



# Datos de investigación: retos, oportunidades y buenas prácticas ante la reforma de la evaluación de la investigación

**Daniel Garijo**

Universidad Politécnica de Madrid

[daniel.garijo@upm.es](mailto:daniel.garijo@upm.es)

@dgarijov

**Madrid, 24 de octubre de 2024**

<https://doi.org/10.5281/zenodo.13971475>

 08/02/2024

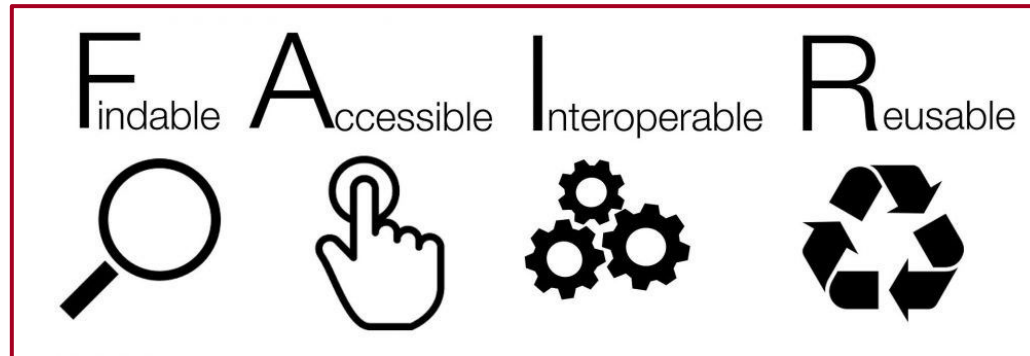
## Publicado el borrador de los nuevos criterios de acreditación a PTU y CU (RD 678/2023)

ANECA publica el [borrador](#) de los nuevos criterios de evaluación y requisitos mínimos de referencia de los méritos y competencias requeridos para obtener la acreditación estatal para el acceso a los cuerpos docentes universitarios



<https://www.aneca.es/-/publicado-el-borrador-de-los-nuevos-criterios-de-acreditaci%C3%B3n-a-ptu-y-cu-rd-678/2023->

- **Evaluación científica:** ¿Cómo hemos llegado hasta aquí?
- Datos de investigación: ¿Qué son y qué **buenas prácticas hay**?
- ¿Cómo **adoptar buenas prácticas** invirtiendo el **mínimo de tiempo**?

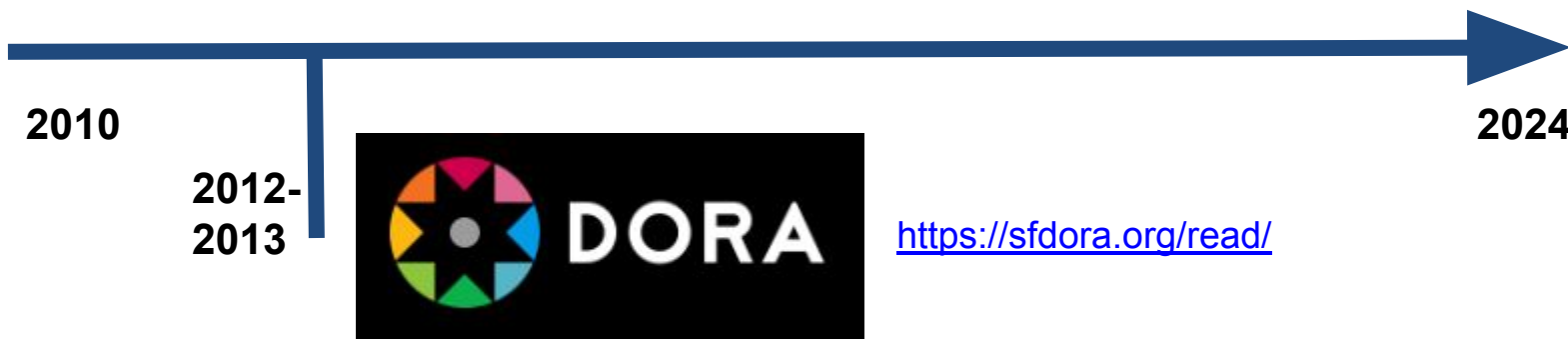


- **Evaluación científica**: ¿Cómo hemos llegado hasta aquí?
- Datos de investigación: ¿Qué son y qué **buenas prácticas hay**?
- ¿Cómo **adoptar buenas prácticas** invirtiendo el **mínimo de tiempo**?



### **Métricas “tradicionales”**

- **Journal Impact Factor (JIF)**
- **h-index (2005)**



# San Francisco Declaration on Research Assessment

A number of themes run through these recommendations:

<https://sfdora.org/read/>

- the need to eliminate the use of journal-based metrics, such as Journal Impact Factors, in funding, appointment, and promotion considerations;
- the need to assess research on its own merits rather than on the basis of the journal in which the research is published; and
- the need to capitalize on the opportunities provided by online publication (such as relaxing unnecessary limits on the number of words, figures, and references in articles, and exploring new indicators of significance and impact).

3. For the purposes of research assessment, consider the value and impact of all research outputs (including datasets and software) in addition to research publications, and consider a broad range of impact measures including qualitative indicators of research impact, such as influence on policy and practice.



Bibliometrics: The Leiden Manifesto for research metrics

Hicks, D., Wouters, P., Waltman, L. *et al.* Bibliometrics: The Leiden Manifesto for research metrics. *Nature* **520**, 429–431 (2015).  
<https://doi.org/10.1038/520429a>



**1) Quantitative evaluation should support qualitative, expert assessment.** Quantitative metrics can challenge bias tendencies in peer review and facilitate deliberation. This should strengthen peer review, because making judgements about colleagues is difficult without a range of relevant information. However, assessors must not be tempted to cede decision-making to the numbers. Indicators must not substitute for informed judgement. Everyone retains responsibility for their assessments.

**5) Allow those evaluated to verify data and analysis.** To ensure data quality, all researchers included in bibliometric studies should be able to check that their outputs have been correctly identified. Everyone directing and managing evaluation processes should assure

**7) Base assessment of individual researchers on a qualitative judgement of their portfolio.** The older you are, the higher your *h*-index, even in the absence of new papers.



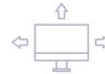


1

We will make openness the default for the research information we use and produce

---

<https://barcelona-declaration.org/>



2

We will work with services and systems that support and enable open research information

---



4

We will support collective action to accelerate the transition to openness of research information

---

# Coalition for Advancing Research Assessment

<https://coara.eu/>

<https://coara.eu/agreement/the-agreement-full-text/> (2022)

Dirección común para proponer cambios en la evaluación investigadora, investigadores y organizaciones que investigan, con el objeto de maximizar la calidad y el impacto de la ciencia

1. Recognise the diversity of contributions to, and careers in, research in accordance with the needs and nature of the research

- Reconocimiento **más allá de publicaciones**
- Distinguir el **rol de los científicos** (gestores de datos, ingenieros software, comunicadores, mentores...)

