del PGD.

cosc
 eosc
 eos

EOSC Multi-Annual Roadmap (MAR) 2025 and 2026–2027

Advisory note

The Multi-Annual Roadmap defines a set of priorities for future investment in EOSC. Within this MAR we have highlighted activities for 2025 as it is expected that two Work Programmes will be developed - one for 2025 and one for the 2026-2027 period.

The priorities under each objective are not placed in order of value. They are thematically clustered on top. They have been labelled purely for ease of referencing in the consultation

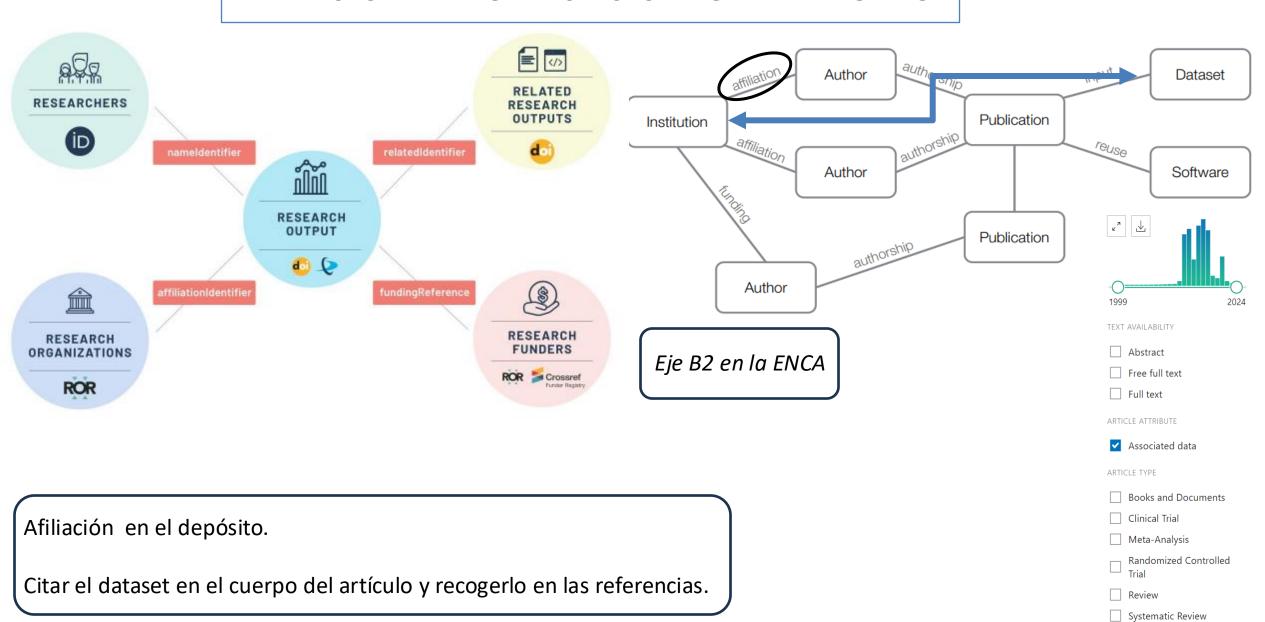
- Objective 1 Ensure that Open Science practices and skills are rewarded and taught, becoming the 'new normal'
- 1.1 European level prioritie
- A. Align on core curricula for data stewardship, Open Science and FAIR for both support and researchers. (2025)
- B. Support the development of networks for data stewards, research software eggd semantic artefact curators and Open Science communities, leveraging on edistinational networks and coordinating these at European level. (2025)
- C. Collect ongoing national best practices and case studies for the research of career assessment of open science practices in collaboration with the Coalition for Advancin Research Assessment (CoARA). (2025)
- D. Support and complement the CoARA movement with developing tools and processes for the research and career assessment of open science practices.
- E. Pilot indicators/metrics at the European level for the research and career assessment of open science practices and disseminate them Europe-wide.
- F. Support and embed appraisal and preservation, and monitor their value and costs. Practices should be transparent, aligned (i.e. differences for national or disciplinary reasons are justified) and interoperable (i.e. through machine actionable rights management).
- G. A dedicated provision for supporting the EOSC Partnership in its concertation actions. This covers stakeholder engagement, coordination, governance and monitoring. A cascading grant mechanism should support widening activities with under-represented cluntries and national coingritence centres, amongst other things.

/files/2023-01/MAR\_2025-

27\_draft.pdf



### RELACIONAR PUBLICACIONES Y DATASETS.





## LINKING DATA: cita y enlace.

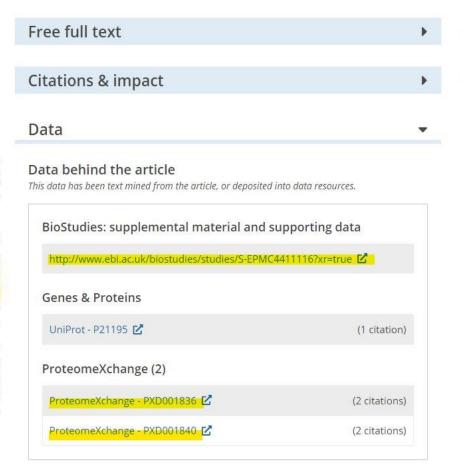
El dataset adquiere valor y visibilidad si se integra con una publicación relacionada (metadato enlazable).

El código identificador es <a href="http://doi.org/10.5281/zenodo.4434086">http://doi.org/10.5281/zenodo.4434086</a> y se hallan disponibles para su uso y distribución bajo una licencia Creative Commons Attribution 4.0 International.

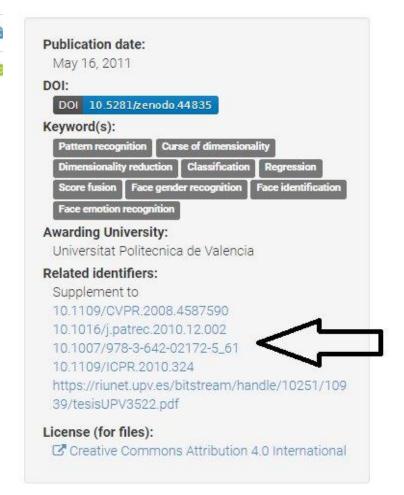
Los conjuntos de datos surgidos y recopilados durante la realización del estudio se encuentran disponibles en el repositorio internacional de datos de investigación Zenodo, en la siguiente dirección: https://zenodo.org/record/4434086#.YhaKWsa-CHp5 Los dataset publicados proporcionan información sobre las iniciativas en materia de alfabetización en datos propuestas por las bibliotecas universitarias-CRAI en España, durante el curso académico 2020-2021. Asimismo se facilita la url de acceso a los portales de alfabetización informacional de las bibliotecas universitarias españolas.

#### Cita del dataset:

Martín González, Yolanda; y Iglesias Rodríguez, Ana. (2021). Alfabetización en datos en bibliotecas-CRAI españolas curso académico 2020-2021 [Data set]. Zenodo. DOI: http://doi.org/10.5281/zenodo.4434086



Data that cites the article





### Metadatos En SENIA:



Metadatos enlazables En Riunet:



### Investigación

Documentos generados por los docentes e investigadores de la Universidad en su labor de investigación científica producida o editada por los departamentos y centros de la Universitat Politècnica de València.

Riunet tiene como objetivos el almacenamiento, la gestión, la preservación y la difusión de estos documentos, y pretende contribuir al aumento de la visibilidad del trabajo desarrollado por los miembros de la comunidad universitaria.



#### Subcomunidades en esta comunidad







Informes técnicos y otro material de investigación [171]



Material investigación. Editorial UPV [16902]



Revistas UPV. Divulgación [539]

### Colecciones en esta comunidad

Artículos, conferencias, monografías [46261]

Colección especial COVID-19 [250]

### Datasets [77]

OpenAIRE (Open Access Infrastructure for Research in Europe) [6045]

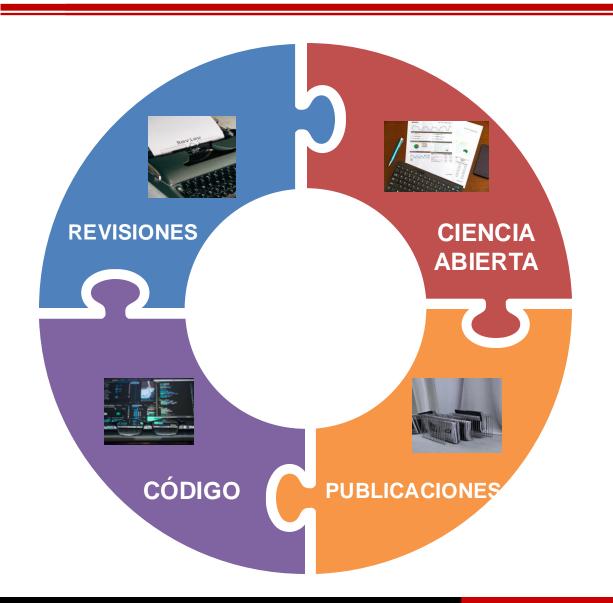
Software [19]

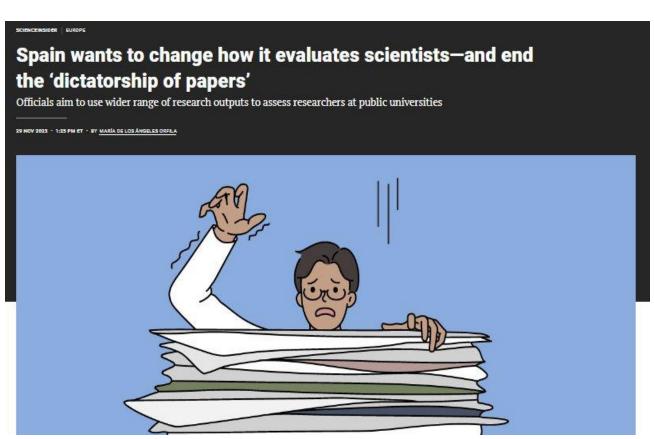
Tesis doctorales [5219]





### NO SOLO PUBLICACIONES





Valverde-Berrocoso, J. (2021). Introducción a la ciencia abierta. REUNID. <a href="https://doi.org/10.5281/zenodo.5747674">https://doi.org/10.5281/zenodo.5747674</a>



## PRÁCTICAS EN CIENCIA ABIERTA



## Horizon Europe (HORIZON)

### HE Programme Guide

https://ec.europa.eu/info/fundingtenders/opportunities/docs/2021-2027/horizon/guidance/programmeguide horizon en.pdf

## **OBLIGATORIAS**

OA inmediato a publicaciones revisadas por pares.

OA a datos (si es posible).

Alineación con los principios FAIR.

Acceso a los resultados necesarios para validar las conclusiones de las publicaciones. Documentación.

## RECOMENDADAS

Intercambio temprano de resultados, <u>pre prints</u>.

Open peer review.

<u>Pre-registros</u> del plan de investigación.

Gestión de otros tipos de resultados de investigación (<u>software</u>).

Depósito de PGD.



# Open is not enough







**UPV** 

BIBLIOTECA Y DOCUMENTACIÓN CIENTÍFICA

Universitat Politècnica de València

www.upv.es · rienma@upvnet.upv.es



# HORIZONTE EUROPA.





**Horizon Europe** 

**Data Management Plan Template** 

Version 1.0

mandate for Research Data Management.

Guides for Researchers. How to comply with Horizon Europe

2008

FP7
OA Pilot

Deposit and open access

2014

H2020

**OA Mandatory** 

Deposit and open access

& ORD/DMP Pilot

2017

H2020

**OA Mandatory** 

Deposit and open access

& ORD/DMP by default (exceptions)

2021

**Horizon Europe** 

OA **Mandatory**Deposit and open

access

DMP in line with FAIR **Mandatory** 

OD by default (exceptions)

& Open Science embedded

LOS DATOS DE INVESTIGACIÓN ALIMENTARÁN LA EUROPEAN OPEN SCIENCE CLOUD (EOSC)

Guía del participante.



# HORIZONTE EUROPA.

Preparación de un Plan de Gestión de Datos (DMP) a los 6 meses y actualización a lo largo del proyecto Depósito de datos de investigación en un repositorio de confianza:

-Repositorios certificados

-Repositorios institucionales/generalistas/temáticos: asignación de identificadores persistentes (DOI, handle..), integridad datos, derechos de acceso, licencias de uso, preservación

-Google drive, Dropbox, webs personales etc NO son repositorios

Art. 17 Dissemination (Grant agreement, Annex 5): Gestión de datos de investigación en línea con Principios FAIR

Los datos de investigación son "as open as posible, as closed as necessary"

### **FAIR NO ES IGUAL A OPEN**

"ACCESSIBLE" hace referencia a dónde y bajo qué condiciones se encuentran alojados los datos

Si los datos no pueden accederse (closed) hay que justificarlo en el DMP

Provisión de información sobre otros resultados/herramientas necesarias para validar o reutilizar los datos de investigación



# Resumen requisitos del Horizon Europe vs open science

### **Publicaciones**

- Depósito de los trabajos (versión publicada o versión aceptada)
- Licencia: Creative Commons BY o similar
- · Cuando: lo antes posible, max en el momento de su publicación
- · Dónde: Repositorio (libre elección)

### **Datos**

- Datos (tan abiertos como sea posible y tan cerrados como sea necesario)
- · Depósito en un repositorio seguro (trusted) de acuerdo al plan de gestión de datos
- Licencia Creative Commons BY o Creative Commons Public Domain Dedication (CCO)
- · Plan de gestión de datos (entregable)
- · Depósito tan pronto como sea posible
- Licencia Creative Commons BY o Creative Commons Public Domain Dedication (CCO)

- Gestionar los datos bajo principios FAIR.
- ► Hacer y mantener un PGD.
- Depositar lo antes posible en un repositorio.
- Licencias:
  - Bajo licencia CC-BY o CCO, siguiendo la máxima de tan abierto como se pueda, tan cerrado como sea necesario.
  - Depositar los metadatos bajo una licencia CCO
- Proporcionar al repositorio cualquier información relevante para la validación o reutilización de los datos.

