

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

5-8 de junio de 2023

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

Tutoría 13 de junio a las 12 h

**Remedios Melero. IATA-CSIC, Valencia, España**

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

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

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

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

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

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

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

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

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

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

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

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

Depurar los datos  
Documentar los datos  
Anonimizar (si corresponde)  
Licenciar  
p.e.

<https://openrefine.org/>

Readme.txt

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

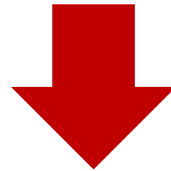
Licencias Creative Commons,

Enriquecer  
estandarizados

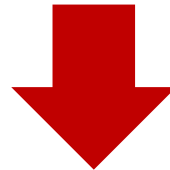
Vocabularios/tesaurus



**Elige repositorio**



**Deposita + Añade metadatos**



**Distribuye en abierto**



**Cita los datos**

# Dónde localizar y depositar datasets

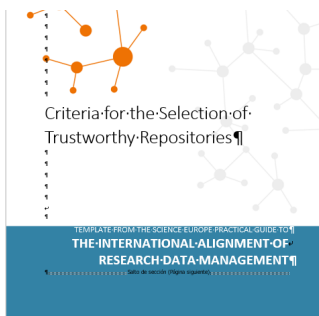


## Criteria for the Selection of Trustworthy Repositories

### “fiables”

#### Ofrecen identificadores persistentes y únicos (PID)

- Permitir descubrir e identificar los datos
- Permitir la búsqueda, citación y recuperación de datos
- Proporcionar soporte para el versionado de datosMetadatos
- Permitir encontrar los datos
- Permitir la referencia a información relevante relacionada, como otros datos y publicaciones
- Proporcionar información que esté disponible y se mantenga públicamente, incluso para los datos no publicados, protegidos, retractados o eliminados.
- Utilizar normas de metadatos ampliamente aceptadas (por la comunidad científica)
- Garantizar que los metadatos sean recuperables por las máquinas.



## **Licencias de acceso y uso de datos**

- Permitir el acceso a los datos en condiciones bien especificadas
- Garantizar la autenticidad e integridad de los datos
- Permitir la recuperación de los datos
- Proporcionar información sobre licencias y permisos (idealmente en formato legible por máquina)
- Garantizar la confidencialidad y respetar los derechos de los sujetos y creadores de los datos

## **Preservación**

- Garantizar la persistencia de los metadatos y los datos
- Ser transparente en cuanto a la misión, el alcance, las políticas de preservación y los planes (incluyendo la gobernanza, la sostenibilidad financiera, el período de conservación y el plan de continuidad).

Metrics 27,132 Downloads [Contact](#) [Share](#)

Repositorio de Datos del Consorcio Madroño | ([www.consorciomadrono.es](http://www.consorciomadrono.es))



Search this dataverse... [Find](#) [Advanced Search](#)

[Dataverses \(13\)](#)   **1 to 10 of 13 Results**   [Sort ▾](#)

[Datasets \(531\)](#)

[Files \(3,100\)](#)

[Catálogos de Barrios Vulnerables de España 1991, 2001 y 2011](#) (vps181.cesvima.upm.es/re-hab/bbvv) [🔗](#)

Jan 14, 2021 [Repositorio de Datos UPM](#)

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

267.9 ms x3

 **Aviso:** Ha cerrado la sesión satisfactoriamente.

## Bienvenido a PGDOnline

PGDOnline ha sido desarrollado por el **Consorcio Madroño** como una herramienta para elaborar planes de gestión de datos.

### Comenzando:

- [Consorcio Madroño](#)
- [INVESTIGAM - Portal de Ciencia Abierta del Consorcio Madroño](#)
- [INVESTIGAM: PAGODA](#)
- [INVESTIGAM — PAGODA: CREAR SU PGD](#)
- [DCC Checklist for a Data Management Plan](#)

[Iniciar sesión](#)   [Registrarse](#)

\* **Correo electrónico**

\* **Clave**

[¿Olvidó su clave?](#)

Recordar correo

[Iniciar sesión](#)

**CORA** Repositori de Dades de Recerca

Home About Search User Guide Contact English Sign Up Log In

Search this dataverse...