2. Base research assessment primarily on qualitative evaluation for which peer review is central, supported by responsible use of quantitative indicators
3. Abandon inappropriate uses in research assessment of journal- and publication-based metrics, in particular inappropriate uses of Journal Impact Factor (JIF) and h-index
4. Avoid the use of rankings of research organisations in research assessment
10. Evaluate practices, criteria and tools based on solid evidence and the state-of-the-art in research on research, and make data openly available for evidence gathering and research


Algunas instituciones que han firmado el acuerdo (de **más de 70 en España**) incluyen:



CSIC  
Consejo Superior de Investigaciones Científicas



UNIVERSIDAD  
POLITÉCNICA  
DE MADRID

 03/04/2023

# ANECA se adhiere a DORA y a CoARA

ANECA, participe de los debates que se han ido desarrollando en el [contexto europeo](#) y global sobre la necesidad de reformar los modelos de evaluación de la investigación, se ha adherido a la *San Francisco Declaration on Research Assessment* (DORA) y a la *Coalition for Advancing Research Assessment* (CoARA) con la convicción de que nos encontramos en un momento óptimo para avanzar hacia una nueva cultura de evaluación de la actividad investigadora en toda su diversidad.

<https://www.aneca.es/-/aneca-se-adhiere-a-la-san-francisco-declaration-on-research-assessment-dora-y-a-la-coalition-for-advancing-research-assessment-coara->

Algunas consecuencias de estos cambios...


# ANECA retira el informe "Análisis bibliométrico e impacto de las editoriales open-access en España" (2021)

En coherencia con la adhesión realizada en el mes de abril a la San Francisco Declaration on Research Assessment ([DORA](#)) y a la Coalition for Advancing Research Assessment ([CoARA](#)), así como con la participación de la Agencia en una [propuesta de Capítulo Nacional de CoARA](#) en el mes de mayo, la dirección de ANECA ha retirado el informe "Análisis bibliométrico e impacto de las editoriales open-access en España" (2021).

---

[https://www.aneca.es/web/guest/-/aneca-retira-el-informe-an%C3%A1lisis-bibliom%C3%A9trico-e-impacto-de-las-editoriales-open-access-en-espa%C3%B1a-2021-?p\\_l\\_back\\_url=%2Fresultados-busqueda%3Fq%3Dbibliometrico](https://www.aneca.es/web/guest/-/aneca-retira-el-informe-an%C3%A1lisis-bibliom%C3%A9trico-e-impacto-de-las-editoriales-open-access-en-espa%C3%B1a-2021-?p_l_back_url=%2Fresultados-busqueda%3Fq%3Dbibliometrico)



 08/02/2024

## Publicado el borrador de los nuevos criterios de acreditación a PTU y CU (RD 678/2023)

ANECA publica el [borrador](#) de los nuevos criterios de evaluación y requisitos mínimos de referencia de los méritos y competencias requeridos para obtener la acreditación estatal para el acceso a los cuerpos docentes universitarios



<https://www.aneca.es/-/publicado-el-borrador-de-los-nuevos-criterios-de-acreditaci%C3%B3n-a-ptu-y-cu-rd-678/2023->

- La evaluación se basará en la **selección de contribuciones y en la narrativa sobre su calidad, relevancia e impacto** aportadas por la persona solicitante. El resultado de la evaluación será favorable o desfavorable.

#### **Beneficios [1]:**

- Mayor flexibilidad para elegir métricas.
- No desaparece JIF/h-index, se complementa
- Se admiten indicadores de impacto social (e.g., menciones en Twitter)
- Se admiten indicadores de uso y visibilidad
- Se valora si está disponible en abierto (FAIR)

#### **Riesgos [1]:**

- Desconfianza de las fuentes de datos usadas
- Uso de métricas poco conocidas
- Falta de soporte en lo reportado

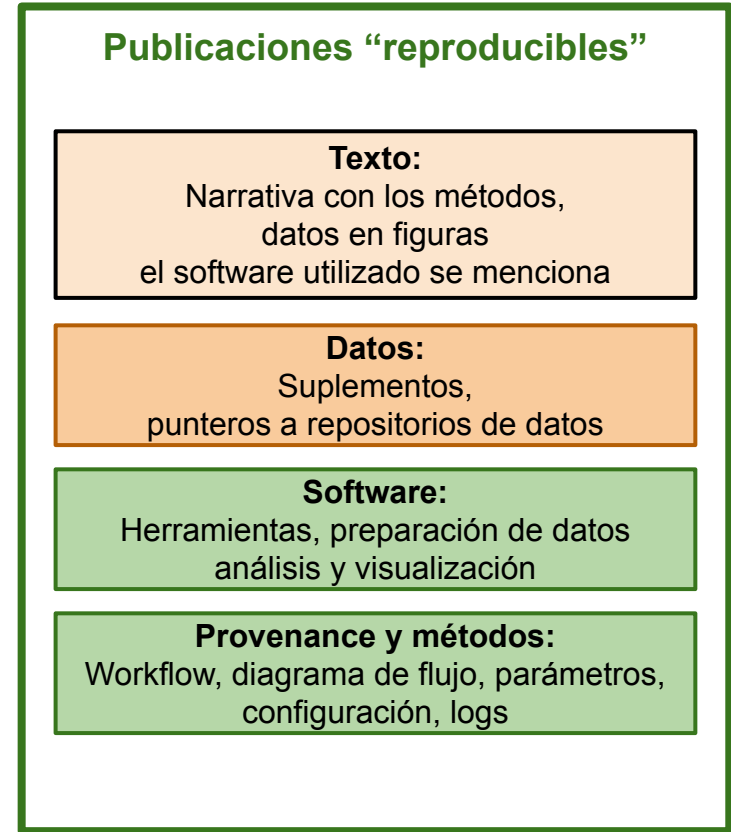
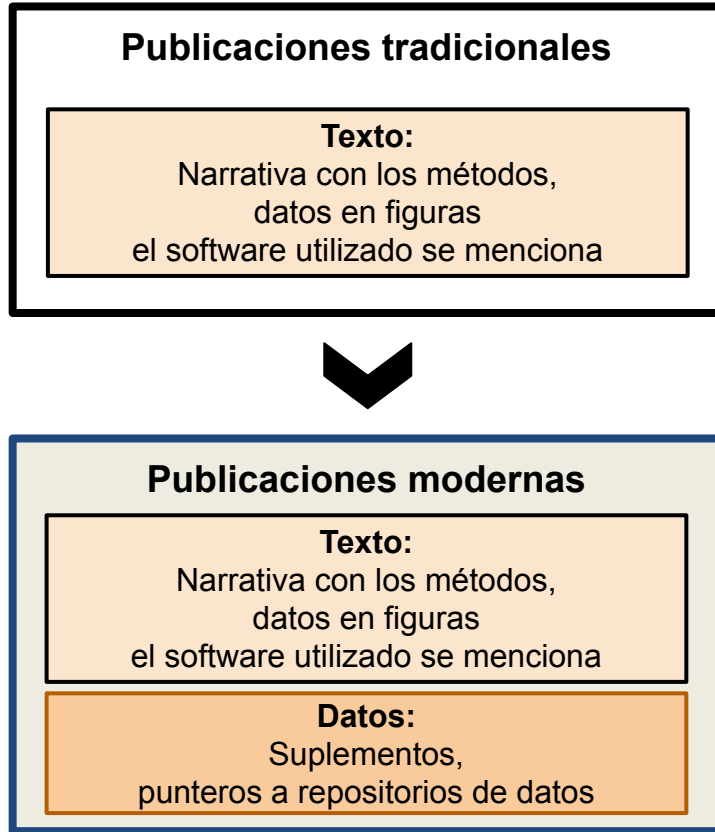
[1] Wenceslao Arroyo-Machado. ANECA: nuevos criterios en la acreditación y en la evaluación de la investigación. [https://hackmd.io/@Wences/ANECA\\_nuevos\\_criterios](https://hackmd.io/@Wences/ANECA_nuevos_criterios)