In Horizon Europe, beneficiaries must manage the digital research data generated in the action ('data') responsibly, in line with the <u>FAIR principles</u>, and should at least do the following:

- Prepare a Data Management Plan (DMP) and keep it updated throughout the course of the project
- Deposit data in a trusted repository and provide open access to it ('as open as possible, as closed as necessary')
- Provide information (via the same repository) about any research output or any other tools and instruments needed to re-use or validate the data



# INFRAESTRUCTURAS.



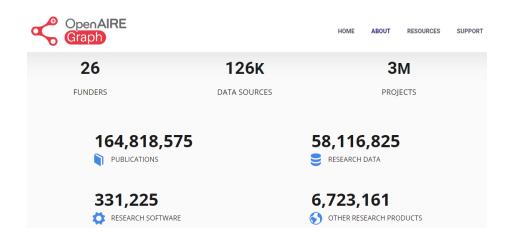








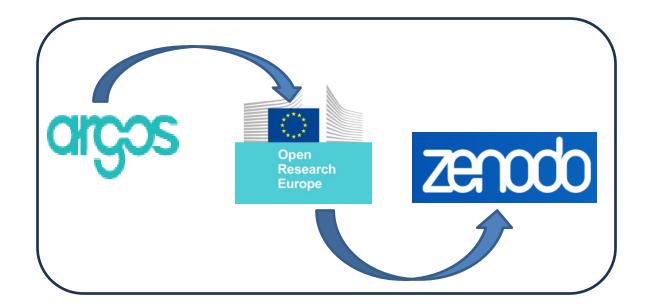
Data shared and preserved across borders and disciplines





# OPEN RESEARCH EUROPE.

H 2020 H EUROPA

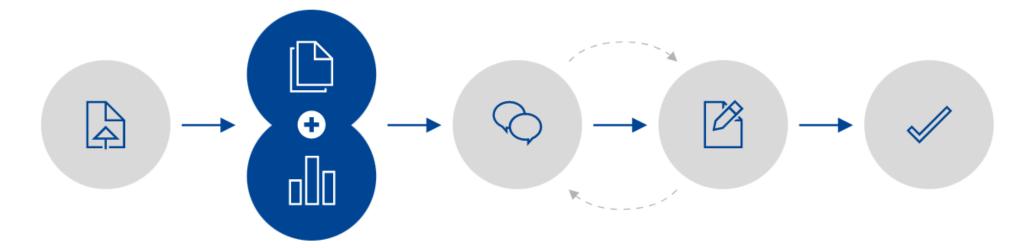


- -Publicación en acceso abierto de preprints que son resultado de proyectos H2020 y Horizonte Europa -Acepta trabajos originales inéditos y que no están en proceso de revisión en otras plataformas/revistas
- -Una vez que se depositan en ORE los preprints no pueden enviarse para publicación a una revista
- -Los preprints pasan por revisiones por pares abiertas
- -Los autores no deben pagar nada
- -Comentarios públicos postpublicación
- -Para todas las áreas de investigación
- -Métricas: estadísticas de uso, citas en Google scholar, altmétricas
- -Indexación por Google Scholar, Pubmed, envío a Zenodo



# **OPEN RESEARCH EUROPE.**





### **Article Submission**

Submission is via a single-page submission system. The in-house editorial team carries out a comprehensive set of prepublication checks to ensure that all policies and ethical guidelines are adhered to.

### **Publication & Data Deposition**

Once the article has passed the prepublication checks, a fully typeset version is published with a DOI, enabling immediate viewing and citation, as well as indexing in Google Scholar. Once the article is published, it cannot be sent to another journal for

# Open Peer Review & Article Revision

Expert reviewers are selected and invited, and their reviews and names are published alongside the article, together with the authors' responses and comments from registered users.

### Send to Indexers & Repositories

Authors are encouraged to publish revised versions of their article. All versions of an article are linked and independently citable. Articles that pass peer review are indexed in external databases such as PubMed, Scopus, and Google Scholar.



# OPEN RESEARCH EUROPE. PUBLICACIÓN.

# What is required when submitting an article

- Your dataset(s) must be deposited in an appropriate data repository.
- Your dataset(s) must have a license applied which allows reuse by others (CC0 or CC-BY).
- Your dataset(s) must have a persistent identifier (e.g. a DOI), allocated by a data repository.
- You must provide a data availability statement as a section at the end of your article, including elements 1-3.
- You must include a data citation and add a reference to data to your reference list.
- Your dataset(s) should not contain any sensitive information, for example in relation to human research participants.
- You should share any related software and code.
- Your dataset(s) must be useful and reusable by others, adhere to any relevant data sharing standards in your discipline and align with the FAIR Data Principles.
- Your dataset(s) should link back to your article, if possible.



# OPEN RESEARCH EUROPE.

**Open Research Europe** 4 Steps to Open Data Prepare your data Link your for sharing datasets to Add a Data your article Availability Statement to your article Select a repository

- 1. Preparar datos.
- 2. Seleccionar repositorio. Depositar.
- 3. Añadir la disponibilidad de los datos en el artículo.
- 4. Enlazar dataset y artículo.

Para cumplir el requisito de simultaneidad y poner la declaración de disponibilidad de datos en el artículo: **se depositan los datos primero**.

"¿Pero cómo ingreso el DOI de los datos en la publicación?"

open-research-europe.ec.europa.eu/





# Data availibility statement. Declaración de disponibilidad de datos

### Indicar:

- REPOSITORIO
- LICENCIA
- IDENTIFICADOR PERSISTENTE

MENU V

The role of miniaturization in the evolution of the mammalian jaw and middle ear

# Data availability

All relevant data (three-dimensional osteological, finite element analysis and multibody dynamics analysis models and computer code) are available via the DataBris repository of the University of Bristol (https://doi.org/10.5523/bris.n5f4ogftag0r2fbffh8u7waok).

El código identificador es <a href="http://doi.org/10.5281/zenodo.4434086">http://doi.org/10.5281/zenodo.4434086</a> y se hallan disponibles para su uso y distribución bajo una licencia Creative Commons Attribution 4.0 International.

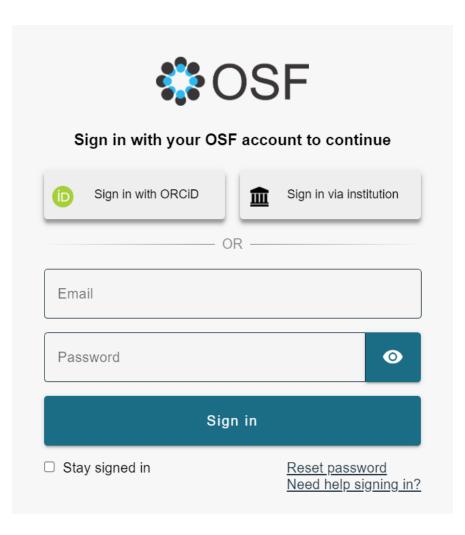
Los conjuntos de datos surgidos y recopilados durante la realización del estudio se encuentran disponibles en el repositorio internacional de datos de investigación Zenodo, en la siguiente dirección: https://zenodo.org/record/4434086#.YhaKWsa-CHp5 Los dataset publicados proporcionan información sobre las iniciativas en materia de alfabetización en datos propuestas por las bibliotecas universitarias-CRAI en España, durante el curso académico 2020-2021. Asimismo se facilita la url de acceso a los portales de alfabetización informacional de las bibliotecas universitarias españolas.

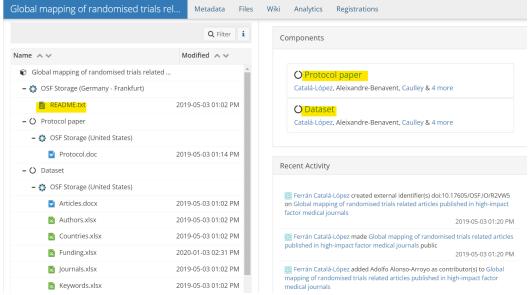
### Cita del dataset:

Martín González, Yolanda; y Iglesias Rodríguez, Ana. (2021). Alfabetización en datos en bibliotecas-CRAI españolas curso académico 2020-2021 [Data set]. Zenodo. DOI: http://doi.org/10.5281/zenodo.4434086



# **Open Science Framework OSF HOME.**





https://osf.io/r2vw5/

- Plataforma de servicios modulares.
   Almacenamiento, comunicación, publicación, gestión, visualización.
- 2. Trabajo de forma privada o pública.
- 3. Facilita un DOI.
- 4. Integrado con otras herramientas.

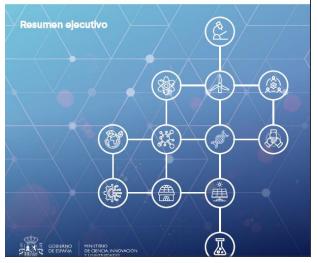




# PEICTI 2024-2027.

# Plan Estatal de Investigación Científica y Técnica y de Innovación





- Apoyar el libre acceso y gestión de los datos generados por la investiga ción, de acuerdo con los principios internacionales FAIR, implantando la *obligatoriedad de realizar un plan de gestión de datos (PGD)* como parte integral de los proyectos de investigación financiados con fondos públicos.
- Considerar la *contribución a la ciencia abierta como criterio evaluable* en la actividad de investigadores/as.
- Perfiles profesionales, infraestructuras digitales.



Ley 17/2022 por la que se modifica la Ley 14/2011, de la Ciencia.

Ley 17/2022	DEPÓSITO
¿Qué?	Versión aceptada + datos
¿Cuándo?	En el momento de la publicación
¿Dónde?	Repositorios institucionales o temáticos de acceso abierto

https://www.boe.es/buscar/act.php?id=BOE-A-2022-14581

«Artículo 37. Ciencia abierta.

1. Los agentes públicos del Sistema Español de Ciencia, Tecnología e Innovación impulsarán que se haga difusión de los resultados de la actividad científica, tecnológica y de innovación, y que los resultados de la investigación, incluidas las publicaciones científicas, datos, códigos y metodologías, estén disponibles en acceso abierto. El acceso gratuito y libre a los resultados se fomentará mediante el desarrollo de repositorios institucionales o temáticos de acceso abierto, propios o compartidos.

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 los datos asociados a las mismas en repositorios institucionales o temáticos de acceso abierto, de forma simultánea a la fecha de publicación.



L.O. 2/2023 del Sistema Universitario (LOSU).

Ley 2/2023	DEPÓSITO	
¿Qué?	Versión aceptada + datos	
¿Cuándo?	En el momento de la publicación	
¿Dónde?	Repositorios institucionales o temáticos de acceso abierto	
¿Cómo?	FAIR	

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

- 2. El personal docente e investigador deberá depositar una copia de la versión final aceptada para publicación y los datos asociados a la misma en repositorios institucionales o temáticos de acceso abierto, *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.
- 4. Los Ministerios ... promoverán otras iniciativas orientadas a facilitar el libre acceso a los datos generados por la investigación (datos abiertos) y a desarrollar infraestructuras y plataformas abiertas.
- 5. Los datos, entendidos como aquellas fuentes primarias necesarias para validar los resultados de las investigaciones, deberán seguir los *principios FAIR* (datos fáciles de encontrar, accesibles, interoperables y reutilizables) y, siempre que sea posible, difundirse en acceso abierto.

https://www.boe.es/buscar/act.php?id=BOE-A-2023-7500&p=20230323&tn=1#a1-4



«Artículo 69. Acreditación de los cuerpos docentes universitarios.

L.O. 2/2023 del Sistema Universitario (LOSU). ▶ Una evaluación basada en la especificidad del área o ámbito de conocimiento, teniendo en cuenta, entre otros criterios, la experiencia profesional, en especial, cuando se trate, entre otras, de profesiones reguladas del ámbito sanitario, la relevancia local, el pluralismo lingüístico y el acceso abierto a datos y publicaciones científicas.

R.D. 678/2023, acreditación estatal para el acceso a los cuerpos docentes universitarios.

«ANEXO. Evaluación del mérito.

Se valorará una amplia gama de resultados, en función de su calidad, grado de difusión e impacto científico y social. Entre ellos, se considerarán publicaciones científicas, datos, códigos, "software", creaciones artísticas, patentes y registros de propiedad intelectual, aportaciones a congresos, divulgación científica, transferencia al sector productivo, transferencia social e intercambio del conocimiento. Se considerarán las evaluaciones positivas de tramos de investigación y de transferencia. Se valorará la disponibilidad en acceso abierto a todos los resultados de la actividad desarrollada y el uso de repositorios institucionales y temáticos, e infraestructuras y plataformas abiertas.



# ANECA. Sexenios de investigación.



### Baremo general de aplicación en la evaluación de los sexenios de investigación 2023 por parte de los Comités Asesores

Contribución al progreso del conocimiento en el área	10%
Impacto científico	60%
Impacto social	10%
Contribución a la ciencia abierta	10%
Aportación preferente	10%
Reiteración de publicaciones	-15%
Autoría insuficientemente justificada	-15%
Malas prácticas de publicación	-25% a -100%

### Contribuye al progreso del conocimiento en el área (10%)

- Grado de originalidad e innovación
- 2. Prioridad temática o reto abordado y aportación al debate científico
- 3. Aportación metodológica

### Presenta impacto científico (60%)

- 1. Uso y lectura de la aportación (en plataforma de revista/editorial/congreso, etc.)
- 2. Citas, excluyendo autocitas / Citas normalizadas / Posición en citas
- 3. Calidad editorial (publicaciones) o reconocimiento (otras aportaciones) del medio de difusión
- 4. Internacionalización del medio de difusión
- Impacto científico del medio de difusión (editorial, revista, congreso, etc.)

Otros indicios: Proyectos financiados vinculados, tesis doctorales vinculadas, estándares de ética, premios, traducciones, reseñas, etc.

### Presenta impacto social (10%)

- 1. Uso y lectura de la aportación (en otras plataformas no académicas)
- Menciones no académicas o sociales a la aportación (working papers, informes, blogs científicos o profesionales, guías para públicos amplios, etc.)
- 3. Interacciones en medios o plataformas sociales sobre la aportación

Otros indicios: Aportación a políticas públicas, soluciones a problemas sociales, etc.

### Contribución a la ciencia abierta (10%)

- Grado de apertura (depósito de preprint, ruta diamante, licencias Creative Commons, etc.)
- En caso de conjunto de datos: Adopción de principios FAIR, posibilidad de descarga, etc.
- En caso de software: Contribución a software libre

Otros indicios: Otras contribuciones no académicas, diversidad de repositorios utilizados para depósito, uso y reutilización de datos, etc.

### Aportación preferente (10%)

Ajuste a lo señalado como 'aportación preferente' en el apartado 1 de cada Campo/Subcampo en la Resolución de criterios de evaluación de 2023: artículo, libro, capítulo de libro, patente, contribución a congreso, etc.

### Doeibles negalizaciones

- Reiteración de publicación en las mismas revistas o editoriales, incluyendo reiteración en special issues y entendiendo por reiteración 3 o más de las 5 aportaciones, siempre que no quede suficientemente justificado (-15% en cada aportación afectada).
- En caso de autoría / coautoría / autoría múltiple que no quede suficientemente justificada (-15% en cada aportación afectada).
- En caso de malas prácticas de publicación (reiteración de contenidos, traducción de obra propia, plagio, alteración de datos, artículos retractados, etc.) (de -25% a -100% en cada aportación afectada).

«Contribución a la ciencia abierta

- Grado de apertura (depósito de preprint, ruta diamante, licencias Creative Commons, etc.).
- 2. En caso de conjunto de datos: Adopción de principios FAIR , posibilidad de descarga, etc.
- 3. En caso de software: Contribución a software libre.

### III. OTRAS DISPOSICIONES

# MINISTERIO DE CIENCIA, INNOVACIÓN Y UNIVERSIDADES

25537 Resolución de 5 de diciembre de 2023, de la Comisión Nacional Evaluadora de la Actividad Investigadora, por la que se publican los criterios para la evaluación de la actividad investigadora.

Sección I. Criterios generales para todos los campos.

Además, las mencionadas órdenes indican que las aportaciones se valorarán considerando la situación general de la ciencia en España, <u>lo que aconseja estimar que en el momento actual también pueden justificar actividad investigadora los conjuntos de datos</u>, las metodologías, los programas de ordenador o los modelos de aprendizaje automático.



# Estrategia Nacional de Ciencia Abierta (2023-2027).



- ▶ Datos, protocolos y metodología abiertos: poner a disposición de la comunidad de investigación nuevas fuentes de información, permitir la reproducibilidad de los análisis, mejorar la comprensión y alcance de los resultados y contribuir a acelerar los descubrimientos a través de la reutilización de los datos.
- Gestión de datos de investigación según los <u>principios FAIR</u> es un eje estratégico.

B1. <u>Perfiles profesionales de apoyo</u> a la gestión de datos de investigación .