26 UNIVERSITIES & CERCA CENTERS    117 DATASETS    1,454 FILES    8,347 DOWNLOADS



- Datasets (46)
  - Datasets (117)
  - Files (1,454)
- Dataverse Category**
- CERCA Centers (17)
  - Universities (11)
  - Research Group (9)
  - Research Project (4)

1 to 10 of 163 Results

Sort



Data of Fecal Immunochemical Test-Based Colorectal Cancer Screening Program According to Ambient Temperature and Humidity  
 Mar 17, 2022 - Universitat de Barcelona

Ibáñez Sanz, Gemma; Milà, Núria; Vives, Núria; Vidal Lancis, Carmen; Binefa i Rodríguez, Gemma; Rocamora, Judith; Atencia, Carmen; Moreno Aguado, Víctor; Sanz Pamplona, Rebeca; García Martínez, Montserrat, 2022, "Data of Fecal Immunochemical Test-Based Colorectal Cancer Screening Program According to Ambient Temperature and Humidity", <https://doi.org/10.34810/data168>, Repositori de Dades de Recerca, V1, UNF:6:peH5fEDC6jc9Nbk3xoFmg== [fileUNF]

Data from a retrospective cohort that included individuals aged 50–69 years who participated in colorectal cancer (CRC)

## Welcome

This tool helps you to create, review, and share data management plans that meet institutional and funder requirements. Join the growing number of researchers that have adopted **eiNa DMP**:



509 Users



306 Plans



15 Organisations



10 simple rules for creating a good data management plan



## LISTAR

### Todo Repisalud

► Comunidades y Colecciones

► Por fecha de publicación

► Autores

► Títulos

► Tipo de documento

► Institución

► Palabras clave

► MeSH

► Agencias financiadoras

### Esta comunidad

► Por fecha de publicación

► Autores

► Títulos

► Tipo de documento

► Institución

[Inicio](#) | [Sobre Repisalud](#) | [Info autores](#) | [FAQs](#) | [Login](#) | [Contacto/Sugerencias](#)

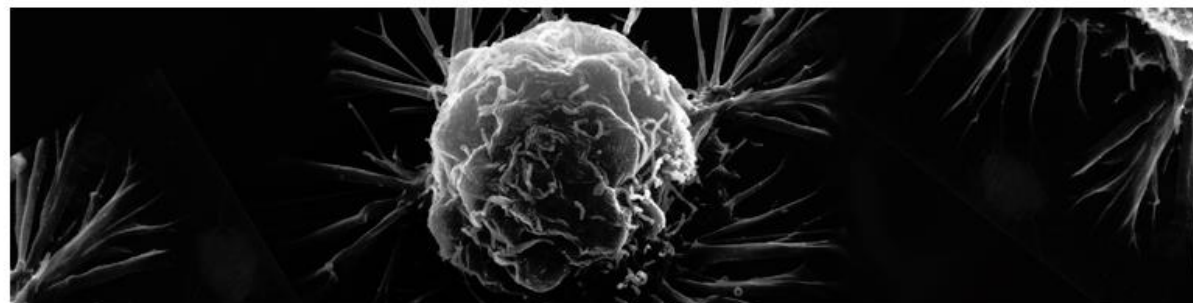
[español](#)

Q Buscar

? **Búsqueda avanzada**

[Repisalud Principal](#) / [Investigación](#)

## Investigación



Artículos, contribuciones a congresos, datos de investigación, libros, capítulos de libro, patentes, tesis y trabajos de investigación.

### Subcomunidades en esta comunidad

**ISCIII** [3452]

Instituto de Salud Carlos III

**CNIC** [1335]

Centro Nacional de Investigaciones Cardiovasculares

**CNIO** [453]

Centro Nacional de Investigaciones Oncológicas

**IIS** [476]

Institutos de Investigación Sanitaria



Comunidad de Madrid | CONSEJERÍA DE SANIDAD

## REPOSITORIO INSTITUCIONAL

DE LA CONSEJERÍA DE SANIDAD DE LA COMUNIDAD DE MADRID

[Inicio](#) | [Sobre el Repositorio](#) | [Ayuda](#) | [Acceso](#) | [Contacto](#)

[ES](#) [EN](#)

Buscar...

[Búsqueda avanzada >>](#)

### Navegar

- ▶ Comunidades/ Colecciones
- ▶ Fecha Publicación
- ▶ Autor
- ▶ Título
- ▶ Palabra clave

### Tipos de documentos

- Artículos
- Congresos
- Datos de Investigación
- Divulgación
- Formación y Docencia
- Informes técnicos
- Libros
- Multimedia

## Hospital Universitario Puerta de Hierro Majadahonda : [103]

### Envíos recientes

- » Plasma IP-10 and IL-6 are linked to Child-Pugh B cirrhosis in patients with advanced HCV-related cirrhosis: a cross-sectional study
- » Risk Factors and Mortality of COVID-19 in Patients With Lymphoma: A Multicenter Study
- » Title: Chess as a cognitive enhancement tool in the healthy and in mental health disorders: A systematic review.

Buscar

[Fecha Publicación](#) [Autor](#) [Título](#) [Palabra clave](#) [Tipos de documento](#)

### Colecciones en esta comunidad

[Artículos \[103\]](#)

[Comunicaciones a congresos \[0\]](#)

[Datos de Investigación \[0\]](#)

# Repositorio “generalista”

<https://zenodo.org/>

zenodo Search Upload Communities rmelero@lata.csic.es

Delete Save Publish

## New upload

Instructions: (i) Upload minimum one file or fill-in required fields (marked with a red star). (ii) Press "Save" to save your upload for editing later. (iii) When ready, press "Publish" to finalize and make your upload public.