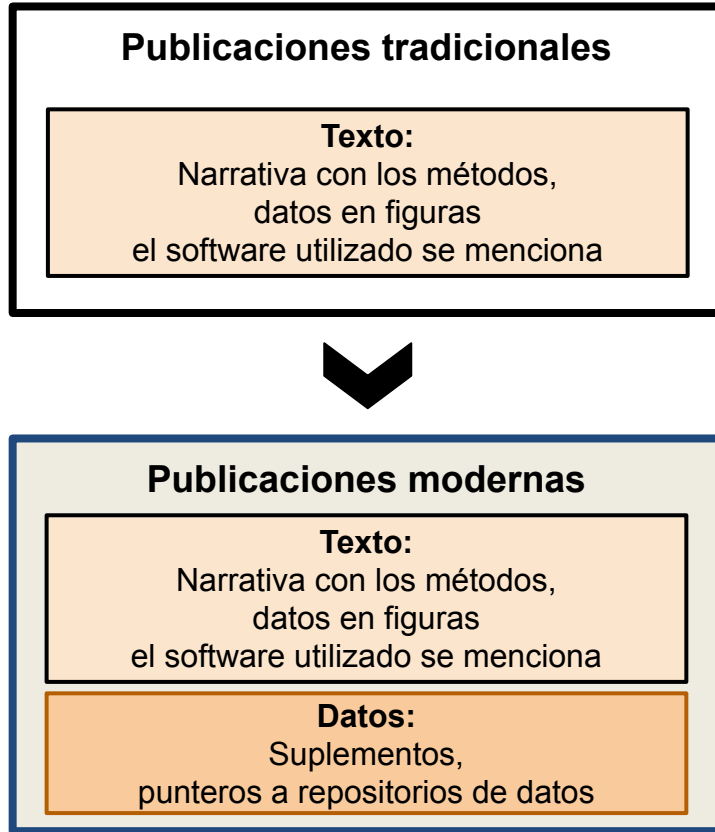
- La evaluación se basará en la **selección de contribuciones y en la narrativa sobre su calidad, relevancia e impacto aportadas por la persona solicitante**. El resultado de la evaluación será favorable o desfavorable.
- Las personas solicitantes deberán destacar la **calidad, la relevancia y el impacto científico y social de sus contribuciones, con apoyo de indicadores contextualizados**. Se deben incluir los **identificadores persistentes** y enlaces que se estimen convenientes a las bases de datos utilizadas comúnmente en los ámbitos de conocimiento correspondientes, así como a los **repositorios institucionales, temáticos o generalistas** y a **agregadores** reconocidos internacionalmente. En todos los casos, se valorará el grado de internacionalización, de innovación, de **apertura** a la sociedad y de multidisciplinariedad o interdisciplinariedad de las contribuciones presentadas.



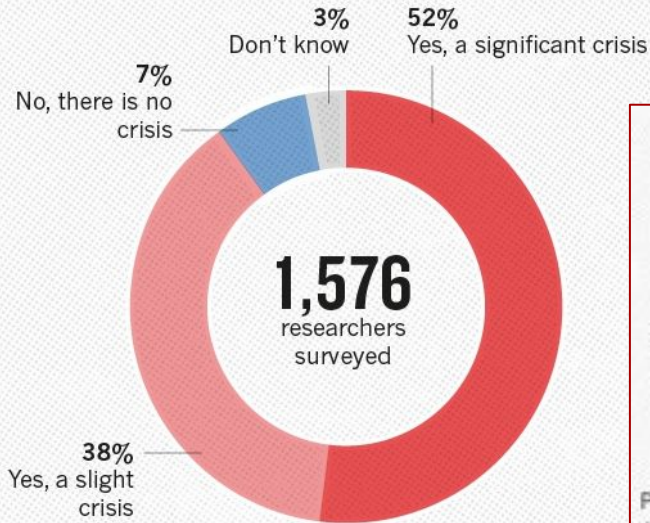
- La evaluación se basará en la **selección de contribuciones y en la narrativa sobre su calidad, relevancia e impacto aportadas por la persona solicitante**. El resultado de la evaluación será favorable o desfavorable.
- Las personas solicitantes deberán destacar la **calidad, la relevancia y el impacto científico y social de sus contribuciones, con apoyo de indicadores contextualizados**. Se deben incluir los **identificadores persistentes** y enlaces que se estimen convenientes a las bases de datos utilizadas comúnmente en los ámbitos de conocimiento correspondientes, así como a los **repositorios institucionales, temáticos o generalistas** y a **agregadores** reconocidos internacionalmente. En todos los casos, se valorará el grado de internacionalización, de innovación, de **apertura** a la sociedad y de multidisciplinariedad o interdisciplinariedad de las contribuciones presentadas.
- También se valorará la actividad en proyectos y actividades de innovación docente, cuyos resultados **deben incluir los identificadores persistentes y enlaces a contenidos y recursos educativos abiertos** que se estimen convenientes, y la tutorización docente tanto reglada como de carácter voluntario.