CÓMO: Se financiará, dentro de las convocatorias de personal de apoyo a la investigación, la creación de perfiles profesionales de apoyo a la gestión FAIR de datos de investigación tales como data stewards y research data analysts, entre otros.



https://www.ciencia.gob.es/InfoGeneralPortal/docume nto/c30b29d7-abac-4b31-9156-809927b5ee49



# Estrategia Nacional de Ciencia Abierta (2023-2027).



- ▶ B2. Incluir una adecuada *valoración de otros resultados de investigación,* adoptar nuevas formas de comunicación científica basadas en ellos, *superar las limitaciones del sistema de comunicación científica*..
- CÓMO: Las publicaciones científicas se deberán <u>vincular con sus datos de investigación subyacentes</u>. Se hará <u>obligatoria la realización de un plan de gestión de datos (PGD)</u> como parte integral de los proyectos de investigación financiados con fondos públicos, incluyendo la evaluación de este plan dentro de las actividades de seguimiento de los proyectos, siempre que la disciplina objeto de la investigación conlleve la existencia de datos. Este PGD tendrá un estándar que facilite su evaluación y seguimiento, de acuerdo con un procedimiento y protocolo acorde.



# POLÍTICA DE INTEGRIDAD CIENTÍFICA Y BUENAS PRÁCTICAS EN INVESTIGACIÓN



- Registro organizado de los datos obtenidos.
- Mantener los datos de 3 a 5 años.
- Mantener el acceso abierto si es compatible con la propiedad intelectual y la privacidad.
- Mantener los datos en línea con los principios FAIR.
- Buscar repositorios en re3data.org
- Colección Datasets dentro de Riunet.
- Asegurar la privacidad y confidencialidad de los datos personales.



### 12 meses

• Proyectos de investigación a corto plazo que sirven solamente para fines de evaluación (estudiantes).

### 15 años

• Ensayos clínicos.

### Permanentemente

- Terapia genética.
- Si el trabajo tiene valor parala comunidad o el patrimonio.

### Recomendación general

• 5 años

### POLÍTICA DE INTEGRIDAD CIENTÍFICA Y BUENAS PRÁCTICAS EN INVESTIGACIÓN

### 3 a 5 años

• De investigaciones publicadas.

### 25 años

• Ensayos clínicos.





### 7 años

• De investigaciones publicadas.



# Herramientas comerciales



Share your data

https://www.globus.org/compute

Microsoft Power BI: paseo quiado

### Bienvenido

En esta visita quiada, le mostraremos cómo Sam, un analista comercial de Contoso, puede actualizar y publicar rápidamente su informe de potencial de ventas. Los compañeros de Sam pueden usar las capacidades de IA integradas en Power BI para analizar rápidamente y obtener información de los datos de ventas y clientes de Contoso y luego compartir esos conocimientos con sus compañeros de equipo.



Paso 2 de 2



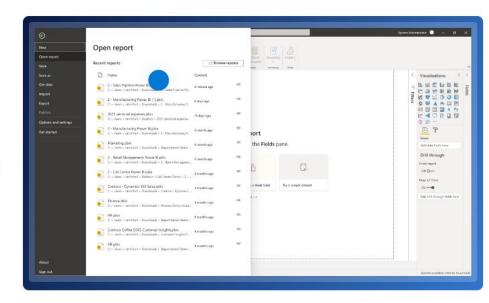
Mejorar un modelo de datos

Crear una nueva página de informes

Descubrir información a partir de objetos visuales inteligentes

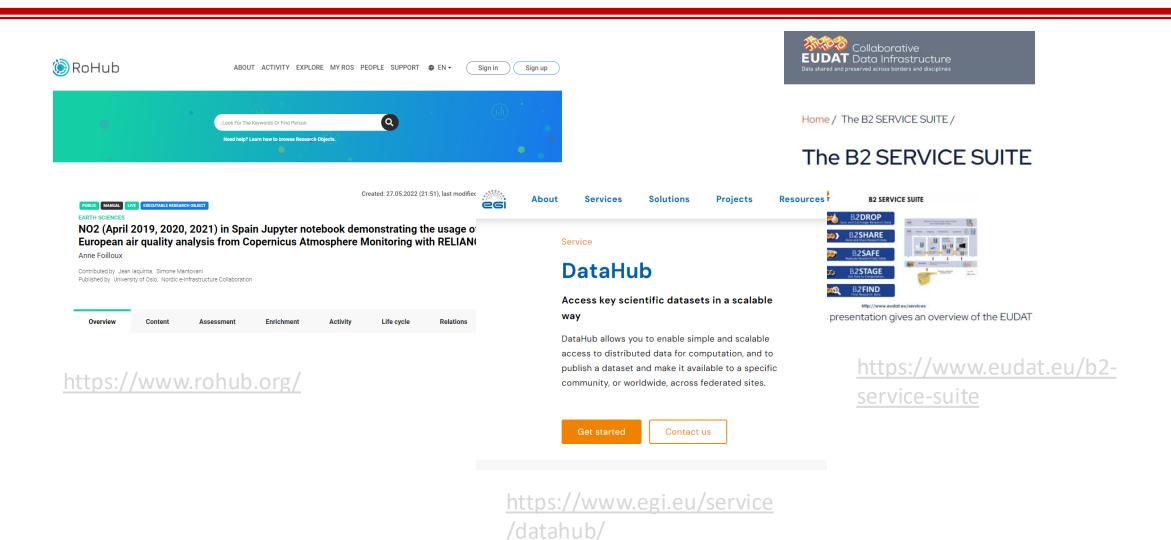
Integración con sus aplicaciones

Resumen



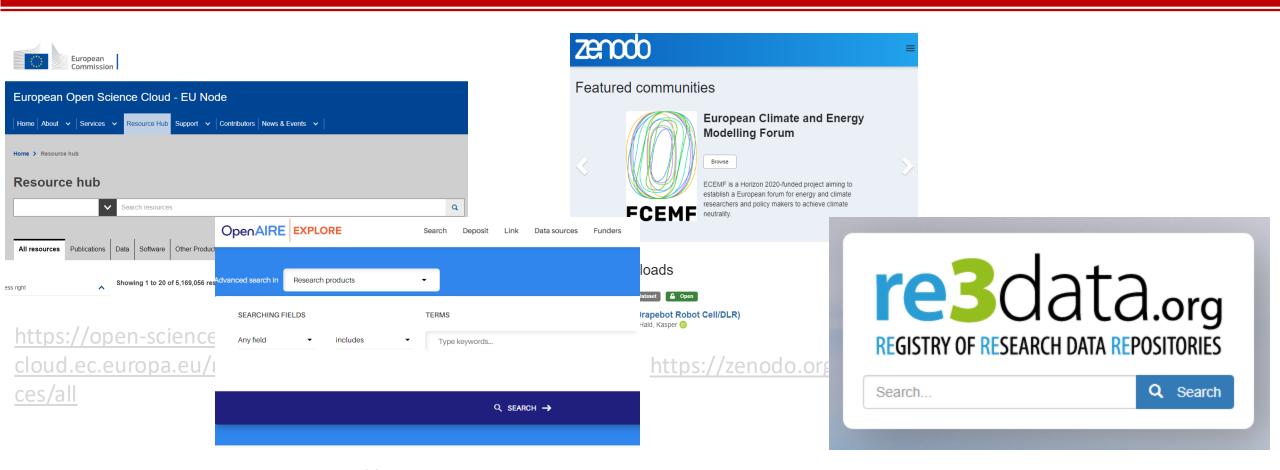


# **ALMACENAR DATOS**





# **ENCONTRAR DATOS**



https://explore.openaire.e u/search/advanced/researc h-outcomes

https://www.re3data.org/



# **SERVICIOS DE APOYO UPV.**

 Delegación de Protección de Datos (DPD\_UPV): encomendadas la supervisión del cumplimiento del Reglamento General de Protección de Datos y de la Ley Orgánica de Protección de Datos Personales y garantía de los derechos digitales en el ámbito de la Universitat Politècnica de València.

Comité de Ética en Investigación de la UPV: cualquier actividad llevada a cabo en la UPV cuya experimentación esté afectada por alguna normativa, por incluir aspectos con implicaciones éticas o de bioseguridad, debe ser previamente autorizada por el Comité de Ética en Investigación de la UPV. El investigador principal debe seguir el procedimiento y cumplimentar el formulario correspondiente para solicitar la evaluación por parte del comité.









BIBLIOTECA Y DOCUMENTACIÓN CIENTÍFICA

Universitat Politècnica de València

www.upv.es · rienma@upvnet.upv.es





### **DMP** timeline

### When should the DMP be ready?

- A short (1-page) DMP is required at the proposal stage.
- · A full, initial version of the DMP is required as deliverable; normally by month 6.
  - Note: by exception, in cases of a public emergency and if the work programme requires so, you should submit a full DMP already with submission of proposals or at the latest by the signature of the grant agreement.
- The DMP is considered 'a living document' and has to be regularly updated to reflect changes that may arise or decisions that are implemented. For projects longer than 12 months, an updated version of the DMP has to be submitted as a deliverable.
- A final version of the DMP that describes how the data is managed and shared has to be delivered at the end of the project.



EJE B. Gestión de datos de investigación siguiendo los Principios FAIR . Se hará <u>obligatoria</u> la elaboración de un plan de gestión de datos como parte integral de los proyectos de investigación incluyendo su evaluación dentro de las actividades de seguimiento. MARCO TEMPORAL: <u>medio plazo</u>.







# Instrucciones para la elaboración del Plan de Gestión de Datos (PGD)

Para completar la justificación científico-técnica final es **indispensable** rellenar el Plan de Gestión

Un Plan de Gestión de Datos es una declaración formal que describe los datos que se utilizan y producen durante el curso de las actividades de investigación. El PGD establece cómo se van a gestionar los datos en un proyecto de investigación, tanto durante el proyecto como una vez que éste ha finalizado. 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.

La Comisión Europea proporciona orientación sobre cómo elaborar un plan de gestión de datos en sus directrices para gestionar datos según el esquema FAIR.

Se puede utilizar cualquier plantilla o herramienta para crear el PGD.

- Resumen de los datos: qué datos se han generado/recogido, qué formatos y estándares se han utilizado, qué valor tienen los datos para otros investigadores, qué datos no se pueden compartir y por qué.
- Responsabilidades: personas responsables, quién determina el acceso a los datos, titularidad de propiedad intelectual, la gestión de los datos y de su preservación y conservación.
- Medidas para garantizar que los datos sean FAIR.
- Seguridad de los datos:
  - Tratamiento de la recuperación de los datos, almacenamiento seguro y el traspaso de datos sensibles y las medidas adoptadas para la seguridad de los datos (almacenamiento y copias de seguridad).
  - Revisión ética (describir, si fuera de aplicación, los problemas éticos en la recogida, almacenamiento, procesamiento y archivo de los datos, así como en los procedimientos de aprobación ética relacionados con el proyecto).

https://www.aei.gob.es/sites/default/files/convocatory\_into/z 023-07/Plan%20de%20Gestión%20de%20Datos%202020.pdf



El PGD es un documento donde se establece cómo se manejarán los datos/códigos de un proyecto de investigación en el corto, mediano y largo plazo. Un PGD es un documento que puede cambiar en el tiempo.

# ¿POR DÓNDE EMPEZAR?

- 1. Establecer responsabilidades (rol de gestión de datos)
- 2. Definir flujos de trabajo según distintos tipos de datos.

DISTINGUIR siempre:

- ✓ Datos ESTÁNDAR que se pueden abrir.
- ✓ Datos que no pueden abrirse (personales-sensibles sin anonimizar, comerciales, potencialmente peligrosos, etc.). JUSTIFICAR.



3 tipos de datos que deben describirse separadamente.

Datos <u>brutos/</u>
<u>primarios</u>: datos
recogidos de la
fuente (hay que
conservar
siempre una
versión de sólo
lectura).

Datos
procesados:
una versión de
los datos que se
ha modificado
para su análisis
o visualización.

Datos <u>finales</u>: datos listos para ser compartidos en una publicación o repositorio de datos.



# Partes comunes en PGD

# Descripción de datasets

 Descripción detallada de los datos que se van a generar y/o reutilizar: relevancia científica, enfoque metodológico para relacionarlos entre ellos, tipos y volumen estimado.

# Estándares y metadatos

- Protocolos y estándares que se van a usar para estructurar los datos de tal manera que otros investigadores puedan descubrirlos, evaluarlos y reproducirlos
- Si es relevante, mención a los estándares de la comunidad científica a la que pertenece

# Identificadores persistentes y curación de los datos

- PIDs que se usarán para que los datos puedan resolverse (abrirse) y citarse
- Estándares que asegurarán la integridad de los datos
- Metodología de preservación y conservación a largo plazo
- Criterios que hacen del repositorio uno de confianza
- Detalles que informan sobre la modalidad de acceso a los datos y términos de uso (licencia)



Descripción del ciclo de vida de la gestión de todos los conjuntos de datos recopilados, procesados o generados en el proyecto.

# Documento mínimo:

- ✓ Tipos de datos.
- ✓ Estándares para registro y metadatos.
- ✓ Políticas de acceso, compartición y protección.
- ✓ Políticas de reutilización, distribución y generación de datos derivados.
- ✓ Archivo y preservación.

<u>Plan de Gestión de Datos de investigación UPV: guía para la elaboración de un Plan de Gestión de Datos de investigación en el contexto de la Universitat Politècnica de València (UPV).</u>





- Datos administrativos.
- Recopilación de datos.
- 3. Documentación y metadatos.
- 4. Ética y cumplimiento legal. Propiedad intelectual.
- 5. Almacenamiento y copias.
- 6. Selección y preservación.
- 7. Intercambio de datos.
- 8. Responsabilidad y recursos.



**Datos administrativos**: financiador, patrocinador, nº de subvención, proyecto, IP (ORCID), fechas de la 1º versión y de las actualizaciones del PGD, políticas relacionadas, datos de contacto...

**Recopilación y descripción de los datos**: resumen (naturaleza, contenido, alcance), propósito, procedencia, técnicas, tipo de datos, formato y volumen, control de versiones.

**Documentación y metadatos**: esquemas, vocabularios y terminologías, documentos que acompañarán a los datos.

**Ética y cumplimiento legal:** Propiedad intelectual: datos personales, consentimientos, privacidad, anonimización, seguridad, licencias, restricciones, embargos.

**Almacenamiento y copias**: previsiones, medios, infraestructura, copias, restauración, responsabilidades, accesos.

Selección y preservación: preservación y compartición, plazos, repositorio, presupuestos.

**Intercambio de datos**: localización, disponibilidad, identificadores, condiciones, restricciones, citación.

**Responsabilidad y recursos**: funciones y responsables, recursos.



Criterios para decidir qué datos conservar y qué datos eliminar.