Files Choose files Start upload

Drag and drop files here

— or —

Choose files

(minimum 1 file required, max 50 GB per dataset - [contact us](#) for larger datasets)  
If you're experiencing issues with uploading larger files, read our [FAQ section](#) on file upload issues.

Communities recommended

Specify communities which you wish your upload to appear in. The owner of the community will be notified, and can either accept or reject your request. Please make sure your record complies with the content policy of the communities you add; reported abuse will be followed by account inactivation.

Start typing a community name...

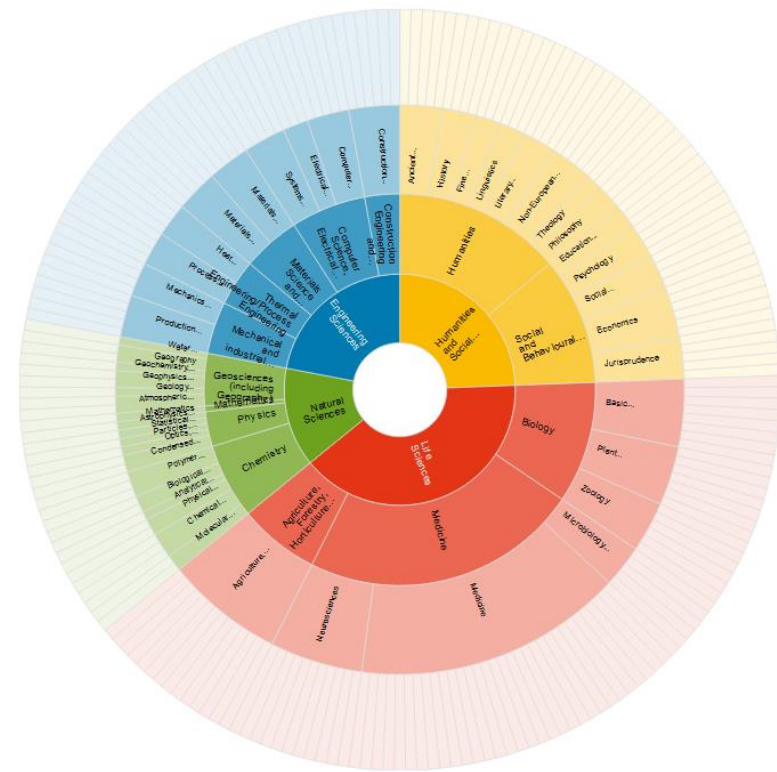
Upload type required

Publication Poster Presentation Dataset Image Video/Audio Software Lesson Physical object Other

Publication type Journal article

Publication type.

<http://www.re3data.org/>



**German Research Foundation to fund new services of re3data**

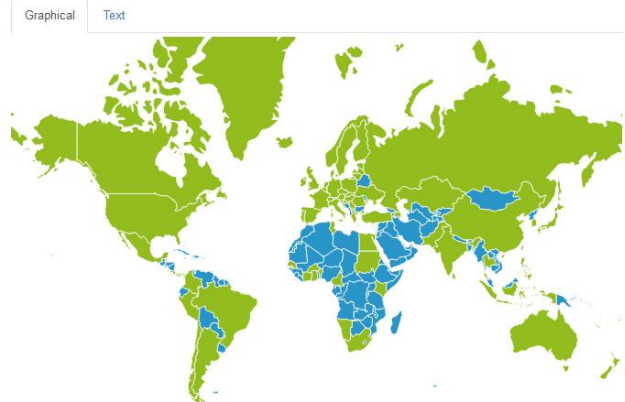
Since its launch in 2012, re3data has become the resource of information about research data repositories. It indexes and provides extensive information about more

**Data sharing made easier: use Repository Finder to find the right repository for your data**

More and more funders and publishers require research data to be made available in appropriate repositories, but