- **Evaluación científica:** ¿Cómo hemos llegado hasta aquí?
- **Datos de investigación:** ¿Qué son y qué **buenas prácticas hay?**
- ¿Cómo **adoptar buenas prácticas** invirtiendo el **mínimo de tiempo?**





## IS THERE A REPRODUCIBILITY CRISIS?



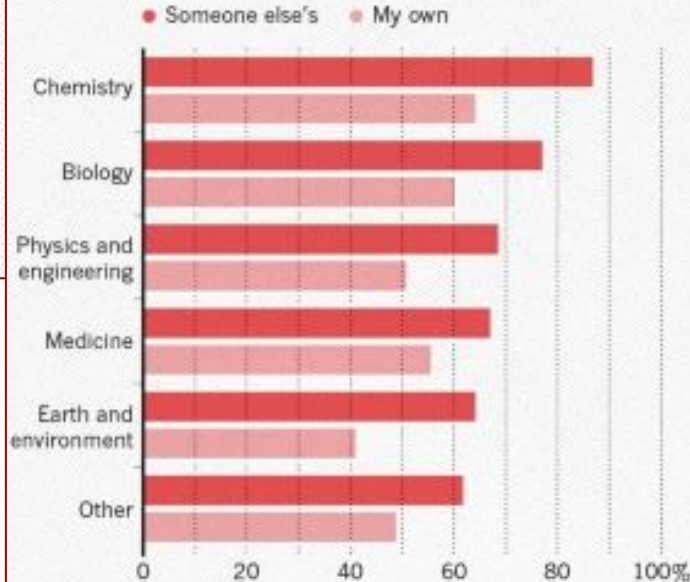
## Retraction Watch

Tracking retractions as a window into the scientific process



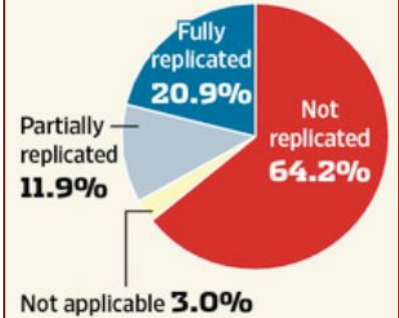
## HAVE YOU FAILED TO REPRODUCE AN EXPERIMENT?

Most scientists have experienced failure to reproduce results.



## No Cure

When Bayer tried to replicate results of 67 studies published in academic journals, nearly two-thirds failed.



Source: Nature Reviews Drug Discovery

<https://www.nature.com/articles/533452a>





# Buenas prácticas de Ciencia Abierta para datos de investigación:

¿Verdad o mito?

## Datos abiertos



## Licencias Abiertas



## Código abierto



## Modelos abiertos



## Contenedores abiertos



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

# CESAER

The strong and united voice of universities  
of science and technology in Europe

## NEXT GENERATION METRICS

WHITE PAPER DATED 10 JUNE 2020

With a full implementation in mind, we propose a set of indicators for measuring research performance through open science ([annexe I](#)). The purpose of this set is to help to: (i) move towards open access publishing of scientific publications; (ii) involve citizens more in science; (iii) promote FAIR data; (iv) store more scientific data on open repositories; (v) strengthen transparency and integrity in research; (vi) inspire policy-makers, research institutions, funding bodies and researchers themselves to use context-appropriate metrics to complement qualitative assessments for monitoring and stimulating development.

<https://www.cesaer.org/publications/next-generation-metrics/>



Data availability statements and data citations policy: guidance for authors

**Policy summary**

All manuscripts reporting original research must include a data availability statement. Authors are also encouraged to include formal cita-

nature.com > scientific data

SCIENTIFIC DATA

 nature.com



**Availability of Software**

PLOS supports the development of open source software and believes that, for submissions appropriate open source standards will ensure that the submission conforms to (1) our requirement that another researcher can reproduce the experiments described, (2) our aim to promote open science so that PLOS journals can be built upon by future researchers. Therefore, if new software or a new article submitted to a PLOS journal can be built upon by future researchers. Therefore, if new software or a new article submitted to a PLOS journal that the software conforms to the [Open Source Definition](#), have deposited the following three items as Supporting Information:

- **The associated source code of the software described by the paper.** This should be licensed under a suitable license such as BSD, LGPL, or MIT (see <http://www.copdess.org>). Commercial software such as Mathematica and MATLAB does not preclude a paper from being open access.
- **Documentation for running and installing the software.** For end-user applications this is a prerequisite; for software libraries, instructions for using the application program interface are also required.
- **A test dataset with associated control parameter settings.** Where feasible, results from running the software on a test dataset should not have any dependencies — for example, a database dump.

Acceptable archives should provide a public repository of the described software. The code should be available for creating user accounts, logging in or otherwise registering personal details. The repository should host more than 1,000 projects. Examples of such archives are: [SourceForge](#), [Bioinformatics.Org](#), [Savannah](#), [GitHub](#) and the [Codehaus](#). Authors should provide a direct link to the deposited software.

# COPDESS

Coalition on Publishing Data  
in the Earth and Space  
Sciences

## COPDESS Suggested Author Instructions and Best Practices for Journals

The Coalition on Publishing Data in the Earth and Space Sciences ([COPDESS](#)) develops and recommends best practices for journal author instructions around data and identifiers as a resource to the community. These best practices are consistent with and based on the COPDESS Statement of Commitment and have been developed with guidance from participants in COPDESS.

[Data Policy Statement](#)

[Data Citation](#)

[Sample Citation and Identification](#)

[Crossref Funder Registry](#)

[ORCIDs](#)

[Presentations on Best Practices](#)



CRT

Contributor Role Taxonomy

<https://credit.niso.org/>



Strategic Research and Innovation Agenda (SRIA)

<https://op.europa.eu/en/publication-detail/-/publication/f9b12d1d-74ea-11ec-9136-01aa75ed71a1/language-en>

**EOSC**

## What the European Open Science Cloud is

The ambition of the European Open Science Cloud (EOSC) is to develop “Web of FAIR Data and services’ for science in Europe. EOSC will be a multi-disciplinary environment where researchers can publish, find and re-use data, tools and services, enabling them to better conduct their work.

EXECUTIVE OFFICE OF THE PRESIDENT  
OFFICE OF SCIENCE AND TECHNOLOGY POLICY  
WASHINGTON, D.C. 20502

February 22, 2013

MEMORANDUM FOR THE HEADS OF EXECUTIVE DEPARTMENTS AND AGENCIES

FROM: John P. Holdren *JPH*  
Director

SUBJECT: Increasing Access to the Results of Federally Funded Scientific Research

### 1. Policy Principles

The Administration is committed to ensuring that, to the greatest extent and with the fewest constraints possible and consistent with law and the objectives set out below, the direct results of federally funded scientific research are made available to and useful for the public, industry, and the scientific community. Such results include peer-reviewed publications and digital data.

**an approach for optimizing search, archival, and dissemination features that encourages innovation in accessibility and interoperability, while ensuring long-term stewardship of the results of federally funded research;**



# Estrategia Nacional de Ciencia Abierta (ENCA)

2023 – 2027



<https://www.boe.es/eli/es/lo/2023/03/22/2/con>

Ley Orgánica 2/2023, de 22 de marzo, del Sistema Universitario.

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.

## EJES ESTRATÉGICOS

A

Infraestructuras digitales para la ciencia abierta.

B

Gestión de datos de investigación siguiendo los principios FAIR.

C

Acceso abierto a publicaciones científicas.

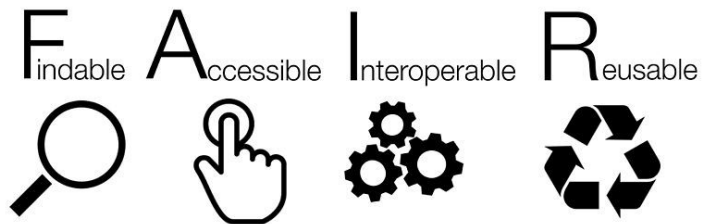
D

Incentivos, reconocimientos y formación.