Valorar los costes

Identificar los datos que deben preservarse por mandato

Identificar los propósitos que pueden cumplir los datos(reutilización + visibilidad)





### 8 Recomendaciones para tus datos de investigación

Enríquez Malavé, R. (2023). 8 Recomendaciones para tus datos de investigación. http://hdl.handle.net/10251/193421

Por favor, use este identificador para citar o enlazar este ítem: http://hdl.handle.net/10251/193421

### Ficheros en el ítem



### Metadatos del ítem

8 Recomendaciones para tus datos de investigación Título:

Nombre: 8 Recomendaciones

Tamaño: 452.9Kb Formato: PDF

Autor: Enríquez Malavé, Ricardo

Universitat Politècnica de València. Biblioteca y Documentación Científica - Biblioteca i Documentació **Entidad UPV:** 

Científica

Fecha difusión: 2023-05-16

Se ofrecen recomendaciones elaboradas por el Servicio de Documentación Científica de la biblioteca de la Resumen:

Universitat Politècnica de València para trabajar con datos de investigación

Ciencia abierta, Open Science, Gestión de datos de investigación, Research data management. Palabras clave:

<u>investigación</u>

Reconocimiento (by)

Material docente

http://hdl.handle.net/10 251/193421

- Asocia siempre los datos con las publicaciones relacionadas.
- Guarda y protege los raw data.
- Realiza copias de seguridad.
- Haz tus datos FAIR.
- Haz los datos amigables para humanos.
- Haz los datos amigables para máquinas.
- Deposita los datos.
- Publica un "Data paper".
- ✓ Utiliza metadatos estandarizados. Ej: países, términos químicos o médicos:
  - o ISO https://www.iso.org/iso-3166-country-codes.html
  - o IUPAC International Chemical Identifier (InChITM) https://iupac.org/who-we-are/divisions/division-details/inchi/
  - Medical Subject Headings (MeSH) https://www.nlm.nih.gov/mesh/meshl\*Ej: 01 Surveydaten2019\_V3\_20201121\_GOM.
- ✓ Acuerda procedimientos internos desde ϵ
- ✓ Recopila el máximo número de variable aunque no sirvan para el propósito inicia
- ✓ Utiliza siempre las mismas estructuras d provectos.
- □ 01 Paquete de trabajo.
- Surveydaten2019 Datos del experimento.
- V3 Versión.
- 20201121 Fecha.
- GOM Autor(a).
- ✓ Denomina por áreas nunca por investigadores.
- ✓ Establece un máximo de niveles de subcarpetas.



## HERRAMIENTAS PARA ELABORAR UN PGD









#### Argos: guía de uso

Enríquez Malavé, R. (2022). Argos: guía de uso. http://hdl.handle.net/10251/190729

Por favor, use este identificador para citar o enlazar este ítem: http://hdl.handle.net/10251/190729

#### Ficheros en el ítem



#### Metadatos del ítem

Título: Argos: guía de uso

Autor: <u>Enríquez Malavé, Ricardo</u>

Entidad UPV: Universitat Politècnica de València. Biblioteca y Documentación Científica - Biblioteca i Documentació

Cientilica

Fecha difusión: 2022-12-15

[ES] Se ofrece un procedimiento para elaborar un Plan General de Datos así como la descripción de los

sets de datos a través de la herramienta online ARGOS.

Resumen:

[EN] A procedure for preparing a General Data Plan with the ARGOS online tool is provided, as well as a

description of the datasets included in the Plan.

Palabras clave: Research data management , Datasets , Gestión de datos de investigación , Data management plan

<u>Datos de investigación</u>, <u>Argos</u>, <u>Plan de gestión de datos</u>, <u>Ciencia abierta</u>, <u>Open science</u>

Derechos de uso: Reconocimiento - No comercial - Compartir igual (by-nc-sa)

Tipo: Material docente

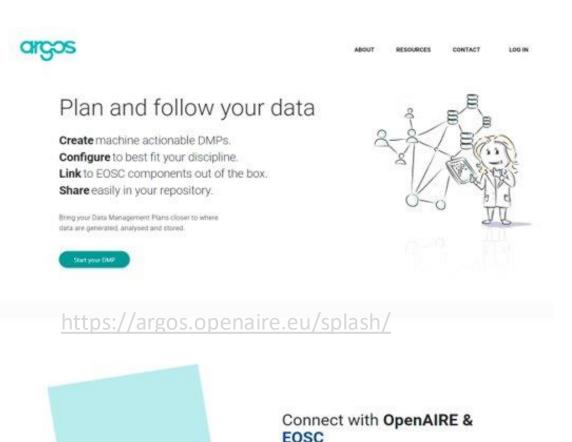
Contexto: Primer ciclo

Idioma del destinatario: Español

http://hdl.handle.net/10 251/190729



## HERRAMIENTAS ON LINE.



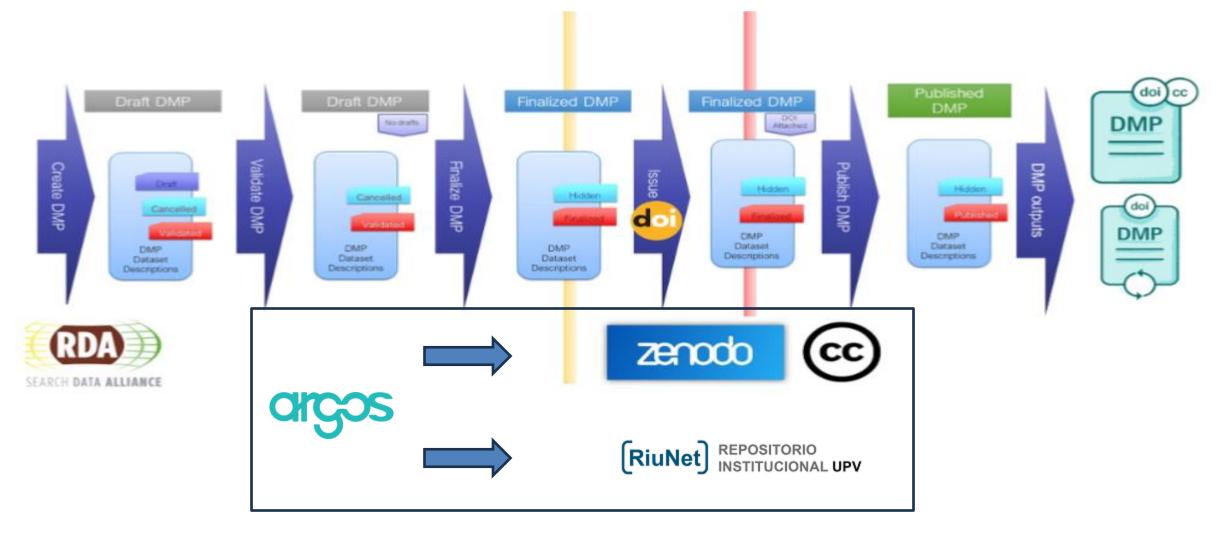
- Gestión colaborativa de los PGD.
- Gestión del PGD como un producto de investigación: asignación de DOI, licencias, etc., lo que permite su publicación y visibilidad.
- Relacionar un PGD con uno o varios datasets dentro de un proyecto de investigación.
- Diferentes plantillas para cargar los sets de datos.
- Funcionamiento con asistentes.
- Vincula con la cuenta que tengamos en ZENODO: podemos publicarlo en abierto.

Guía de uso: <a href="https://riunet.upv.es/handle/10251/190729">https://riunet.upv.es/handle/10251/190729</a>

Link your plan directly to underlying OpenAIRE and EOSC services

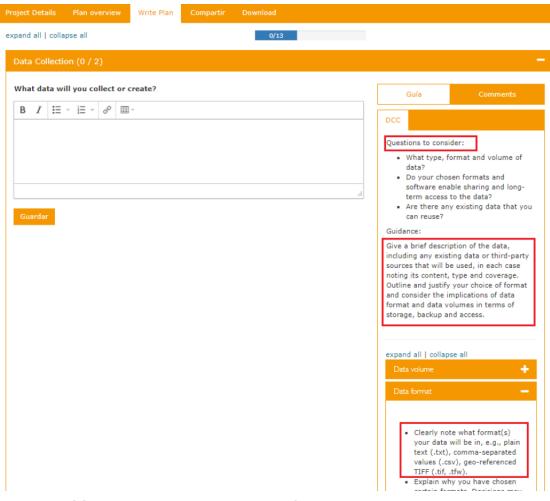
sources and semantics and trace the quality of your research. Use







## HERRAMIENTAS ON LINE.



Ejemplos: <a href="https://dmponline.dcc.ac.uk/public plans">https://dmponline.dcc.ac.uk/public plans</a>



https://www.consorciomadrono.es/investigam/pagoda/



Data Stewardship Wizard (DSW).

Provided by 券UC3 10 D C C

https://dmponline.dcc.ac.uk/



# PRÁCTICA: Redactar un PGD (versión inicial)



# Plan and follow your data

Create machine actionable DMPs.

**Configure** to best fit your discipline.

**Link** to EOSC components out of the box.

**Share** easily in your repository.

Bring your Data Management Plans closer to where data are generated, analysed and stored.

Start your DMP

- 1. Entrad en ARGOS https://argos.openaire.eu/splash/
- 2. Identificaos con el ORCID
- 3. Elegid plantilla Horizonte Europa o similar.
- 4. Redactar un entregable de 1 hoja con los elementos mínimos de vuestra investigación.
- 5. Subidlo al Poliforma-T

Tiempo: 15 minutos.

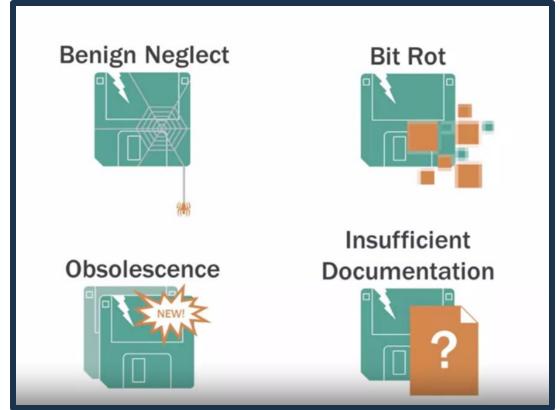
#### O bien:

- 1. Seleccionar uno de: <a href="https://libereurope.eu/working-group/research-data-management/plans/">https://libereurope.eu/working-group/research-data-management/plans/</a>
- 2. Adaptarlo.
- 3. Subidlo al Poliforma-t









Principales riesgos asociados con la pérdida de datos digitales

- ✓ Negligencia
- ✓ Podredumbre de bits
- ✓ Obsolescencia
- ✓ Documentación insuficiente



Los principios FAIR minimizan el riesgo de pérdida de datos.



## 5. PRINCIPIOS FAIR.

Los datos de investigación deben cumplir con unos <u>requisitos</u> para que sean localizables, accesibles, interoperables y reutilizables: Pautas y buenas prácticas desarrolladas por los propios investigadores.

Decisiones tomadas durante la investigación. Proveedores cuyos servicios permitan alcanzar el carácter FAIR de los datos.





#### FINDABLE

Data has rich metadata and unique identifier

#### CONSEJOS:

- Revisar los perfiles ResearchID y Scopus ID y la firma bibliográfica.
- Sincronizarlos con ORCID para actualizar las publicaciones en ambos.
- Introducirlos en SENIA.
- FIRMAR los datasets como firmamos las publicaciones.
- VIGILAR cómo se hace el depósito en repositorios temáticos (si no lo hacemos nosotros).

Identificadores únicos y persistentes (PID) a nivel de objeto, persona o institución.

Metadatos completos que permiten que puedan ser comprendidos conforme a los estándares de cada disciplina.

Indexación por parte de agregadores académicos y motores de búsqueda.

Localizar el repositorio en Re3data.

# ResearchId ROR Handle DOI









#### **ACCESSIBLE**

Data can be easily downloaded or used by using standard protocols

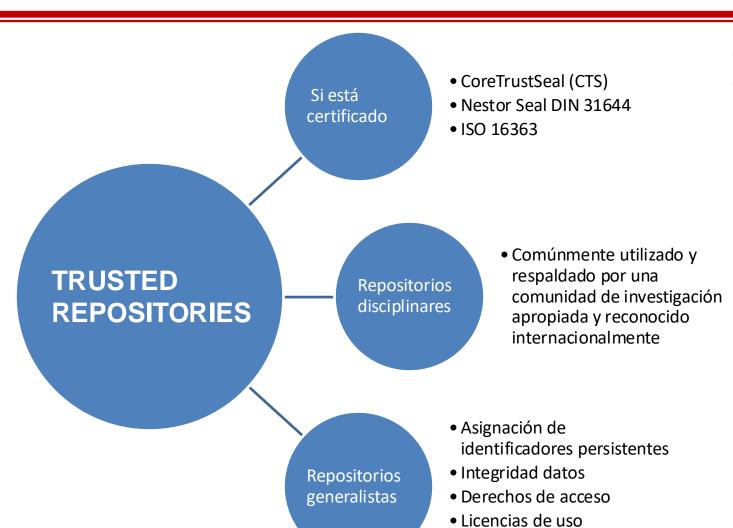
*Protocolos* estandarizados, abiertos, gratuitos e implementables.

Procesos claros de *autenticación*. *Documentar* métodos y herramientas complementarias de acceso.

Descarga segura.

*Metadatos* siempre accesibles, incluso si los datos dejan de estar disponibles.





Preservación



#### **Certified repositories**

E.g. CoreTrustSeal, nestor Seal DIN31644, ISO16363 etc. https://www. coretrustseal. org/

https://www.is o.org/standard /56510.html

# \* Disciplinar

## Disciplinary or domain specific repos

Commonly used, endorsed by the research communities and internationally recognized





General purpose or institutional repositories



## Criterios a tener en cuenta para evaluar un repositorio

- Madurez de la infraestructura tecnológica.
- Sostenibilidad, financiación y transparencia. Políticas.
- Persistencia digital a largo plazo y preservación
- Servicios de apoyo y de valor añadido, equipos de apoyo. Relevancia de contenidos. Usabilidad.
- Adopción de estándares y buenas prácticas ampliamente conocidas.

https://www.rebiun.org/grupos-trabajo/repositorios/evaluacion



 $Universitat\ Politècnica\ de\ València.\ Unidad\ de\ Documentación\ Científica\ de\ la\ Biblioteca\ (+34)\ 96\ 387\ 70\ 85\cdot \underline{RiuNet@bib.upv.es}$ 



El contenido de este sitio está bajo una licencia Creative Commons Reconocimiento – No Comercial – Sin Obra Derivada (by-nc-nd), salvo que se indique lo contrario.



Los metadatos de este sitio están bajo una licencia Dominio Público.



ombre del repositorio: RiuNet RL del repositorio: https://riunet.upv.es/ iombre de la institución: Politécnica de Valencia riterios cumplidos: 24



## **ESQUEMAS DE METADATOS.**

- ✓ List of Metadata Standards
- ✓ <u>Directorio de metadatos</u> del RDA Metadata Standards Directory Working Group.
- ✓ <u>Vocabularios controlados.</u>

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

#### **General Purpose**

- ISO 639-2 Codes for Representation of Names of Languages
- Library of Congress Subject Headings (LCSH)
- · Library of Congress Name Authority File (LCNAF)
- · Virtual International Authority File (VIAF)

#### **Arts & Humanities**

- Art and Architecture Thesaurus (AAT)
- Cultural Objects Name Authority (CONA)
- Getty Thesaurus of Geographic Names (TGN)
- Rare Books and Manuscripts Section (RBMS) Controlled Vocabularies
- · Thesaurus for Graphic Materials (TGM)
- Union List of Artist Names (ULAN)

#### Physical & Life Sciences

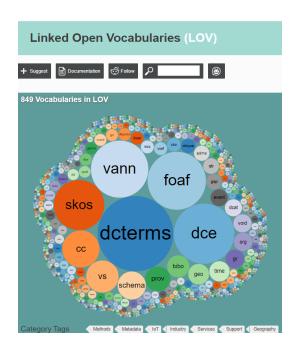
- AGROVOC Multilingual Agricultural Thesaurus
- · International Classification of Diseases (ICD)
- Medical Subject Headings (MeSH)
- National Agriculture Library Thesaurus (NALT)

#### Social & Behavioral Sciences

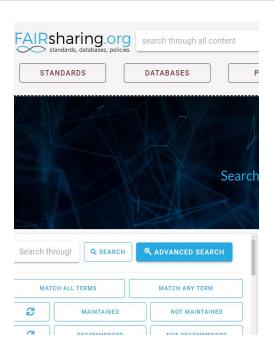
- · Data Documentation Initiative (DDI) Controlled Vocabularies
- Inter-university Consortium for Political and Social Research (ICPSR) Thesaurus
- United Nations Educational, Scientific and Cultural Organisation (UNESCO) Thesaurus



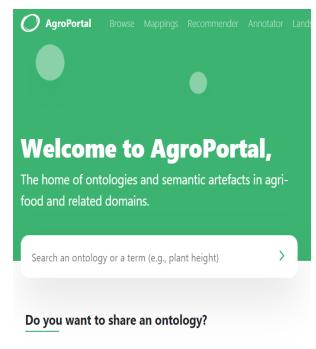
## ¿Dónde buscar vocabularios controlados?