**2,000 Data Repository and Science Europe Framework for Discipline-specific Research Data Management**

**Browse by country**



**Filter**  
Reset all

**Subjects**

- Humanities and Social Sciences (1)
- Life Sciences (8)
  - Biology (5)
    - Basic Biological and Medical Research (2)
    - Bioinformatics and Theoretical Biology (1)
  - Plant Sciences (5)
    - Plant Systematics and Evolution (1)
    - Plant Biochemistry and Biophysics (1)
    - Plant Genetics (2)
- Medicine (1)
  - Nutritional Sciences (1)
- Agriculture, Forestry, Horticulture and Veterinary Medicine (8)
  - Agriculture, Forestry, Horticulture and Veterinary Medicine (8)
  - Soil Sciences (3)
    - Plant Cultivation (8)
    - Plant Nutrition (3)
    - Ecology of Agricultural Landscapes (2)
    - Plant Breeding (5)
    - Agricultural and Food Process Engineering (1)
    - Agricultural Economics and Sociology (3)
    - Inventory Control and Use of Forest Resources (1)
    - Basic Forest Research (1)
- Natural Sciences (1)
  - Chemistry (1)
    - Biological Chemistry and Food Chemistry (1)
    - Biological and Biomimetic Chemistry (1)
- Engineering Sciences (1)
  - Computer Science, Electrical and System Engineering (1)
  - Computer Science (1)
    - Operating, Communication and Information Systems (1)

**Content Types**

**Countries**

**AID systems**

**API**

**Data access**

**Data access restrictions**

Search...

← Previous 1 Next →

Toggle short help

Sort by ▾

Found 8 result(s)

**International Institute of Tropical Agriculture datasets**  
International Institute of Tropical Agriculture CKAN repository

**Subject(s)**  
Agriculture, Forestry, Horticulture and Veterinary Medicine | Life Sciences | Plant Breeding | Plant Cultivation | Plant Nutrition | Plant Genetics | Humanities and Social Sciences  
Agriculture, Forestry, Horticulture and Veterinary Medicine | Plant Sciences | Biology

**Content type(s)**  
Standard office documents | Databases | Scientific and statistical data formats | Structured text | other

**Country**  
International

IIITA conducts research on the following thematic areas: Biotechnology and genetic improvement, Natural resource management, Social science and agribusiness, and Plant production and plant health.

**CSISA Data Repository**  
Cereal Systems Initiative for South Asia (CSISA) Research Data

**Subject(s)**  
Agriculture, Forestry, Horticulture and Veterinary Medicine | Life Sciences | Plant Nutrition | Agricultural Economics and Sociology | Plant Breeding | Plant Cultivation  
Agriculture, Forestry, Horticulture and Veterinary Medicine

**Content type(s)**  
Standard office documents | Scientific and statistical data formats | Software applications | Source code

**Country**  
International | United States | India

In keeping with the open data policies of the U.S. Agency for International Development (USAID) and Bill & Melinda Gates Foundation, the Cereal Systems Initiative for South Asia (CSISA) has launched the CSISA Data Repository to ensure public accessibility to key data sets, including crop cut data- directly observed, crop yield estimates, on-station and on-farm research trial data and socioeconomic surveys. CSISA is a science-driven and impact-oriented regional initiative for increasing the productivity of cereal-based cropping systems in Bangladesh, India and Nepal, thus improving food security and farmers' livelihoods. CSISA generates data that is of value and interest to a diverse audience of researchers, policymakers and the public. CSISA's data repository is hosted on Dataverse, an open source web application developed at Harvard University to share, preserve, cite, explore and analyze research data. CSISA's repository contains rich datasets, including on-station trial data from 2009–17 about crop and resource management practices for sustainable future cereal-based cropping systems. Collection of this data occurred during the long-term, on-station research trials conducted at the Indian Council of Agricultural Research – Research Complex for the Eastern Region in Bihar, India. The data include information on agronomic management for the sustainable intensification of cropping systems, mechanization, diversification, futuristic approaches to sustainable intensification, long-term effects of conservation agriculture practices on soil health and the pest spectrum. Additional trial data in

We guide consumers to discover, select and use these resources with confidence, and producers to make their resource more discoverable, more widely adopted and cited.



### Researchers in academia, industry and government

Identify and cite the standards, databases or repositories that exist for your discipline when creating a data management plan, releasing data or submitting a manuscript to a journal...

[\[read more\]](#)

- [Researchers](#)
- [Developers & Curators](#)
- [Journal Publishers](#)
- [Librarians & Trainers](#)
- [Societies & Alliances](#)
- [Funders](#)

Find

standards, databases, policies

Search all of FAIRsharing

Standards | Databases | Policies | Collections | Add/Claim Content

**Recommendations**

Standards and/or databases recommended by policies.

View as Table | View as Grid

Sort by: Name

**Recommended Records**

Associated Publication? No Publication Has Publication

Claimed? No Maintainer Has Maintainer

**Record Status**

Uncertain Deprecated In development Ready

**Record Type**

Journal: 92

Funder: 23

Project: 13

Society: 12

**Domains**





Bibliography: 52

Nucleic Acid Sequence: 3

Showing records 1 - 50 of 140.