<https://www.ciencia.gob.es/en/Estrategias-y-Planes/Estrategias/ENCA.html>

<https://www.force11.org/group/fairgroup/fairprinciples>

<https://doi.org/10.1038/sdata.2016.18>



## Implementing FAIR Data Principles: The Role of Libraries

### What are the FAIR Data Principles?

The FAIR Data Principles are a set of guiding principles in order to make data findable, accessible, interoperable and reusable (often referred to as FAIR). These principles provide guidance for scientific data management and stewardship and are relevant to all stakeholders in the current digital ecosystem. This activity address data producers and data consumers to promote maximum use of research data. Research libraries can use the FAIR Data Principles as a framework for leading and extending research data services.

### What is FAIR DATA?



### To be Findable:

- F1. (meta)data are assigned a globally unique and eternally persistent identifier.
- F2. data are described with rich metadata.
- F3. (meta)data are registered or indexed in a searchable resource.
- F4. metadata specify the data identifier.

### TO BE ACCESSIBLE:

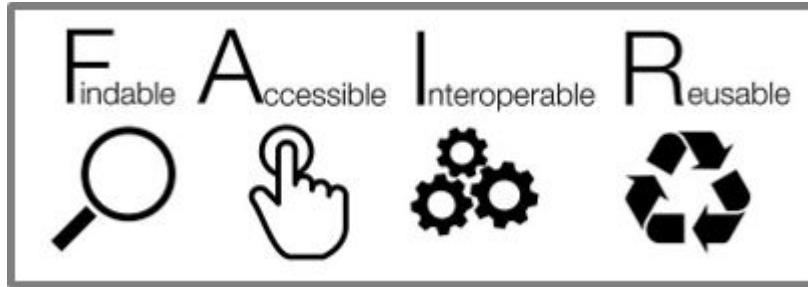
- A1 (meta)data are retrievable by their identifier using a standardized communications protocol.
  - A1.1 the protocol is open, free, and universally implementable.
  - A1.2 the protocol allows for an authentication and authorization procedure, where necessary.
- A2 metadata are accessible, even when the data are no longer available.

### TO BE INTEROPERABLE:

- 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 other (meta)data.

### TO BE RE-USABLE:

- R1. meta(data) have 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 their provenance.
  - R1.3. (meta)data meet domain-relevant community standards.



Datos (inicialmente) [1]



Research Software (FAIR4RS)



Methods (FAIR Workflows)



FAIR Semantic artefacts

Otras guías:

- Guidelines for Transparency and Openness Promotion (TOP) [2]
- Reproducibility Enhancement Principles (REP) [3]
- ...



[1] Wilkinson, M., Dumontier, M., Aalbersberg, I. *et al.* The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* 3, 160018 (2016). <https://doi.org/10.1038/sdata.2016.18>

[2] <https://www.cos.io/initiatives/top-guidelines>

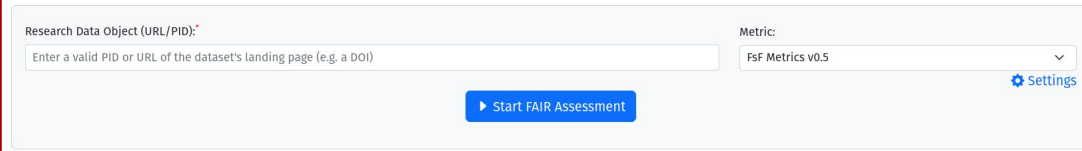
[3] Stodden, V *et al* Enhancing reproducibility for computational methods  
<https://www.science.org/lookup/doi/10.1126/science.aah6168>



## FAIR assessment

F-UJI is a web service to programatically assess FAIRness of research data objects (aka data sets) based on metrics developed by the [FAIRsFAIR](#) project.

Please use the form below to enter an identifier (e.g. DOI, URL) of the data set you wish to assess. Optionally you also can enter a metadata service (OAI-PMH, SPARQL, CSW) endpoint URI which F-UJI can use to identify additional information.




The screenshot shows a web form for FAIR assessment. It has two input fields: 'Research Data Object (URL/PID):' with a placeholder 'Enter a valid PID or URL of the dataset's landing page (e.g. a DOI)' and 'Metric:' with a dropdown menu showing 'Fsf Metrics v0.5'. A 'Settings' link with a gear icon is next to the metric dropdown. A blue button labeled 'Start FAIR Assessment' is at the bottom.

Librería para evaluar Software científico:

<https://github.com/fair-software/howfairis/>



**foops!** Ontology Pitfall Scanner for FAIR (Beta)



The screenshot shows a web form for the foops! Ontology Pitfall Scanner. It has a blue header with the text 'foops! Ontology Pitfall Scanner for FAIR (Beta)'. Below the header is a light blue box labeled 'URI'. Inside this box is a text input field with a placeholder 'Enter URI...' and an example: 'Example: https://w3id.org/example (click here to enter this ontology)'. A blue button labeled 'RUN' is at the bottom.

Evaluador para ontologías:

<https://w3id.org/foops/>

- **Evaluación científica:** ¿Cómo hemos llegado hasta aquí?
- Datos de investigación: ¿Qué son y qué **buenas prácticas hay**?
- ¿Cómo **adoptar buenas prácticas** invirtiendo el **mínimo de tiempo**?
  - Datos

*Nature Genetics* **41**, 149 - 155 (2009)

Published online: 28 January 2008 | doi:10.1038/ng.295

## Repeatability of published microarray gene expression analyses

scientists. Here we evaluated the replication of data analyses in 18 articles on microarray-based gene expression profiling published in *Nature Genetics* in 2005–2006. One table or figure from each article was independently evaluated by two teams of analysts. We reproduced two analyses in principle and six partially or with some discrepancies; ten could not be reproduced. The main reason for failure to reproduce was data unavailability, and discrepancies were mostly due to incomplete data annotation or specification of data processing and analysis.

- Datos **no disponibles**

PLOS ONE | DOI:10.1371/journal.pone.0115253 December 26, 2014

RESEARCH ARTICLE

## Scholarly Context Not Found: One in Five Articles Suffers from Reference Rot

Martin Klein<sup>1\*</sup>, Herbert Van de Sompel<sup>1</sup>, Robert Sanderson<sup>1</sup>, Harihar Shankar<sup>1</sup>, Lyudmila Balakireva<sup>1</sup>, Ke Zhou<sup>2</sup>, Richard Tobin<sup>2</sup>

We analyze a vast collection of articles from three corpora that span publication years 1997 to 2012. For over one million references to web resources extracted from over 3.5 million articles, **we observe that the fraction of articles containing references to web resources is growing steadily over time.** We find **one out of five STM articles suffering from reference rot**, meaning it is impossible to revisit the web context that surrounds them some time after their publication. When only considering STM articles that contain references to web resources, this fraction increases to seven out of ten.