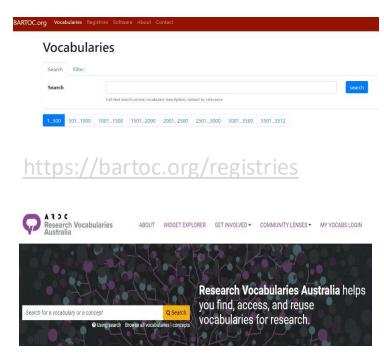
https://lov.linkeddata.es
/dataset/lov/



https://fairsharing.org /search/?q=Ecology



https://agroportal.lirmm.fr/



https://vocabs.ardc.edu.au/





### INTEROPERABLE

Metadata use an accessible and standard language

#### **FORMATOS:**

- ✓ Considerar normas y costumbres en cada disciplina.
- ✓ Formatos abiertos garantizan el acceso a largo plazo.
- Si se usan otro tipo de formatos, lo recomendable es convertir los datos a formatos abiertos.
- ✓ Recomendados : <u>UK DataService</u>, <u>Library of</u> Congress.

Vocabulario estándar, controlado, formal, accesible, reconocido y compartido (≠ palabras clave, etiquetas.)

Datos integrables y combinables con otras aplicaciones, herramientas y otros datos: formatos abiertos.

Referencias cruzadas y enlaces.







## REUSABLE

Data is well-described and provides clear usage of licences

Licencias abiertas.

CreativeCommons OpenDataCommons

https://www.openaire.eu/ how-do-i-license-myresearch-data

Documentación asociada: plantilla README, información contextual, guías, procedimientos, etc.

Plantilla READme

Cornell University



## DOCUMENTACIÓN.

## Plantilla READme

Incluye la información necesaria para <u>que personas ajenas a la investigación puedan</u> <u>comprender y utilizar los datos</u>.

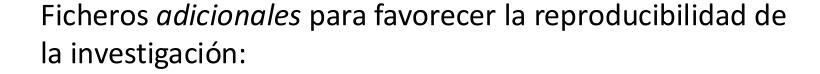
- Información general: título del dataset, equipo investigador (contacto, afiliaciones ...), fechas, financiación, descripción, palabras clave.
- Información para compartir y acceder a los datos: licencias, enlaces a las publicaciones relacionadas, propuesta de cita para los datos, etc..
- Información sobre la organización de archivos: listado de archivos y sus contenidos, versiones, tamaño, etc.
- Información metodológica: métodos utilizados para la recopilación, generación y procesamiento de los datos y software necesario para interpretar los datos.
- Información específica sobre los datos: contenido exacto, número de variables, listas de códigos, abreviaturas utilizadas, etc.











#### ENCUESTA SOBRE ACTITUDES HACIA LA INMIGRACIÓN Y LOS INMIGRANTES EN ESPAÑA (EASIE SURVEY)

LIBRO DE CÓDIGOS para usuarios del fichero de datos depositado en DIGITAL.CSIC

V1 (septiembre de 2023)

Financiación: Encuesta realizada en el marco del proyecto de I+D+i "Explicando Actitudes Sosegadas hacia los Inmigrantes en España" (CSO2017-87364-R), financiado por MCIN/ AEI/10.13039/501100011033/ y por "FEDER Una manera de hacer Europa".

Ponderación: Para mejorar la representatividad de la muestra es necesario utilizar la ponderación por la variable "PESO".

Número de filas en el archivo de datos: 2.344 registros.

- ✓ Codebook: conjunto de códigos, definiciones y ejemplos que se utililizan a menudo como guía para proporcionar contexto y ayudar a analizar los datos, esenciales para analizar la investigación cualitativa (por ejemplo, encuestas).
- ✓ Data dictionary: colección de nombres, definiciones y atributos sobre elementos de datos que se utilizan o captura, en una base de datos, sistema de información o parte de un proyecto de investigación.



## Redactar una plantilla README



### Cornell Data Services

Services Data Management ✓ News Workshops

Home > Data Management > Sharing > Guide to writing "readme" style metadata

### Guide to writing "readme" style metadata

A readme file provides information about a data file and is intended interpreted, by yourself at a later date or by others when sharing or generally preferable, but where no appropriate standard exists, for ir an appropriate strategy.



Download a template and adapt\* it for your own da

- Entrad en <a href="https://data.research.cornell.edu/data-">https://data.research.cornell.edu/data-</a> management/sharing/readme/
- Descargad la plantilla: <a href="https://cornell.app.box.com/v/ReadmeTemplate">https://cornell.app.box.com/v/ReadmeTemplate</a> o en: https://upv-es.libguides.com/ld.php?content\_id=35338971
- Cumplimentad los datos con vuestro último proyecto.
- Subidlo al Poliforma-T
- Consultad ejemplos en RIUNET, colección DATASETS.

Tiempo: 10 minutos.

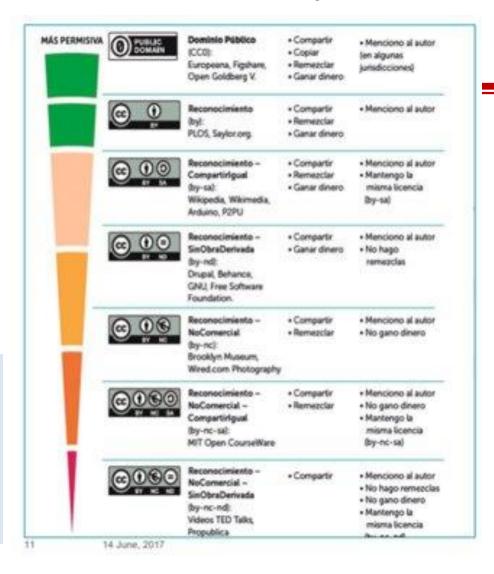


# **Creative Commons / Open Data Commons** ¿Dónde?

- Al cargar los datos en el repositorio.
- En los metadatos.
- En el archivo README.txt

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

 Versión 4.0 o posteriores (las anteriores no son específicas para datos).



https://www.openaire.eu/how-do-i-know-if-my-research-data-is-protected



## **Licencias** Open Data Commons.

Son más específicas para bases de datos <a href="https://opendatacommons.org/">https://opendatacommons.org/</a>

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

This {DATA(BASE)-NAME} is made available under the Open Data Commons Attribution License: <a href="http://opendatacommons.org/licenses/by/{version}">http://opendatacommons.org/licenses/by/{version}</a>.

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

This {DATA(BASE)-NAME} is made available under the Open Database License: <a href="http://opendatacommons.org/licenses/odbl/1.0/">http://opendatacommons.org/licenses/odbl/1.0/</a>.

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



#### Licencias de Dominio Público:

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

# CC0 1.0 Universal (CC0 1.0) Public Domain Dedication



Implica renuncia a los derechos personales sobre el trabajo.

### Open Data Commons Public Domain Dedication and License (PDDL)

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

#### License (SPDX IDs)

Creative Commons CCZero (CC0-1.0)

Open Data Commons Public Domain Dedication and Licence (PDDL-1.0)

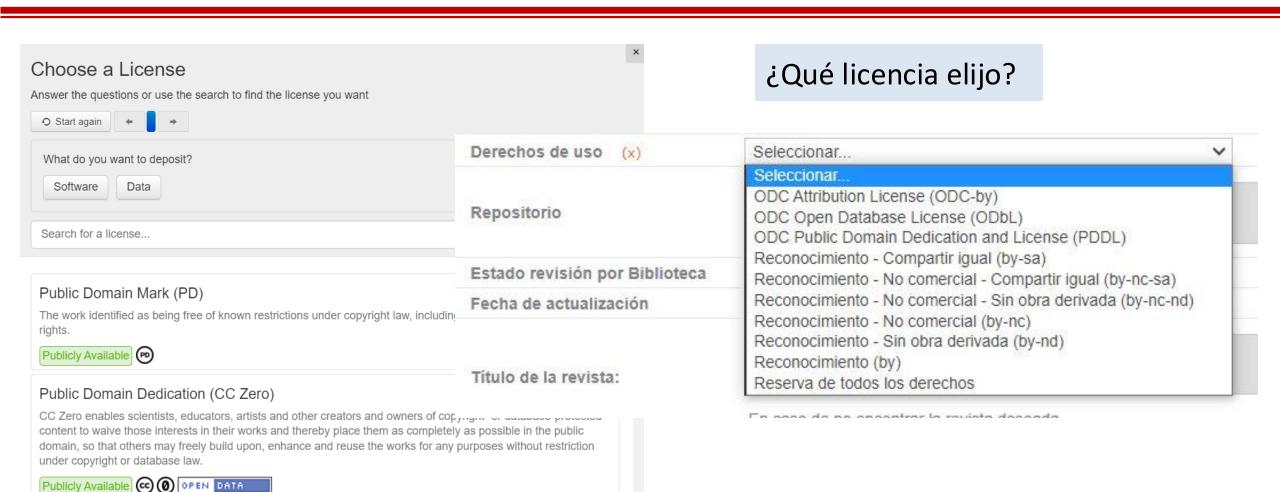
Creative Commons Attribution 4.0 (CC-BY-4.0)

Open Data Commons Attribution License (ODC-By-1.0)

Creative Commons Attribution Share-Alike 4.0 (CC-BY-SA-4.0)

Open Data Commons Open Database License (ODbL-1.0)





http://ufal.github.io/public-license-selector/ https://creativecommons.org/choose/?lang=es\_ES



## Buenas prácticas de trabajo con datos

### Carpetas

- Seguir procedimientos internos.
- Utilizar siempre las mismas estructuras de carpetas incluso entre proyectos.
- Denominar por áreas nunca por investigadores.
- Establecer un máximo de niveles de subcarpetas.
- Separar trabajos en curso de los ya finalizados.
- Evaluar periódicamente los materiales y eliminar archivos intermedios innecesarios.

#### **Archivos**

- No utilizar nombres de ficheros demasiado largos.
- Evitar caracteres especiales ~ ¡ ! @# \$ % ^ & \* ( ) `; < > ¿ ? , [ ] { } ' " |
- Guion bajo\_ en lugar de espacio en blanco
- Indicar al final del nombre el número de versión, v01, v02, etc.
- Utilizar nombres únicos y numerar de forma continua: 001, 002, 003, etc.
- Comenzar con el nombre más común, el más específico detrás.
- Si hay que poner al autor especificar por iniciales.
- Evitar interfaz y programas intermedios para operaciones masivas, mejor con comandos.

#### Formatos

- Bases de datos: XML, CSV.
- Texto: RFT, TXT, XML.
- Estadísticas: ASCII, DTA, POR, SAS, SAV.
- Datos tabulados: CSV, TSV.
- Geoespaciales: SHP, DBF, GeoTIFF, NetCDF.
- Vídeo: OGG, MP4.
- Sonido: FLAC, WAV, AIFF, MP3.
- Imágenes: TIFF, BMP.
- Ficheros comprimidos: no se recomienda el uso de ficheros comprimidos. Hay formatos que pierden información en la compresión.

https://dans.knaw.nl/en/file-formats/



#### 1. **FINDABLE**:

- ✓ Datos con id. persistente (evitamos el 404 not found). El IP debe aparecer en el registro de metadatos.
- ✓ Metadatos en línea en recursos de búsqueda (los motores de búsqueda localizarán los datos).
- ✓ Elegir un repositorio adecuado (si los datos no se pueden compartir compartiremos los metadatos).
- ✓ Incluir el ID persistente en la publicación asociada.

#### 2. ACCESSIBLE:

- ✓ Elegir un <u>repositorio</u> adecuado. Metadatos siempre disponibles, aunque los datos no estén abiertos.
- ✓ Elegir formatos abiertos o estándar, convertir si es necesario.
- ✓ <u>Documentación</u> adecuada: README, metadatos en XML, vocabularios y nomenclaturas según disciplinas...
- ✓ Documentar los procesos de autenticación y autorización si los hay.

#### 3. INTEROPERABLE:

- ✓ Elegir <u>formatos</u> *abiertos* o estándar *reconocibles*, convertir si es necesario.
- ✓ Documentación adecuada: README, metadatos en XML, vocabularios y nomenclaturas según disciplinas...
- ✓ Lenguaje formal y accesible. Vocabulario controlado.
- ✓ Referencias y enlaces a otros datos.

### 4. **REUSABLE**:

- ✓ Elegir formatos abiertos o estándar, convertir si es necesario.
- ✓ Documentación adecuada: README, metadatos en XML, vocabularios y nomenclaturas según disciplinas...
- ✓ Usar <u>licencias</u> abiertas.

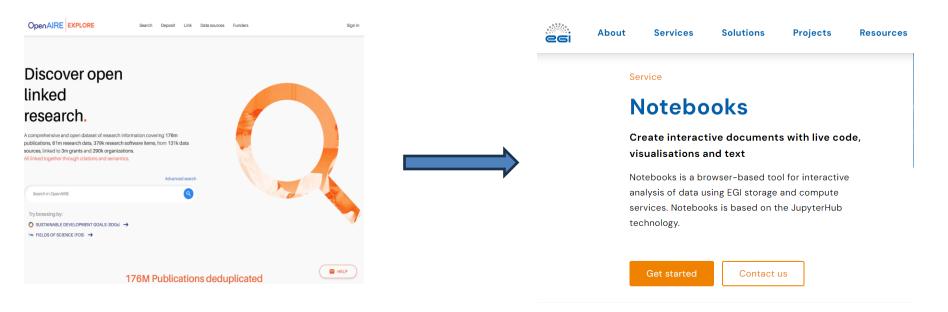


## **UN EJEMPLO EXITOSO**

Demo showing what RELIANCE project has achieved on Open Science, FAIR and EOSC:



En <u>OpenAIRE EXPLORE</u> buscamos "**Copernicus air quality**". Encontramos como "Software" un "notebook EOSC Jupyter" creado por <u>Simone Mantovani</u> con un DOI y metadatos adicionales para que OpenAIRE explore pueda "asociarlo" a un servicio EOSC específico: <u>EGI Notebook.</u>





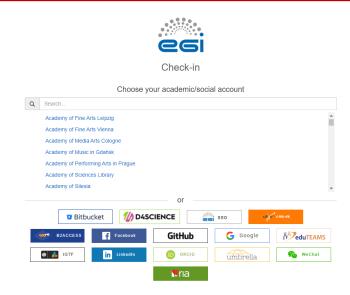
## UN EJEMPLO EXITOSO



Para el acceso se realiza una solicitud de pedido donde es posible que deba explicar por qué le gustaría acceder a este servicio EOSC. Para autenticarse en cualquier servicio EOSC se recomienda utilizar el identificador ORCID.