< 1 2 3 >

Registry	Name	Abbreviation	Type	Subject	Domain	Taxonomy	Related Database	Related Standard	Related Policy	In Collection
	National Child Development Study 1958BC data deposition policy	NCDS Policy	Project	<span>Biomedical Science</span> <span>Practical Studies</span>	None	<span>Homo sapiens</span>	None	None	Material Transfer Agreement for 1958BC samples Policy for use and oversight of samples and data arising from the Biomedical Resource of the 1958 Birth Cohort (National Child Development Study) Conditions of use of 1958BC data including policy on incidental findings (National Child Development Study)	
	African Academy of Sciences Open Research Data Guidelines for Authors	AAS Policy	Journal	<span>Biomedical Science</span> <span>Life Science</span>	<span>Journal Article</span>	<span>All</span>	GenBank BioSD ChEMBL COD Expression Atlas Plus 42 more...	ARRIVE PRISMA CONSORT DataCite Metadata Schema CARE Plus 1 more...	None	None
	Advanced Genetics	None	Journal	<span>Biology</span> <span>Biomedical Science</span> <span>Comparative Genomics</span> <span>Evolutionary Genomics</span> <span>Functional Genomics</span> <span>Plus 3 more...</span>	<span>Annotation</span> <span>Knowledge Representation</span> <span>Protocol</span>	<span>All</span>	GenBank ChEMBL FlyBase TAIR WormBase Plus 36 more...	ISA-Tab OBI EFO CellML ENVO Plus 19 more...	None	None
	Agency for Healthcare Research and Quality Public Access to Federally Funded Research	AHRQ Policy	Funder	<span>Biomedical Science</span> <span>Life Science</span>	None	<span>All</span>	ClinicalTrials.gov	None	None	None
	American Association for the Advancement of Science - Science - Editorial Policies: Data Materials Availability	AAAS Science Policy	Journal	<span>Biomedical Science</span> <span>Environmental Science</span> <span>Life Science</span>	<span>Bibliography</span> <span>Journal Article</span> <span>Publication</span>	<span>All</span>	GenBank CSD ArrayExpress ENA PDB Plus 7 more...	CONSORT MIAME	None	None

 About us ▾ Services ▾ Resources ▾ Community ▾ Become a member   

# WELCOME TO DATACITE

Locate, identify, and cite research data with the leading global provider of DOIs for research data.

[Learn more](#)



Find what you're looking for by searching millions of records with extensive, reliable metadata.



Share your data and reuse the data of others to create the highest impact in the research community.



Cite your research sources with confidence, and receive proper credit when your work is reused.



Connect your research – publications, datasets, software, authors, institutions, and funding data all in one place.


Get started with DataCite!

## DataCite Commons

 Works

 People

 Organizations

 Repositories

Search works by keyword(s) and/or identifier. Documentation is available in [DataCite Support](#).


# Dataset Search



Buscar conjuntos de datos 

Probar [coronavirus covid-19](#) o [education outcomes site:data.gov](#).

[Más información sobre Búsqueda de Datasets](#)

Se han encontrado más de 100 conjuntos de datos

 Evolución de enfermedad por el coronavirus (COVID-19)  
datos.gob.es  
data.europa.eu  
Última actualización: Nov 6, 2020

 **Coronavirus (Covid-19) Data in the United States**  
github.com  
www.nytimes.com  
 csv

 **Coronavirus Disease (COVID-19) – the data**  
ourworldindata.org  
 csv

## Evolución de enfermedad por el coronavirus (COVID-19)

[Ver en datos.gob.es](#)

[Ver en data.europa.eu](#)

72 artículos académicos citan este conjunto de datos ([Ver en Google Académico](#))

### Fecha de actualización del conjunto de datos

Nov 6, 2020

### Conjunto de datos creado y proporcionado por

Ministerio de Sanidad

### Licencia

<https://www.mscbs.gob.es/avisoLegal/home.htm>

### Descripción

Este conjunto de datos ofrece datos sobre el brote de la enfermedad por un nuevo tipo de virus de la familia Coronaviridae, denominado como SARS-CoV-2. La enfermedad causada por este consenso internacional COVID-19.

Los datos aquí recogidos provienen de los reportados por los servicios de salud de las Comunidades Autónomas.

Se está trabajando para mejorar los datos ofrecidos a medida que sea posible.

<https://data.mendeley.com/>

## Share and discover datasets

Mendeley Data is a secure cloud-based repository where you can store your data, ensuring it is easy to share, access and cite, wherever you are.