- Datos disponibles pero la **URL no resuelve**

## Artículos de datos

Ecological Research  
July 2013, Volume 26, Issue 4, p 541  
Date: 10 May 2013



### Monitoring records of plant species in the Hakone region of Fuji-Hakone-Izu National Park, Japan, 2001–2010

Takeshi Osawa

#### Abstract

The monitoring of species occurrences is a crucial aspect of biodiversity conservation, and regional volunteerism can serve as a powerful tool in such endeavors. The Fuji-Hakone-Izu National Park in the Hakone region of Kanagawa Prefecture, Japan, boasts a volunteer association of approximately 100 members. These volunteers have monitored plant species occurrences from 2001 to the present along several hiking trails in the region. In this paper, I present the annual observation records of plant occurrences in Hakone from 2001 to 2010. This data set includes 1,071 species of plants from 151 families. Scientific names follow the Y List, and this data set includes several threatened plant species. Data files are formatted based on the Darwin Core and Darwin Core Archives, which are defined by the Biodiversity Information Standards (BIS) or Biodiversity Information Standards Taxonomic Databases Working Group (TDWG). Data files filled on required and some additional item on Darwin Core. The data set can download from the author's personal Web site as of July 2012. These data will soon be published for the Global Biodiversity Information Facility (GBIF) through GBIF Japan. All users can then access the data from the GBIF portal site.

\* The complete data set for this abstract published in the Data Paper section of the journal is available in electronic format in Ecological Research Data Paper Archives at [http://db.cger.nies.go.jp/Ja/ITER/IER\\_DataPapers/archives/2013/ERDP-2013-01](http://db.cger.nies.go.jp/Ja/ITER/IER_DataPapers/archives/2013/ERDP-2013-01).

## Repositorios de dominio



+/- NTL LTER "WDNR Yahara Lakes Fisheries: Fish Lengths and Weights 1987-1998" - Lathrop

#### LTER Identifier:

knb-iter-ntl.279.1

#### Abstract:

These data were collected by the Wisconsin Department of Natural Resources (WDNR) from 1987-1998. Most of these data (1987-1993) precede 1995, the year that the University of Wisconsin's NTL-LTER program took over sampling of the Yahara Lakes. However, WDNR data collected from 1997-1998 (unrelated to LTER sampling) is also included. In 1987 a joint project by the WDNR and the University of Wisconsin-Madison, Center for Limnology (CFL) was initiated on Lake Mendota. The project involved biomanipulation o...

#### Owners/Creators:

Lathrop

#### Metadata:

Select [here](#) for full metadata

#### Data File(s):

- [wdnr\\_hyke\\_miniflyke\\_seine\\_lengths\\_weights.csv](#)
- [wdnr\\_boomshock\\_lengths\\_weights.csv](#)
- [wdnr\\_gillnet\\_lengths\\_weights\\_93.csv](#)
- [wdnr\\_waiteye\\_age\\_lengths\\_weights\\_87.csv](#)
- [wdnr\\_creel\\_survey\\_lengths\\_weights.csv](#)
- [wdnr\\_creel\\_survey\\_angler\\_counts.csv](#)

February 5, 2023

Dataset Open Access

# A large-scale COVID-19 Twitter chatter dataset for open scientific research - an international collaboration

Bandu, Juan M.; Tekumalla, Ramya; Wang, Guanyu; Yu, Jingyuan; Liu, Tuo; Ding, Yuning; Artemova, Katya; Tutubalina, Elena; Chowell, Gerardo

Version 152 of the dataset. **MAJOR CHANGE NOTE:** The dataset files: `full_dataset.tsv.gz` and `full_dataset_clean.tsv.gz` have been split in 1 GB parts using the Linux utility called `Split`. So make sure to join the parts before unzipping. We had to make this change as we had huge issues uploading files larger than 2GB's (hence the delay in the dataset releases). The peer-reviewed publication for this dataset has now been published in *Epidemiologia an MDPI Journal*, and can be accessed here: <https://doi.org/10.3390/epidemiologia2030024>. Please cite this when using the dataset.

File Name	Size
2020-01-04_clean-emoji_char.tsv	11 Bytes
2020-01-04_clean-emoji_text.tsv	21 Bytes
2020-01-06_clean-emoji_char.tsv	1 Byte
2020-01-06_clean-emoji_text.tsv	1 Byte
2020-01-08_clean-emoji_char.tsv	1 Byte
2020-01-08_clean-emoji_text.tsv	1 Byte
2020-01-09_clean-emoji_char.tsv	1 Byte
2020-01-09_clean-emoji_text.tsv	1 Byte
2020-01-10_clean-emoji_char.tsv	1 Byte
2020-01-10_clean-emoji_text.tsv	1 Byte
2020-01-11_clean-emoji_char.tsv	29 Bytes
2020-01-11_clean-emoji_text.tsv	104 Bytes

Files (15.9 GB)

**Publication date:**

February 5, 2023

**DOI:**

DOI: 10.5281/zenodo.7608146

**Keyword(s):**

social media twitter nlp covid-19 covid19

**Published in:**

Epidemiologia: 2 pp. 315-324 (3).

**Related identifiers:**

Continued by  
<http://www.panacealab.org/covid19/> (Other)  
 Supplement to  
<https://arxiv.org/abs/2004.03688> (Preprint)

**Alternate identifiers:**

10.3390/epidemiologia2030024 (Journal article)  
[https://github.com/thepanacealab/covid19\\_twitter](https://github.com/thepanacealab/covid19_twitter)  
 (Software)

**Communities:**

Bio-Hackathon  
 Coronavirus Disease Research Community - COVID-19  
 Zenodo

**License (for files):**

Other (Public Domain)

**Cite as**

Bandu, Juan M., Tekumalla, Ramya, Wang, Guanyu, Yu, Jingyuan, Liu, Tuo, Ding, Yuning, Artemova, Katya, Tutubalina, Elena, & Chowell, Gerardo. (2023). A large-scale COVID-19 Twitter chatter dataset for open scientific research - an international collaboration [Data set]. In *Epidemiologia* (Version 152, Vol. 2, Number 3, pp. 315–324). Zenodo. <https://doi.org/10.5281/zenodo.7608146>

Start typing a citation style...

1. Publicar en un repositorio público
2. Añadir metadatos mínimos
3. Añadir metadatos de dominio
4. Usar identificadores persistentes
5. Indicar cómo citar el recurso



February 5, 2023

Dataset Open Access

# A large-scale COVID-19 Twitter chatter dataset for open scientific research - an international collaboration

Bandu, Juan M.; Tekumalla, Ramya; Wang, Guanyu; Yu, Jingyuan; Liu, Tuo; Ding, Yuning; Artemova, Katya; Tutubalina, Elena; Chowell, Gerardo

Version 152 of the dataset. MAJOR CHANGE NOTE: The dataset files: `full_dataset.tsv.gz` and `full_dataset_clean.tsv.gz` have been split in 1 GB parts using the Linux utility called `Split`. So make sure to join the parts before unzipping. We had to make this change as we had huge issues uploading files larger than 2GB's (hence the delay in the dataset releases). The peer-reviewed publication for this dataset has now been published in *Epidemiologia an MDPI Journal*, and can be accessed here: <https://doi.org/10.3390/epidemiologia2030024>. Please cite this when using the dataset.