Iniciamos sesión en el EGI notebook usando el identificador ORCID y cargamos (manualmente) el cuaderno Jupyter que encontramos en OpenAIRE (siguiendo el enlace, por ejemplo, de zenodo (<a href="https://doi.org/10.5281/zenodo.5554786">https://doi.org/10.5281/zenodo.5554786</a>) El cuaderno de Jupyter utiliza el análisis de calidad del aire europeo CAMS del Servicio de Monitoreo de la Atmósfera Copernicus. Se puede acceder a los datos de entrada a través de un servicio externo llamado plataforma ADAM (Plataforma avanzada de gestión de datos geoespaciales).



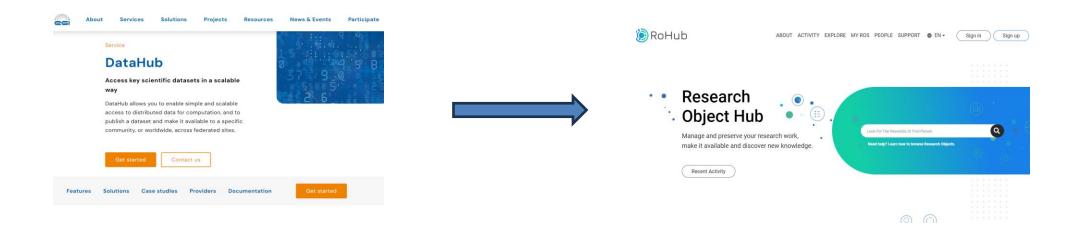




## UN EJEMPLO EXITOSO



- ✓ Podemos volver a ejecutar el cuaderno Jupyter y podemos crear trabajos derivados si la licencia de los datos y software original lo permite.
- ✓ Duplicamos el cuaderno Jupyter y lo personalizamos, utilizando un sistema de almacenamiento que permita compartir. Utilizamos "<u>EGI datahub</u>".
- ✓ El nuevo análisis es muy similar al anterior pero sobre una zona geográfica diferente (Francia).
- ✓ Finalmente, creamos un nuevo objeto de Investigación que agrega todos los recursos. Usamos otro servicio externo llamado <u>RoHub (Research Object Hub</u>) y creamos un "Objeto de investigación ejecutable" que esperamos sea encontrado, accedido y reutilizado.





# PRÁCTICA:¿Cómo de fair son tus datos?



#### SATIFYD

Self-Assessment Tool to Improve the FAIRness of Your Dataset

Welcome to SATIFYD: the DANS Self-Assessment Tool to Improve the FAIRness of Your Dataset. This tool will show you how FAIR (Findable, Accessible, Interoperable, Reusable) your dataset is and will provide you with tips to score (even) higher on FAIRness. Ideally, you use this tool prior to the deposit in EASY.

The 12 questions touch upon the FAIR data principles of but do not strictly follow them. While answering the questions, the score per letter will be displayed underneath each letter. The more 'blue' the letters get, the more FAIR your dataset is. An overall score is provided at the end of the page. Some questions are posed more than once (e.g. on metadata and data standards or usage licences), because the topics are relevant in more than one letter.

Want to know more? Please click here ←

If you have any questions, please let us know by sending an e-mail .

FINDABLE		i
Did you provide sufficient metadata (information) about your data for others to find, understand and reuse your data?	v	•
Did you use standards such as controlled vocabularies, taxonomies (thesauri) or ontologies to describe your dataset?	Controlled vocabularies Taxonomies (thesauri) Ontologies There are no standards for my discipline	
3. Did you provide rich and detailed additional documentation?	□ Readme file □ Versioning □ Provenance	
ACCESSIBLE		ā
4. Is the metadata publicly accessible even if the data is no longer available? $\ensuremath{\textcircled{\scriptsize 1}}$	Yes No	
5. Does your dataset contain personal data? 1	○ Yes     No	
6. Which of the usage licenses provided by EASY did you choose in order to comply with the access rights attached of the data?		•
INTEROPERABLE		i
7. Are the data in your dataset stored in preferred formats?		•
8. Do you link to other (meta)data and is this (meta)data online resolvable? $\ensuremath{\mathbb{G}}$		•
9. Did you provide contextual information about your dataset?	Persistent Identifier(s) Reference to other datasets	

- 1. Entrad en <a href="https://satifyd.dans.knaw.nl/">https://satifyd.dans.knaw.nl/</a>
- 2. CUMPLIMENTA el formulario y comprueba si tus datos son FAIR con el resultado obtenido.

Tiempo: 10 minutos.

#### ¿Cómo comprobar si sus datos son FAIR? Cualquier ha de poder encontrar y localizar sus datos. Los metadatos enriquecidos deben estar disponibles en línea en un recurso de búsqueda y los datos han tener asignado un identificador persistente. Sus datos tienen un identificador persistente. Los datos están descritos con metadatos enriquecidos. Los metadatos están en línea en un recurso de búsqueda, por ejemplo, en un catálogo o en repositorio El registro de metadatos especifica el identificador persistente. Ha de ser posible que tanto personas como máquinas tengan acceso a sus datos, bajo condiciones específicas o restricciones cuando sea oportuno. [FAIR no significa que los datos sean siempre abiertos! Sin embargo, los metadatos sí han de estar disponibles, independientemente de que los datos sean accesibles El ID persistente enlaza con los datos o metadatos asociados para que estos puedan ser recuperados. El protocolo mediante el cual se pueden recuperar los datos sigue estándares reconocidos, por ☐ El procedimiento de acceso incluye pasos de autenticación y autorización en caso de que sea necesario. Los metadatos serán accesibles, siempre que sea posible, aunque los datos no sean. Los datos y metadatos han de ajustarse a formatos y estándares reconocidos para permitir su combinación ☐ Los datos se proporcionan en formatos conocidos y usados ampliamente, y preferiblemente abiertos. Los metadatos proporcionados siguen estándares reconocidos. Siempre que sea posible se utilizarán vocabularios controlados, palabras clave, tesauros u ontologías. Se incluyen referencias y enlaces a otros datos relacionados. nttps://zenodo.org/record/5111307#.YpehpWhBxD8





UPV

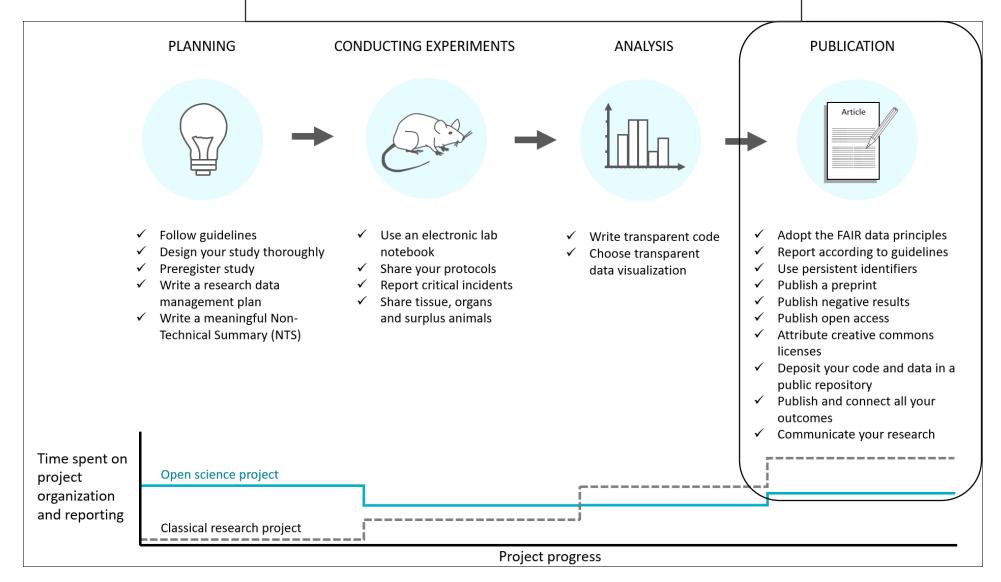
BIBLIOTECA Y DOCUMENTACIÓN CIENTÍFICA

Universitat Politècnica de València

www.upv.es · rienma@upvnet.upv.es



## Prácticas de ciencia abierta







## Estrategia de publicación y difusión de los datos

### PRINCIPIOS FAIR

- Elegir un repositorio FAIR para el depósito.
- No son válidas infraestructuras alternativas: web del proyecto, portal de difusión, redes sociales, Google Drive, etc.

# SELECCIONAR LOS DATOS

- Definir los datos del proyecto que serán públicos.
- Definir la persona responsable que recogerá los datos y metadatos.
- Comprobar requerimientos institucionales o de agencias financiadoras.

# ESTRATEGIA DE ACCESO ABIERTO.

- Definir cuándo se harán públicos en repositorios los conjuntos de datos.
- Cumplir con requerimientos institucionales y agencias financiadoras.

#### **ELEGIR LICENCIAS.**

- Clasifican los usos permitidos.
- Deben ser conocidas por los miembros del equipo.
- El software primero debe protegerse y luego asociarle una licencia.
- Los datos restringidos o confidenciales no deben publicarse ni llevar licencias.





INTEGRAR los datos en la publicación.

Integración datos y artículo.

Difíciles de encontrar, extensión limitada, no son FAIR, sin reutilización.



Datos adjuntos como MATERIAL SUPLEMENTARIO en un fichero.

Integración de datos y artículo, posibilidad de reutilizar.

Difíciles de encontrar, no son FAIR.



Depósito en REPOSITORIO.

Cumplimiento de mandatos de acceso abierto, FAIR, mayor extensión, posibilidad de autorías diferentes, preservación y difusión.



DATA PAPER.

Nueva publicación: valor añadido a los datos, mayor extensión, posibilidad de autorías diferentes.



### DISTINTIVOS DE PRÁCTICAS ABIERTAS A FAVOR DE LA REPRODUCIBILIDAD



Datos abiertos: datos necesarios para reproducir los resultados comunicados están disponibles públicamente.



Código Analítico Abierto: disponible el código analítico necesario para reproducir el análisis comunicado.



Materiales Abiertos: la metodología de investigación necesaria para reproducir el procedimiento (contribuciones a congresos, modelos, posters, software, encuestas..)



Papers: todos los manuscritos, documentos, informes y artículos que se generaron a partir del estudio.



Suplementos: cualquier otro recurso relacionado con un estudio que no se ajuste a ninguno de los elementos anteriores (p.e, planes de gestión de datos, diccionario de datos, figuras, eventos..).

## APA Journals offering open science badges

The following journals are now offering open science badges:

- American Psychologist
- American Journal of Orthopsychiatry
- Asian American Journal of Psychology
- Consulting Psychology Journal
- Canadian Journal of Experimental Psychology
- Cultural Diversity & Ethnic Minority Psychology
- Decision
- Journal of Applied Research in Memory Cognition
- Journal of Comparative Psychology
- Journal of Educational Psychology
- Journal of Psychopathology and Clinical Science
- <u>Journal of Experimental Psychology: Learning, Memory, and Cognition</u>
- Journal of Threat Assessment and Management

https://www.apa.org/pubs/journals/resources/open-science-badges



## Revistas con Sello.

#### Journals That Issue Open Science Badges\*

Addiction Research & Theory | Taylor & Francis

Advances in Archaeological Practice | Cambridge University Press

Advances in Methods and Practices in Psychological Science | SAGE

AIS Transactions on Replication Research | Elsevier

American Journal of Orthopsychiatry | APA

American Journal of Political Science | Wiley

American Journal of Primatology | Wiley

Analyses of Social Issues and Public Policy (ASAP) | Wiley

Annual Review of Applied Linguistics | Cambridge University Press

Archive for the Psychology of Religion | SAGE

Asian American Journal of Psychology | APA

Big Earth Data | Taylor & Francis

BMC Microbiology (uses modified badge criteria) | BMC

BMJ Open Science | BMJ

Brain and Neuroscience Advances | SAGE

Canadian Journal of Experimental Psychology (CJEP) | APA

Career Development and Transition for Exceptional Individuals | SAGE

Clinical Psychological Science | APS

Cognitive Science | Wiley

Communication Studies | Taylor & Francis

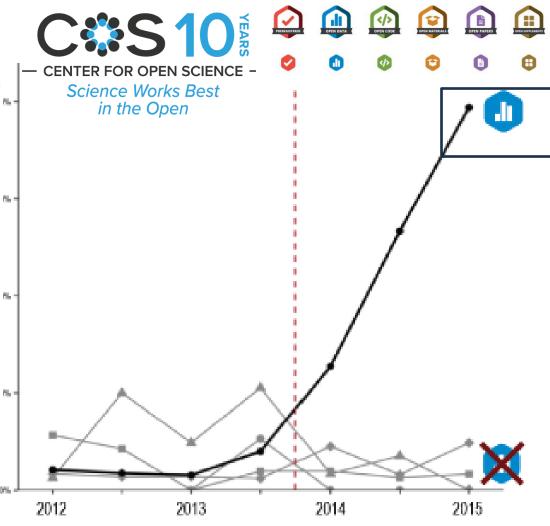
Communication Research Reports | Taylor & Francis

Cortex | Elsevier

Cultural Diversity & Ethnic Minority Psychology | APA

Decision I APA

Journal of Comparative Psychology | APA Journal of Experimental Psychology: Learning, Memory, and Cognition | APA Journal of Experimental Social Psychology | Elsevier Journal of International Crisis and Risk Communication Research | Nicholson School of Communication and Media Journal of Maps | Taylor & Francis Journal of Neuroendocrinology | Wiley Journal of Neurochemistry | Wiley Journal of Neuroscience Research (JNR) | Wiley Journal of Personality Assessment | Taylor & Francis Journal of Psychiatric and Mental Health Nursing | Wiley Journal of Social Psychology | Taylor & Francis Journal of Research in Personality | Elsevier Journal of Research on Educational Effectiveness | Taylor & Francis Journal of Threat Assessment and Management | APA Language Awareness | Taylor & Francis Language Learning | Wiley Language Testing | SAGE Law and Human Behavior | APA Management and Organization Review | Cambridge University Press Media Psychology | Taylor & Francis Meta-Psychology | Linnaeus University Press Neuropsychology | APA Neuroscience of Consciousness | Oxford University Press





## PUBLICAR LOS DATOS

"Tan abiertos como sea posible, tan cerrados como sea necesario" **EXCEPCIONES** (debidamente motivadas en el PGD):

- Protección derechos de propiedad industrial: proteger resultados si se prevé explotación comercial.
- Protección datos personales
- Confidencialidad en asuntos de seguridad
- El objetivo principal del proyecto se vea comprometido
- No se generen datos en el proyecto
- Otras razones a justificar

Se puede optar por no compartir datos en distintos momentos del proyecto.



# CITAR LOS DATASET

### ¿Por qué es importante citar los datos?

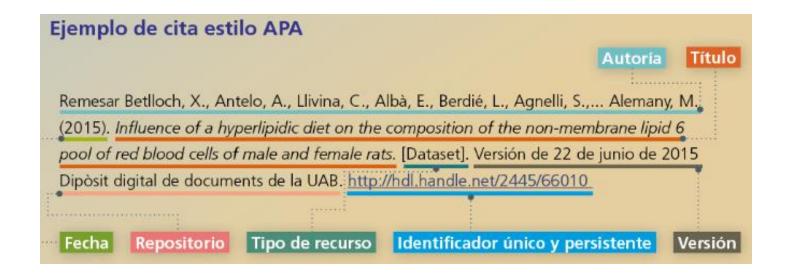
- son resultados de investigación.
- identificación y el acceso a los datos.
- reconocer autoría de sus creadores.
- métrica e impacto de los datos.
- transparencia de la investigación científica.
- incorporarse en el CV y en el registro ORCID.