Create a Dataset

Already using Mendeley? [Sign in](#)

## Find research data

Search 24.2 million datasets from domain-specific and cross-domain repositories

Search

[Advanced search help](#)

Or try: covid-19, coronavirus or epidemic



B2FIND

**B2FIND**

Find research data, research data portal



B2SAFE

**B2SAFE**

Keep research data safe via data management policies



B2SHARE

**B2SHARE**

Store and publish research data



B2DROP

**B2DROP**

Sync and share research data



B2NOTE

**B2NOTE**

Data annotation service to create annotations on research data



B2ACCESS

**B2ACCESS**

Identity & authorisation



B2HANDLE

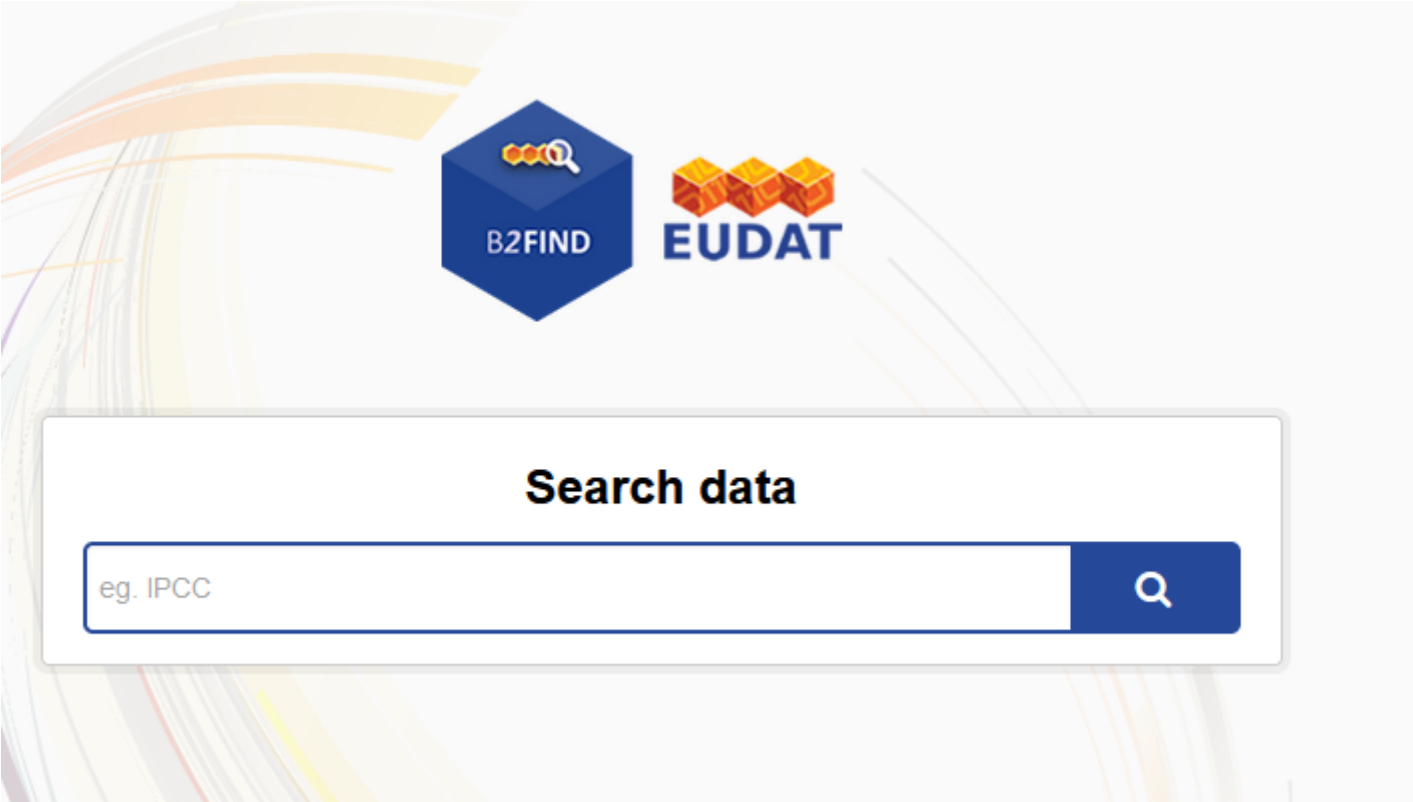
**B2HANDLE**

Register your research data with a persistent identifier

<https://sp.eudat.eu/catalog/>

<http://b2find.eudat.eu/>

<http://b2find.eudat.eu/>



# A global clinical research data sharing platform

The Vivli team is dedicated to helping researchers share and access from clinical trials to advance science.