**Publication date:**

February 5, 2023

**DOI:**DOI: [10.5281/zenodo.7608146](https://doi.org/10.5281/zenodo.7608146)**Keyword(s):**

social media twitter nlp covid-19 covid19

**Published in:**

Epidemiologia: 2 pp. 315-324 (3).

**Related identifiers:**Continued by  
<http://www.panacealab.org/covid19/> (Other)

## Supplement to

<https://arxiv.org/abs/2004.03688> (Preprint)**Alternate identifiers:**[10.3390/epidemiologia2030024](https://doi.org/10.3390/epidemiologia2030024) (Journal article)  
[https://github.com/thepanacealab/covid19\\_twitter](https://github.com/thepanacealab/covid19_twitter) (Software)**Communities:**Bio-Hackathon  
Coronavirus Disease Research Community - COVID-19  
Zenodo**License (for files):**[Other \(Public Domain\)](#)**Cite as**

Bandu, Juan M., Tekumalla, Ramya, Wang, Guanyu, Yu, Jingyuan, Liu, Tuo, Ding, Yuning, Artemova, Katya, Tutubalina, Elena, & Chowell, Gerardo. (2023). A large-scale COVID-19 Twitter chatter dataset for open scientific research - an international collaboration [Data set]. In *Epidemiologia* (Version 152, Vol. 2, Number 3, pp. 315–324). Zenodo. <https://doi.org/10.5281/zenodo.7608146>

Start typing a citation style...

1. Publicar en un repositorio público
2. Añadir metadatos mínimos
3. Añadir metadatos de dominio
4. Usar identificadores persistentes
5. Indicar cómo citar el recurso

## Sin curación



## Curados



"Pangaea logo hg" by Hannes Grobe/AWI - Own work. Licensed under CC BY 3.0 via Wikimedia Commons - [http://commons.wikimedia.org/wiki/File:Pangaea\\_logo\\_hg.png#mediaviewer/File:Pangaea\\_logo\\_hg.png](http://commons.wikimedia.org/wiki/File:Pangaea_logo_hg.png#mediaviewer/File:Pangaea_logo_hg.png)

<http://www.arqhys.com/articulos/ingeniero-inspector.html>

February 5, 2023

Dataset Open Access

# A large-scale COVID-19 Twitter chatter dataset for open scientific research - an international collaboration

👤 Banda, Juan M.; 👤 Tekumalla, Ramya; Wang, Guanyu; Yu, Jingyuan; Liu, Tuo; Ding, Yuning; Artemova, Katya; Tutubalina, Elena; 👤 Chowell, Gerardo

Version 152 of the dataset. MAJOR CHANGE NOTE: The dataset files: `full_dataset.tsv.gz` and `full_dataset_clean.tsv.gz` have been split in 1 GB parts using the Linux utility called `Split`. So make sure to join the parts before unzipping. We had to make this change as we had huge issues uploading files larger than 2GB's (hence the delay in the dataset releases). The peer-reviewed publication for this dataset has now been published in *Epidemiologia an MDPI Journal*, and can be accessed here: <https://doi.org/10.3390/epidemiologia2030024>. Please cite this when using the dataset.

File Name	Size
2020-01-04_clean-emoji_char.tsv	11 Bytes
2020-01-04_clean-emoji_text.tsv	21 Bytes
2020-01-06_clean-emoji_char.tsv	1 Byte
2020-01-06_clean-emoji_text.tsv	1 Byte
2020-01-08_clean-emoji_char.tsv	1 Byte
2020-01-08_clean-emoji_text.tsv	1 Byte
2020-01-09_clean-emoji_char.tsv	1 Byte
2020-01-09_clean-emoji_text.tsv	1 Byte
2020-01-10_clean-emoji_char.tsv	1 Byte
2020-01-10_clean-emoji_text.tsv	1 Byte
2020-01-11_clean-emoji_char.tsv	29 Bytes
2020-01-11_clean-emoji_text.tsv	104 Bytes

**Publication date:**

February 5, 2023

**DOI:**

DOI: 10.5281/zenodo.7608146

**Keyword(s):**

social media twitter nlp covid-19 covid19

**Published in:**

Epidemiologia: 2 pp. 315-324 (3).

**Related identifiers:**

Continued by  
<http://www.panacealab.org/covid19/> (Other)  
 Supplement to  
<https://arxiv.org/abs/2004.03688> (Preprint)

**Alternate identifiers:**

10.3390/epidemiologia2030024 (Journal article)  
[https://github.com/thepanacealab/covid19\\_twitter](https://github.com/thepanacealab/covid19_twitter)  
 (Software)

**Communities:**

Bio-Hackathon  
 Coronavirus Disease Research Community - COVID-19  
 Zenodo

**License (for files):**

Other (Public Domain)

**Cite as**

Banda, Juan M., Tekumalla, Ramya, Wang, Guanyu, Yu, Jingyuan, Liu, Tuo, Ding, Yuning, Artemova, Katya, Tutubalina, Elena, & Chowell, Gerardo. (2023). A large-scale COVID-19 Twitter chatter dataset for open scientific research - an international collaboration [Data set]. In *Epidemiologia* (Version 152, Vol. 2, Number 3, pp. 315–324). Zenodo. <https://doi.org/10.5281/zenodo.7608146>

Start typing a citation style...

1. Publicar en un repositorio público
2. **Añadir metadatos mínimos**
3. Añadir metadatos de dominio
4. Usar identificadores persistentes
5. Indicar cómo citar el recurso



Preview

emojis.zip

! The previewer is not showing all the files

- extracted\_elements
  - emojis
    - 2020-01-04\_clean-emoji\_char.tsv 11 Bytes
    - 2020-01-04\_clean-emoji\_text.tsv 21 Bytes
    - 2020-01-06\_clean-emoji\_char.tsv 1 Byte
    - 2020-01-06\_clean-emoji\_text.tsv 1 Byte
    - 2020-01-08\_clean-emoji\_char.tsv 1 Byte
    - 2020-01-08\_clean-emoji\_text.tsv 1 Byte
    - 2020-01-09\_clean-emoji\_char.tsv 1 Byte
    - 2020-01-09\_clean-emoji\_text.tsv 1 Byte
    - 2020-01-10\_clean-emoji\_char.tsv 1 Byte
    - 2020-01-10\_clean-emoji\_text.tsv 1 Byte
    - 2020-01-11\_clean-emoji\_char.tsv 29 Bytes
    - 2020-01-11\_clean-emoji\_text.tsv 104 Bytes