### **DOI Citation Formatter**

Paste your DOI:	
For example 10.1145/2783446.2783605	
Select Formatting Style:	
ара	,
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.	1

### Recomendación de DataCite:

Creator (Publication Year). Title. [Dataset]. Version. Publisher. Identifier (DOI)



Format



## scientific data

Check for updates

# OPEN COVID-19 patient accounts of DATA DESCRIPTOR illness severity, treatments and lasting symptoms

Moriah E. Thomason (31,2,3 10), Denise Werchan 1,2 & Cassandra L. Hendrix 1

First-person accounts of COVID-19 illness and treatment can complement and enrich data derived from electronic medical or public health records. With patient-reported data, it is uniquely possible to ascertain in-depth contextual information as well as behavioral and emotional responses to illness. The Novel Coronavirus Illness Patient Report (NCIPR) dataset includes complete survey responses from 1,584 confirmed COVID-19 patients ages 18 to 98. NCIPR survey questions address symptoms, medical complications, home and hospital treatments, lasting effects, anxiety about illness, employment impacts, guarantine behaviors, vaccine-related behaviors and effects, and illness of other family/ household members. Additional questions address financial security, perceived discrimination, pandemic impacts (relationship, social, stress, sleep), health history, and coping strategies. Detailed patient reports of illness, environment, and psychosocial impact, proximal to timing of infection and considerate of demographic variation, is meaningful for understanding pandemic-related public health from the perspective of those that contracted the disease.

#### Background & Summary

Major discoveries about COVID-19 illness, susceptibility, transmission, and human behavior have been unearthed through utilization of rich medical and public digital record systems. Chasms in health inequity have been revealed1-3. Discrete spatiotemporal patterns of public health behavior have been characterized4-5 Individual- and community-level risk factors underlying local transmission of COVID-19 have been identified<sup>6.7</sup>, Examination of internet searches has revealed that (i) information flow about COVID-19 is inversely relates to positive cases\* and (ii) there have been population-level shifts over the course of the pandemic from searches pertaining to activity/fitness to more sedentary activities and dietary supplements\*. Overall, publicly available large-scale databases are significant sources of information that have been rapidly deployed to identify crucial determinants of health, aspects of transmission, and core human behavioral adaptations in the context

The challenge, however, is that these studies are limited by the constraints of electronic fields that serve narrow functions and lack nuance that is intrinsic of individual human stories. As a specific example, electronic medical records (EMR) systems include lists of symptoms, physical examination and laboratory results, treatments, diagnoses and basic demographic information. EMR are not intended to capture information about patient perceptions, and yet we know that the subjective experiences and outcomes of one person to the next are variable and also predictive of future health 10,51. Circumstances of illness occur in diverse socioemotional contexts and relate to other events occurring in an individual's life. The goal of NCIPR was to aggregate a sizable dataset of first-person accounts of COVID-19 illness, risk, and recovery. NCIPR survey data can be used to weave individual strands of history, environment, perspective, and health together to make new discoveries that will compliment and enrich knowledge about COVID-19 that has been derived from medical and public digital record systems. Further, the NCIPR dataset contains measures of lasting secondary effects of COVID-19 infection that are not readily available in medical record systems. In summary, the data available in this data set are different from many other publicly available data sets because these are self-reported patient data in a large sample with notable variability in illness severity, timing, and treatment.

<sup>1</sup>Department of Child and Adolescent Psychiatry, NewYork University Medical Center, NewYork, USA. <sup>2</sup>Department of Population Health, New York University Medical Center, New York, NY, USA. <sup>1</sup>Neuroscience Institute, NYU Langone Health, New York, NY, USA. Fe-mail: moriah.thomason@myulangone.org

Data Paper: A scholarly publication describing a particular dataset or collection of datasets and usually published in the form of a peer-reviewed article in a scholarly journal. The main purpose of a data paper is to provide facts about the data (metadata, such as data collection, access, features etc) rather than analysis and research in support of the data, as found in a conventional research article.

# Data Paper > Data Journal

- Describe los conjuntos de datos generados durante la investigación.
- Facilita la comprensión de los datasets.
- Fomenta el reconocimiento científico y académico de los autores.
- Proceso de *revisión por pares* que garantiza la calidad de los datos.



# Buena práctica: incluir en el PGD la previsión de publicar algún data paper.

Title	URL	Charge	Notes for authors (N.B. we suggest checkin submission of data already published)
Data			
Journal of Open Archaeology Data	http://openarchaeologydata.metajnl.com/	Article publication charge	http://openarchaeologydata.metajnl.com/abo
Open Health Data	http://openhealthdata.metajnl.com/	Article- processing charge	http://openhealthdata.metajnl.com/about/sub
Journal of Open Psychology Data	http://openpsychologydata.metajnl.com/	Article publication charge	http://openpsychologydata.metajnl.com/abou
Scientific Data	http://www.nature.com/scientificdata/	Article- processing charge	http://www.nature.com/sdata/for-authors http://www.nature.com/sdata/for-authors#dat
Data in Brief	https://www.journals.elsevier.com/data-in-brief	Article publishing charge	https://www.journals.elsevier.com/data-in-briedata-are-suitable-for-data-in-brief
Geoscience Data Journal	http://www.geosciencedata.com/ //www.wiki.ed.ac.uk/dis	splav/da	http://onlinelibrary.wiley.com/journal/10.1002, 6060/homepage/ForAuthors.html#dataset_sul

Listados de Data Papers

### Data@MLibrary

Updates from the Research Data Services Team at the University of Michigan Library



← Conference Highlights: Research Data Access & Preservation Summit 2014

### A Growing List of Data Journals

Posted on May 9, 2014 by RDS

by Katherine Akers\*

- RSS Posts
- RSS Comments

**®** Follow Data@MLibrary

Search

https://mlibrarydata.wordpress.com/2 014/05/09/data-journals/



# DATA JOURNALS.





**About the journal** <a href="https://www.sciencedirect.com/journal/data-in-brief">https://www.sciencedirect.com/journal/data-in-brief</a>

FAQs Data in Brief.

Data in Brief is a multidisciplinary, open access, peer-reviewed journal, which publishes short, digestible articles that describe and provide access to research data. It contributes to open science and improves reproducibility by making data and the associated research more ...

## Consejos:

- 1. En el caso de tratarse de <u>datos de poca entidad asociados</u> (tablas, ...), depositarlos en <u>DRYAD</u> o en <u>FIGSHARE</u> (según políticas editoriales).
- 2. En el caso de tratarse de <u>datos aun sin publicación o derivados de</u> <u>financiación</u> europea o nacional con <u>mandatos</u> de depósito, escoger repositorio temático, institucional <u>Riunet</u> o <u>Zenodo</u>.
- 3. Intentar publicar un <u>data paper en revista de datos</u>: el dataset será revisado por pares recibiendo un valor añadido.

View all

Acknowledgements



### DATA JOURNALS.

### A synthesis of hydroclimatic, ecological, and socioeconomic data for transdisciplinary research in the Mekong

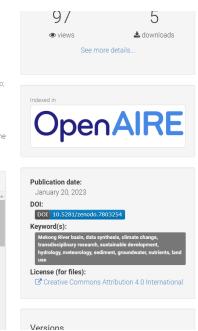


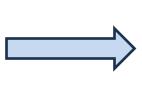
(b) Tiwari, Amar Deep

Various climate, hydro-meteorological, ecological, and socio-economic datasets are synthesized and made available for the Mekong River Basin. The sources of each dataset are also mentioned in the associated readme file.

Dam attribute data, inundation data, and Cambodia census data can be made available upon request to the authors.









## Published January 20, 2023 | Version v3

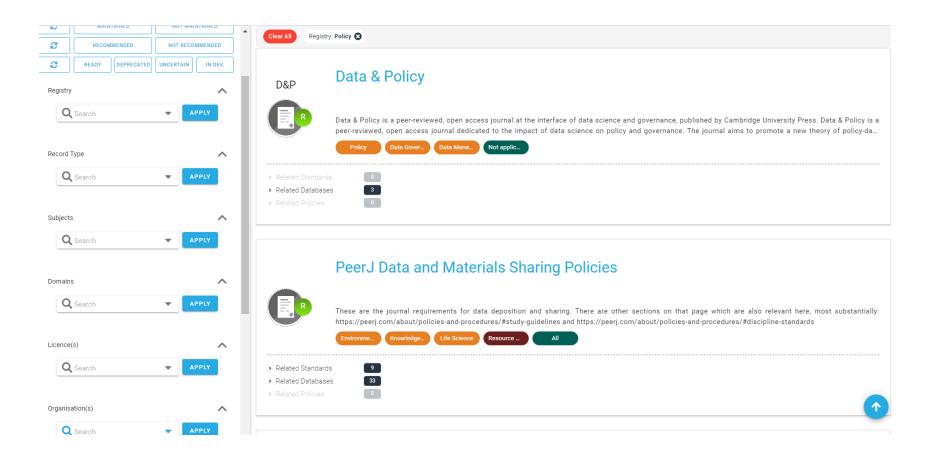
https://zenodo.org/record/780325 4#.ZGJ0AnZBwtx

Published: 15 May 2023

https://www.nature.com/articles/s41597-023-02193-0#Sec23



### CÓMO BUSCAR UN DATA JOURNAL







### Política editorial del Grupo PLOS

- Definición mínima de dataset: "the data required to replicate all study findings reported in the article, as well as related metadata and methods".
- No se aceptan publicaciones sin datos: "PLOS journals require authors to make all data necessary to replicate their study's findings publicly available without restriction at the time of publication".

Closing the gender gap at academic conferences: A tool for monitoring and assessing academic events

#### Table 4

#### Focus group participant

	Knowledge area	Profile / Expertise	
1	Economy	Research and innovation evaluation. Working for think tanks and government administration. Leading of multi-country European projects.	
2	Political Sciences	Gender and Corporate Social Responsibility.	
		Researcher at a European Horizon 2020 project and Member of the gender commission of the Spanish observatory 'Mujeres, Ciencia e Innovación' (Women, Science and Innovation) for gender equality.	
3	Chemistry	Public policies for scientific research and technological innovation, and the interactions between science and innovation. Extensive experience in management positions in various Spanish R&D institutions and leader of multiple national and international projects.	
4	Psychology	Bibliometrics and cyber-metrics, databases, humanities information, information searching and retrieval, and scientific-technical information.	
5	History	History of Science, history of medicine and cultural history.	
		One of the first women to lead an inter-university research institute in Spain.	
6	Sociology	Gender studies in the area of women and equality policies. Feminist activist.	
7	Engineering	Findingering and project management evaluation of competitiveness husiness	



https://doi.org/10.1371/journal.pone.0243549.t004





**UPV** 

BIBLIOTECA Y DOCUMENTACIÓN CIENTÍFICA

Universitat Politècnica de València

www.upv.es · rienma@upvnet.upv.es



Comprobar si hay un repositorio temático de datos en el ámbito

Utilizar el repositorio institucional

(RiuNet)

Utilizar un repositorio de datos multidisciplinar

- Dataverse
- Dryad
- <u>Figshare</u>
- ► <u>IEEE Dataport</u>
- Mendeley Data
- Open Science Framework
- Synapse
- Vivli
- Zenodo

BUSCAMOS EN re3data

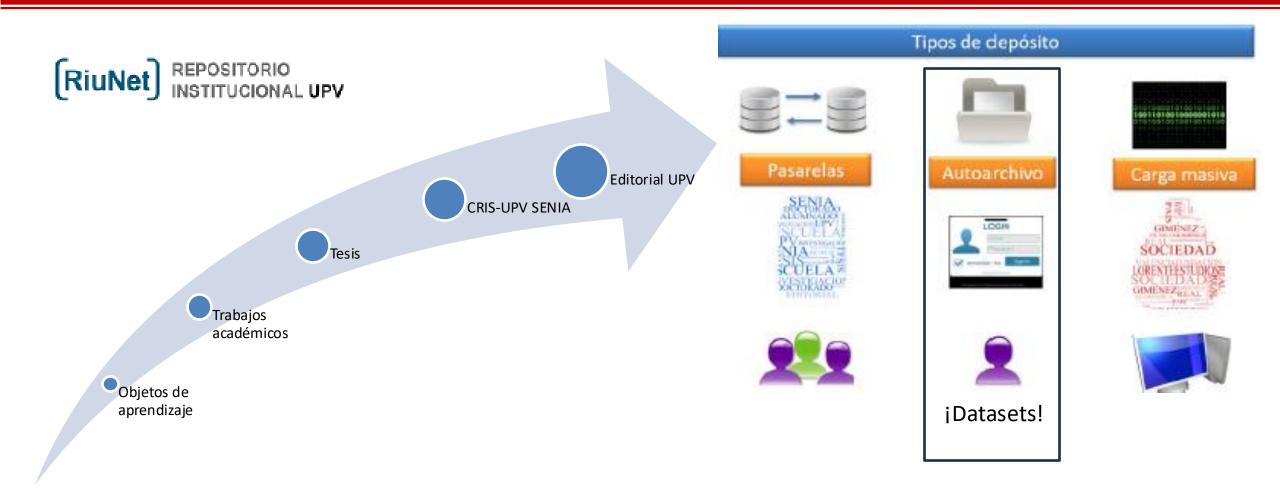
Recomendaciones para seleccionar un repositorio donde depositar datos de investigación del CSUC

# DEPÓSITO EN REPOSITORIO TEMÁTICO.

- ·¿Área?
- ¿Capacidad de almacenamiento?
- ¿Facilidad de recuperación de datos?
- ¿Identificador único y persistente?
- ¿Selección de la licencia?
- ¿Preservación a largo plazo?
- ¿Certificación <u>CoreTrustSeal?</u>
- Introducir.
  - Afiliación: Universitat Politècnica de València
  - Firma científica: NBU.



# **DEPOSITAR DATOS**







Cada repositorio tiene una política, gobernanza y servicios a los usuarios.



Conviene comprobar los tipos de datos, dimensiones y ficheros que soporta.



Conviene comprobar qué PIDs asigna y si hay algún coste asociado.



Cada repositorio describe los recursos según esquemas de metadatos.



Puede ofrecer a los usuarios plantillas de descripción README u otros requisitos de documentación específicos.





Biblioteca y Documentación Científica

# Política de RiuNet

25 de octubre de 2021

Politics de contanidos

Politica de metadatos

Politica de dapósito

Politica de formatos

Politica de preservación y edición de registros

#### Política de contenidos

Riutiet albeiga la prosoción académica de la UPV, tans cientifica some docente, estructurándose en suatro grandes ámbitos. Docencia, investigación, lestitucional y Patrimonial.

Su objetivo es diable, de una parte difundir la producción de la Universidad, en eles sertado apuedas radicialmente por las certamentes apopalistes. Poro Ristrict terrader comple sina farcidad de accisión, de ablique en distaminador conocionnes abintiles contresidos cercentades supre organistes dise será entides a randicio encidados. Una situación intermedia es la formidad por contresidos embalgados, que pertenencien en certado por un período de terrago apendinados partir de usa facilida dada.

La apuesta por el acceso abiento a la producción inbélecidade la institución i manifestada en la <u>política (le Acceso Albertoch III UP</u>), aprinciada en 21 de Judi de 2011, la comorta pasa les autores, dependendo de las colecciones, en un mandado un insignidad colecciones.

Colección	Política, tipo de obligación		
Objetos de aprendizaje: Polimedias, laboratorios virtuales y artículos docentes	Recomendación incentivada. Se gratifican económicamente los objetos depositados en abierto.		
Trabajos académicos	Recomendación. La decisión es del alumno.		
Revistas UPV, Editorial UPV	,		
Tesis doctorales:	Mandato. (Real Decreto 99/2011, de 28 de enero, por el que se regulan las enseñanzas oficiales de doctorado)		
Otras colecciones de investigación	Ley de la ciencia. Política de la editorial. Los registros incluidos en estas colecciones corresponden a publicaciones, se sigue la política de la editorial. La Universidad recomienda a los autores UPV que firmen con las editoriales acuerdos "sin exclusividad" reservando parte de sus derechos de explotación.		
Ámbito institucional	oito institucional Por defecto mandato. Puede haber situaciones especiales		
Ámbito patrimonial	to patrimonial Por defecto mandato. Puede haber situaciones especiales		
Datasets	/ opcional.		