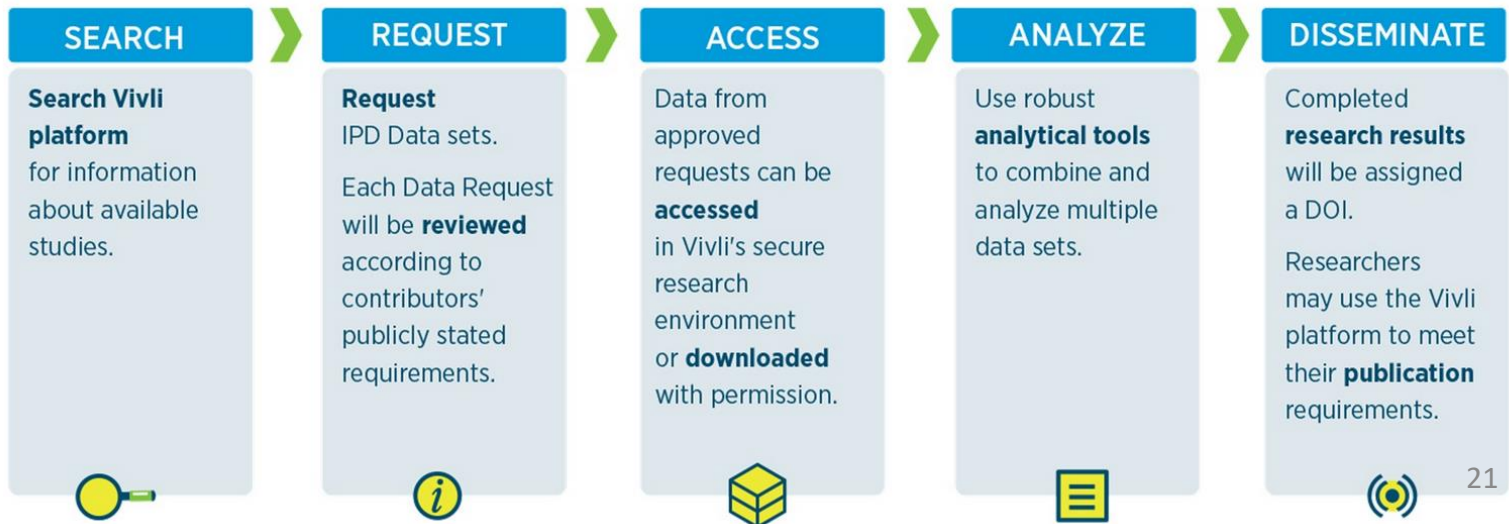
BEGIN SEARCHING FOR STUDIES

<https://vivli.org/>



Vivli provides a **global data-sharing and analytics platform** to serve all elements of the international research community.

Users may **search for studies, request data packages, and analyze data sets** within a secure research environment.



the  
**YODA**  
PROJECT

Forging a unified  
scientific community

ABOUT REQUEST TRIALS METRICS LOG IN

“Discovery consists of looking at the same thing as everyone else and thinking something different.”

*Albert Szent-Györgyi*

**Our Mission**

The Yale University Open Data Access (YODA) Project’s mission is to advocate for the responsible sharing of clinical research data, open science, and research transparency. The Project is committed to supporting research focused on improving the health of patients and informing science and

**Our Model**

The YODA Project seeks mutually beneficial partnerships with Data Partners, promoting independence, responsible conduct of research, good stewardship of data, and the generation of knowledge in the best interest of society. To participate, each Data Partner must transfer full

**Request Data**

Are you ready to request data? To date, 419 trials have been identified as available. The YODA Project and Data Partners continue to identify and add more.

GET STARTED

**¿Cumplen mis datos con los principios FAIR?**

*Note: This page uses iframes. If you do not see the tool directly below, please try a different browser or turn on iframes in your settings.*

Total across F.A.I.R

**Findable** i

**Does the dataset have any identifiers assigned?** No identifier

**Is the dataset identifier included in all metadata records/files describing the data?** No

**How is the data described with metadata?** The data is not described

**What type of repository or registry is the metadata record in?** The data is not described in any repository

**Accessible** i

**How accessible is the data?** No access to data or metadata

**Is the data available online without requiring specialised protocols or tools once access has been approved?** No access to data

**Will the metadata record be available even if the data is no longer available?** Unsure

<https://fairaware.dans.knaw.nl/>



Let's assume you have research data almost ready for uploading to a repository: do you already know how you and the repository can work together to make the data as findable, accessible, interoperable and reusable (FAIR) as possible? By guiding you through the assessment process, the FAIR-Aware tool can help you to better understand the FAIR Principles [↗](#) and how making data FAIR can increase the potential value and impact of your data.

FAIR-Aware is an disciplinary-agnostic online tool developed by the FAIRsFAIR [↗](#) project. Different scientific communities can adapt it to their own use. You should, however, have a target dataset in mind to be able to answer the questions and complete the assessment.