# **COLECCIÓN DATASETS**

# CONDICIONES DE

**ACEPTACIÓN** 

- ✓ Los datos deben haber sido producidos:
  - Dentro de proyectos en los que participen investigadores de la UPV.
  - Entidades y/o servicios de la UPV.
- ✓ Los autores de los datos deben estar en condiciones de conceder los derechos necesarios a la UPV para asegurar la correcta distribución y preservación de los datos a través de RiuNet.
- ✓ Si el conjunto de datos contiene **datos personale**s, se debe tener en cuenta lo expresado en la normativa nacional e internacional sobre protección de datos personales.
- ✓ Los datos deben estar debidamente organizados para facilitar su comprensión y reutilización. Para ello, se seguirán las recomendaciones habituales sobre nombre y formatos de los ficheros, así como la documentación.
- ✓ Aunque se pueden depositar varios ficheros con datos de investigación, cada uno de los ficheros no debe superar los 2 GB de tamaño.



# **DEPÓSITO EN ZENODO.**



Plataforma Online

Repositorio a largo plazo

Privados (no publicados) o Públicos (publicados)

Requiere un admin

50 GB por carga

Universitat Politècnica de València (UPV). Research Data.

View Curate

Datasets collection from Universitat Politècnica de València researchers, managed by Documentación científica at University Library.

Curated by: Documentacion\_Cientifica\_Biblioteca

Depositad los datasets dentro de la Comunidad, los hará más accesibles y localizables.

### **Publication date:**

May 10, 2023

DOI:

DOI 10.5281/zenodo.7942786

Communities:

Gene Ontology Zenodo

License (for files):

Versions

☑ Creative Commons Attribution 4.0 International

### Version 2023-05-10 May 10, 2023 10.5281/zenodo.7942786

10.5281/zenodo.7796232

Apr 1, 2023

Jul 2, 2018

Jul 2, 2018

Version 2023-03-06 Mar 6, 2023 10.5281/zenodo.7709866

Version 2023-01-01 10.5281/zenodo.7504797

Version 2022-12-04 10.5281/zenodo.7407024

Version 2023-04-01

View all 55 versions

Cite all versions? You can cite all versions by using the DOI 10.5281/zenodo.1205166. This DOI represents all versions, and will always resolve to the latest one. Read more.



# DEPÓSITO EN ZENODO.

- Seguridad y confianza.
- DOI.
- Modo restringido.
- Versionado.
- Integración con GitHub.
- Estadísticas de uso.
- Recogido en OpenAIRE.
- Autoarchivo por parte del personal investigador.



### Guías y Ayudas

#### MANUAL DE DEPÓSITO DE DATASETS EN ZENODO.

Fecha última actualización: 07/06/22 Tiempo aprox. de consulta: 10 min.

¿Necesitas conocimientos previos?: Sí. Si te ha interesado, puede resultarte útil:

 Biblioguía de <u>Datos de investigación</u> de la Biblioteca de la Universitat Politècnica de Valencia.

Autor/es de la guía: Documentación científica.

- Qué es ZENODO.
- Por qué usar ZENODO.
- Procedimiento para depositar datasets en la comunidad UPV de Zenodo.



Esta obra se comparte bajo la licencia: Creative Commons Reconocimiento-NoComercial-Compartir Igual

#### Qué es Zenodo.

Zenodo es un repositorio de datos de investigación en acceso abierto gestionado por el CERN. Forma parte de OPENAIRE y está financiado por la Unión Europea.

### GUÍA DE USO:

http://hdl.handle.net/10251/190726



# ADHESIÓN DE RIUNET A LOS PRINCIPIOS FAIR.

Protocolos de interoperabilidad abiertos y gratuitos (OAI-PMH).	
Esquema de metadatos estándar (DC).	
Plantilla de descripción.	
Asignación de PIDs (Handle, DOIs).	
Indexación de agregadores.	
Servicio de apoyo y asesoramiento.	
Presencia en re3data.	



# **METADATOS FAIR.**

<u>ORCID</u> (dc.contributor.orcid): Identificador ORCID para cada autor, campo normalizado; Apellidos, Nombre [ORCID].

### Metadatos del ítem

Titulo: Surface modification of jute fabrics by rGO-ICP for eco-friendly supercapacitors, Journal of Energy Storage

2023

Otro titulo: Surface modification of jute fabrics by reduced graphene oxide-conducting polymer coatings for their

application in low-cost and eco-friendly supercapacitors

Autor: Cases Iborra, Francisco Javier ( Bonastre Cano, José Antonio ( Molina Puerto,

<u>Javier</u>

<u>Título.</u> (dc.title): El título del artículo y del dataset asociado no deberían coincidir, si lo hacen es necesario identificarlos de forma diferenciada con aclaración [Dataset] o frase introductoria "Supplementary material for: ..."

# Coronavirus en YouTube [Dataset]

Orduña Malea, E.; Font Julián, Cl.; Ontalba Ruipérez, JA. (2020). Coronavirus en YouTube [Dataset]. Universitat Politècnica de València. https://doi.org/10.4995/Dataset/10251/143671

Por favor, use este identificador para citar o enlazar este ítem: http://hdl.handle.net/10251/143671

### Ficheros en el ítem



Referencia bibliográfica del conjunto de datos (dc.identifier.citation):

Aplicamos la recomendación de FORCE11: Campo con una sintaxis determinada y todos los elementos separados por punto y coma: Autor 1; Autor2; Autor 3; Año; Título; Repositorio; Versión; DOI

### OVS-SCM Modulation (raw data)

Pérez Pascual, MA. (2020). OVS-SCM Modulation (raw data); Riunet; https://doi.org/10.4995/Dataset/10251/202149

Por favor, use este identificador para citar o enlazar este ítem: http://hdl.handle.net/10251/202149

### *DOI* (dc.identifier.doi):

- Imprescindible para conjuntos de datos ya presentes en otros repositorios de datos (el DOI identifica de forma unívoca el conjunto de datos. Un mismo dataset no debe tener varios DOI asociados).
- En conjuntos de datos publicados junto a los artículos asociados, el DOI será el mismo que el del artículo junto con el que se publica.
- RIUNET asigna DOIs a conjuntos de datos que se depositan en su infraestructura (no están en ninguna otra), en estos casos, este metadato se genera automáticamente.



# **METADATOS FAIR.**

### <u>Tipología</u> (dc.type) (Vocabulario COAR).

• Tipologías que generan DOI: Dataset, tesis, preprint, software, PGD.

### Palabras clave (dc.subject).

• Recomendado usar entradas procedentes de un vocabulario controlado. El uso de vocabularios controlados de este tipo es una de las principales recomendaciones para cumplir con los Principios FAIR.

59	dc.subject.ddc	Dewey Decimal Classification Number	
505	dc.subject.editorial	Materia de la editorial	
60	dc.subject.lcc	Library of Congress Classification Number	
61	<u>dc.subject.lcsh</u>	Library of Congress Subject Headings	
62	dc.subject.mesh	MEdical Subject Headings	
725	dc.subject.ods	Asignación en Senia de documento a alguno de los ods	
63	<u>dc.subject.other</u>	Titulacion del proyecto-para las colecciones PFCs	
221	dc.subject.unesco	Nomenclatura Internacional de la UNESCO para los campos de Ciencia y Tecnología	



### Entidades financiadoras (dc.contributor.funder).

• Las organizaciones o agencias que financiaron el dataset (pueden o no coincidir con los de la publicación asociada si la hay).

### Referencia del proyecto (dc.relation.projectID).

 Campo con una sintaxis determinada para trabajos subvencionados. Hay que utilizar la API FECYT: <a href="https://buscador.recolecta.fecyt.es/api/rest/proyectos?codigo="https://buscador.recolecta.fecyt.es/api/recolecta.fecyt.es/api/recolecta.fecyt.es/api/recolecta.fecyt.es/api/recolecta.fecyt.es/api/recolecta.fecyt.es/api/recolecta.fecyt.es/api/recolecta.fecyt.es/api/recolecta.fecyt.es/api/recolecta.fecyt.es/api/recolecta.fecyt.es/api/recolecta.fecyt.es/api/recolecta.fecyt.es/api/recolecta.fecyt.es/api/recolecta.fecyt.es/api/recolecta.fecyt.es/api/recolecta.fecyt.es/api/recolect

### Código del Proyecto:

info:eu-repo/grantAgreement/ //FPU19%2F01759//AYUDA PREDOCTORAL FPU-IVORRA MARTINEZ. PROYECTO: PROCESADO Y OPTIMIZACIÓN DE MATERIALES AVANZADOS DERIVADOS DE ESTRUCTURAS PROTEICAS Y COMPONENTES LIGNOCELULÓSICOS/ info:eu-repo/grantAgreement/ //FPU20%2F01732//Desarrollo y caracterización de materiales de alto rendimiento medioambiental derivados de residuos de la industria de los cítricos bajo un planteamiento de economía circular y desarrollo sostenible/ info:eu-repo/grantAgreement/GVA//AICO%2F2021%2F025//NUEVOS MATERIALES SOSTENIBLES DERIVADOS DE LA REVALORIZACIÓN DE RESIDUOS DE LA INDUSTRIA DE CÍTRICOS ¿ CitricNewLife (GENERALITAT VALENCIANA)/



<u>Derechos</u> (dc.rights): abierto, embargado, cerrado.

Licencia (dc.rights.licens): condiciones de uso y reutilización, evitar usos indebidos.

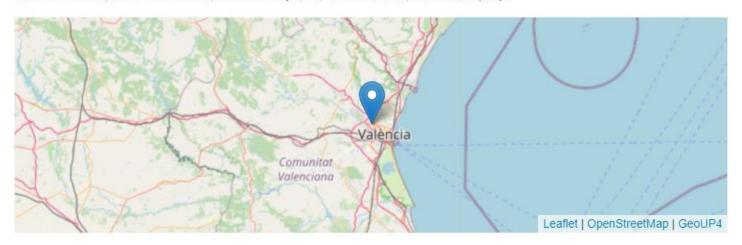
Tipos de autoría: dc.contributor.autor, dc.contributor, ...

Cobertura geográfica (dc. coverage. Spatial):

- Hace referencia al lugar de recogida de datos o al lugar al que hacen referencia los datos
- Es muy recomendable, ya que es una información contextual significativa para muchos datasets y la base para realizar análisis y visualizaciones de geolocalización.
- Se recomienda consultar las entradas normalizadas de vocabularios controlados.

### Localización

east=-0.4431618; north=39.5037093; name=C/ Major, 50, 46980 Paterna, Valencia, Espanya



Other acceptable formats

SPSS portable format (.por).

MS Access (.mdb/.accdb).



### *Formatos (IANA)* (dc.format)

- Una buena práctica es ofrecer los datos en varios formatos, aquéllos que se usan más en una disciplina específica y después en un formato abierto
- Escogeremos el formato adecuado del siguiente listado de IANA (Internet Assigned Numbers Authority) en la columna "Template".

### FORMATOS RECOMENDADOS:

- https://ukdataservice.ac.uk/learning-hub/research-data-management/format-yourdata/recommended-formats/
- https://www.loc.gov/preservation/resources/rfs/TOC.html

<metadata></metadata>	Recommended formats				
	Recommended formats				
<dc:title>Video Educativo sobre Matemáticas</dc:title>	Proprietary formats of statistical packages e.g. SPSS (.sav),				
<dc:creator>Juan Pérez</dc:creator>	Stata (.dta), .sas7bdat.				
<dc:format>video/mp4</dc:format>	Delimited text and command ('setup') file (SPSS, Stata, SAS,				
	etc.) containing metadata information.				
<dc:format>1920 x 1080 pixels</dc:format>	Some structured text or mark-up file containing metadata				
<dc:format>500 MB</dc:format>	information, e.g. DDI XML file.				
<dc:format>01:30:00</dc:format>					
<dc:date>2024-05-27</dc:date>	<dc:date>2024-05-27</dc:date>				
<dc:description>Un video educativo que explica conceptos básicos de matemáticas.</dc:description>					





# Media Types

### Last Updated

2024-05-24

### Registration Procedure(s)

Expert Review for Vendor and Personal Trees

### Expert(s)

Alexey Melnikov, Darrel Miller, Murray Kucherawy (backup)

#### Reference

[RFC6838][RFC4855]

#### Note

Per Section 3.1 of [RFC6838], Standards Tree requests made through IETF documents will be reviewed and approved by the IESG, while requests made by other recognized standards organizations will be reviewed by the Designated Expert in accordance with the Specification Required policy. IANA will verify that this organization is recognized as a standards organization by the IESG.

#### Note

 $[ ext{RFC2046}]$  specifies that Media Types (formerly known as MIME types) and Media Subtypes will be assigned and listed by the IANA.

Procedures for registering Media Types can be found in [RFC6838], [RFC4289], and [RFC6657]. Additional procedures for registering media types for transfer via Real-time Transport Protocol (RTP) can be found in [RFC4855].

The following is the list of Directories of Content Types and Subtypes. If you wish to register a Media Type with the IANA, please see the following for the online application:

[Application for registration of Media Types]

### Registries included below

- application
- audio
- font
- example
- <u>image</u>
- message
- model
- multipart
- text
- video

# application

### **Available Formats**

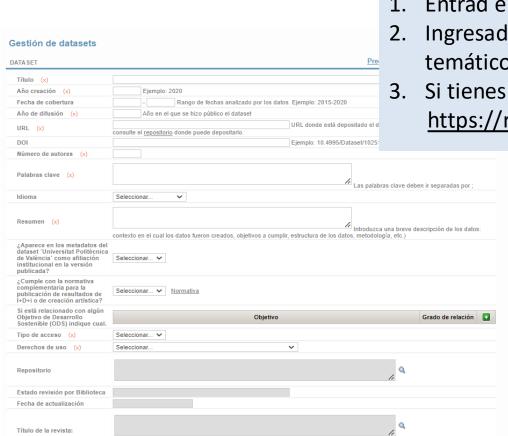


CSV

https://www.iana.org/assignments/mediatypes/media-types.xhtml



# PRÁCTICA: ingresa los metadatos de un dataset en SENIA



- 1. Entrad en https://intranet.upv.es/pls/soalu/est intranet.ni Dual?p cua=senia
- 2. Ingresad los metadatos de un DATASET que hayas depositado en algún repositorio temático (o inventa uno de prueba).
- 3. Si tienes datos y quieres depositarlos acude a la colección "Datasets": https://riunet.upv.es/handle/10251/55048



- POLÍTICA DE INTEGRIDAD CIENTÍFICA Y BUENAS PRÁCTICAS (2021).
- Servicios de <u>ALMACENAMIENTO</u>: Unidades de red, permisos y cuotas
- Repositorios: <u>Riunet</u>, <u>Zenodo</u>.
- Metadatos en <u>Senia</u>.

### Materiales de formación:

Biblioguía Datos de investigación

Biblioguía Riunet

Plan de Gestión de Datos de investigación UPV: guía

Depositar en ZENODO.

ARGOS, guía de uso.

8 Recomendaciones para tus datos de investigación.



# GRACIAS POR LA ATENCIÓN

MMM. 116 A. G.S.