<https://satisfyd.dans.knaw.nl/>



## 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 [↗](#) 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 [✉](#)

## 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 [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

1. Did you provide sufficient metadata (information) about your data for others to find, understand and reuse your data? i

2. Did you use standards such as controlled vocabularies, taxonomies (thesauri) or ontologies to describe your dataset? i

- Controlled vocabularies
- Taxonomies (thesauri)
- Ontologies
- There are no standards for my discipline

3. Did you provide rich and detailed additional documentation? i

- Readme file
- Versioning
- Provenance

#### ACCESSIBLE i

4. Is the metadata publicly accessible even if the data is no longer available? i

Yes  No

I can't find this information in EASY

F




A

I



R

## FAIR questions


### FINDABLE

1. Are you aware that a dataset should be assigned a globally unique persistent and resolvable identifier when deposited with a data repository?   Yes  No
2. Are you aware that when you deposit a dataset with a repository, you will need to provide some details (known as discovery metadata) in order to make the data findable, understandable and reusable to others?   Yes  No
3. Are you aware that the repository providing access to your dataset should make the metadata describing your datasets available in a format readable by machines as well as humans?   Yes  No





### ACCESSIBLE

4. Are you aware that access to your dataset may need to be controlled and that metadata should include licence information under which the data can be reused?   Yes  No
5. Are you aware that metadata should remain available over time, even if the data is no longer accessible?   Yes  No

### INTEROPERABLE

6. Are you aware that the metadata describing your datasets should use semantic vocabularies?   Yes  No

### REUSABLE











7. Are you aware that provenance information about the collection and/or generation of data should be included in the metadata?   Yes  No
8. Are you aware that metadata describing your data should follow the specifications of a community-endorsed standard?   Yes  No
9. Are you aware that data should be deposited preferably in a file format that is open – to support reuse – and supported by the repository for long-term preservation?   Yes  No
10. Are you aware that maintaining your dataset FAIR over time requires professional data curation and preservation?   Yes  No

# Formulario de FAIR AWARE

## Feedback

Please answer the four questions below to help us improve the tool and make it even more relevant for the community.

### • Which of the following issues do you find hard to understand and/or to answer?

1. Globally unique persistent identifier (PID) 
2. Metadata for citation and discovery including PID 
3. Metadata available online is readable by humans and machines 
4. Metadata includes licence, level of access and conditions to access the data 
5. Persistence of metadata 
6. Use of controlled vocabularies in metadata 
7. Metadata includes provenance 
8. Community-endorsed metadata 
9. Data in a preferred format for reuse & preservation 
10. Digital curation and preservation 

### • Are there any issues relevant to your discipline and/or needed for enabling reusability of your dataset that are missing from this assessment? (Please do not include any personal data, such as your name and email address in your response.)

### • Please submit any other feedback on how we might improve the FAIR-Aware assessment tool, including possible additional guidance. (Please do not include any personal data, such as your name and email address in your response.)



The tool FairDataBR, aimed to evaluate data sets, was written from the perception of the need to develop a software application that contributed to the automation process of verifying the adherence of data sets to the FAIR Principles. The tool was designed by researchers from the Federal University of Paraíba (PPGCI / MPGOA - UFPB) and it is characterized by being simple and intuitive to use.

The FAIR principles idea emerged from the work of a diversified group of people who saw the immediate need to improve the infrastructure to support the sharing of research data and its subsequent reuse. Thus, based on these premises, a set of principles was developed to serve as guidelines for all who wish to share and enhance the reuse of research data (WILKINSON et al., 2016).

## FAIR PRINCIPLES

The FAIR Principles are an acronym for Findable, Accessible, Interoperable and Reusable. On the pioneer publication of a scientific paper on these principles, Wilkinson et al. (2016) stated that to research data be under these principles, they must follow the following guidelines:

To be Findable:	To be Accessible:	To be Interoperable:	To be Reusable:
F1. (Meta)data are assigned a globally unique and persistent identifier. F2. Data are described with rich metadata (defined by R1 below). F3. metadata clearly and explicitly include the identifier of the data it describes.	A1. (Meta)data are retrievable by their identifier using a standardized communications <a href="#">protocol</a> . A1.1. The protocol is open, free, and universally implementable. A1.2. the protocol	I1. (Meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation. I2. (Meta)data use vocabularies that follow FAIR principles. I3. (Meta)data include qualified references to	R1. meta(data) are richly described with a plurality of accurate and relevant attributes. R1.1. (Meta)data are released with a clear and accessible data usage license. R1.2. (Meta)data are associated with

Avaliação Fair

**FINDABLE PRINCIPLE**

F1. Do (meta)data have a unique, global and persistent identifier? 1 2 3 4 5 6 \*

Persistent Identifier  
 Web Address  
 Local Identifier  
 No Identifier

F2. Are data described with rich metadata? 1 3 4 5 \*

Rich metadata  
 Structured metadata  
 Simple metadata  
 No metadata

F3. metadata clearly and explicitly include the identifier of the data they describe? 1 4 5 \*

Yes  
 No

F4. Are (meta)data registered or indexed in a searchable resource? 1 5 \*

Yes  
 No

F5. Are (meta)data published in a repository? 1 2 \*

General repository  
 Domain specific  
 Institutional repository  
 Not published in a repository

# ¿Son FAIR mis datos?



<https://fair.csic.es>



<https://www.f-uji.net/>