Files (15.9 GB)

**Datos:** los ficheros (tsv, zip, txt, csv, etc)

zenodo

Search

Upload Communities

February 5, 2023

Dataset Open Access

## A large-scale COVID-19 Twitter chatter dataset for open scientific research - an international collaboration

Banda, Juan M.; Tekumalla, Ramya; Wang, Guanyu; Yu, Jingyuan; Liu, Tuo; Ding, Yuning; Artemova, Katya; Tutubalina, Elena; Chowell, Gerardo

*Version 152 of the dataset. MAJOR CHANGE NOTE: The dataset files: full\_dataset.tsv.gz and full\_dataset\_clean.tsv.gz have been split in 1 GB parts using the Linux utility called Split. So make sure to join the parts before unzipping. We had to make this change as we had huge issues uploading files larger than 2GB's (hence the delay in the dataset releases). The peer-reviewed publication for this dataset has now been published in Epidemiologia an MDPI journal, and can be accessed here: <https://doi.org/10.3390/epidemiologia2030024>. Please cite this when using the dataset.*

**Metadatos:** campos describiendo los datos (título, autor, licencia, palabras clave...)

Choose a License

Creative Commons Corporation creativecommons.org/choose/

WINGS WINGS-Portal ODS DII EC ECC ISD ISI

creative commons About Licenses Public Domain Support CC Projects News

### License Features

Your choices on this panel will update the other panels on this page.

**Allow adaptations of your work to be shared?**

Yes  No



Yes, as long as others share alike


**Allow commercial uses of your work?**

Yes  No

### Selected License

## Attribution 4.0 International

This is a Free Culture License! 

### Help others attribute you!

This part is optional, but filling it out will add machine-readable metadata to the suggested HTML!

Title of work

Attribute work to name

Attribute work to URL


Source work URL

More permissions URL

Format of work

License mark

### Have a web page?



This work is licensed under a Creative Commons Attribution 4.0 International License.

**Copy this code to let your visitors know!**

```
<a rel="license"
href="http://creativecommons.org/licenses/by/4.0/">
</a><br />This work is licensed under a <a rel="license"
href="http://creativecommons.org/licenses/by/4.0/">Creativ
a Commons Attribution 4.0 International License</a>
```

Normal icon  Compact icon

## Recomendación: CC-BY and CC0

### CC0 (datasets)



CC0 can be particularly important for the sharing of data and databases, since it otherwise may be unclear whether highly factual data and databases are restricted by copyright or other rights. Databases may contain facts that, in and of themselves, are not protected by copyright law.

CC0 is recommended for data and databases and is used by hundreds of organizations. It is especially recommended for scientific data. Although CC0 doesn't legally require users of the data to cite the source, it does not take away the moral responsibility to give attribution, as is common in scientific research.



### Attribution CC BY

This license lets others distribute, remix, tweak, and build upon your work, even commercially, as long as they credit you for the original creation. This is the most accommodating of licenses offered. Recommended for maximum dissemination and use of licensed materials.

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

February 5, 2023

Dataset Open Access

# A large-scale COVID-19 Twitter chatter dataset for open scientific research - an international collaboration

Bandu, Juan M.; Tekumalla, Ramya; Wang, Guanyu; Yu, Jingyuan; Liu, Tuo; Ding, Yuning; Artemova, Katya; Tutubalina, Elena; Chowell, Gerardo

Version 152 of the dataset. MAJOR CHANGE NOTE: The dataset files: `full_dataset.tsv.gz` and `full_dataset_clean.tsv.gz` have been split in 1 GB parts using the Linux utility called `Split`. So make sure to join the parts before unzipping. We had to make this change as we had huge issues uploading files larger than 2GB's (hence the delay in the dataset releases). The peer-reviewed publication for this dataset has now been published in *Epidemiologia an MDPI Journal*, and can be accessed here: <https://doi.org/10.3390/epidemiologia2030024>. Please cite this when using the dataset.

**Publication date:**

February 5, 2023

**DOI:**DOI: [10.5281/zenodo.7608146](https://doi.org/10.5281/zenodo.7608146)**Keyword(s):**

social media twitter nlp covid-19 covid19

**Published in:**

Epidemiologia: 2 pp. 315-324 (3).

**Related identifiers:**Continued by  
<http://www.panacealab.org/covid19/> (Other)Supplement to  
<https://arxiv.org/abs/2004.03688> (Preprint)**Alternate identifiers:**[10.3390/epidemiologia2030024](https://doi.org/10.3390/epidemiologia2030024) (Journal article)  
[https://github.com/thepanacealab/covid19\\_twitter](https://github.com/thepanacealab/covid19_twitter)  
(Software)**Communities:**BioHackathon  
Coronavirus Disease Research Community -  
COVID-19  
Zenodo**License (for files):**[Other \(Public Domain\)](#)**Cite as**

Bandu, Juan M., Tekumalla, Ramya, Wang, Guanyu, Yu, Jingyuan, Liu, Tuo, Ding, Yuning, Artemova, Katya, Tutubalina, Elena, & Chowell, Gerardo. (2023). A large-scale COVID-19 Twitter chatter dataset for open scientific research - an international collaboration [Data set]. In *Epidemiologia* (Version 152, Vol. 2, Number 3, pp. 315–324). Zenodo. <https://doi.org/10.5281/zenodo.7608146>

Start typing a citation style...

1. Publicar en un repositorio público
2. Añadir metadatos mínimos
3. **Añadir metadatos de dominio**
4. Usar identificadores persistentes
5. Indicar cómo citar el recurso

- Nombre/título
- Descripción
- Autor
- Fecha de publicación
- Licencia
- Contacto
- Versión
- Tipo de recurso
- Lugar de recogida

**General**

- Palabras clave de dominio
- Preprocesado
- Variables
- Visualización

**Específico**

Estándares de dominio se usan para describir datos específicos



Global climate observing system

February 5, 2023

Dataset Open Access

# A large-scale COVID-19 Twitter chatter dataset for open scientific research - an international collaboration

👤 Banda, Juan M.; 👤 Tekumalla, Ramya; Wang, Guanyu; Yu, Jingyuan; Liu, Tuo; Ding, Yuning; Artemova, Katya; Tutubalina, Elena; 👤 Chowell, Gerardo

Version 152 of the dataset. MAJOR CHANGE NOTE: The dataset files: `full_dataset.tsv.gz` and `full_dataset_clean.tsv.gz` have been split in 1 GB parts using the Linux utility called `Split`. So make sure to join the parts before unzipping. We had to make this change as we had huge issues uploading files larger than 2GB's (hence the delay in the dataset releases). The peer-reviewed publication for this dataset has now been published in *Epidemiologia an MDPI Journal*, and can be accessed here: <https://doi.org/10.3390/epidemiologia2030024>. Please cite this when using the dataset.

**Publication date:**

February 5, 2023

**DOI:**

DOI: 10.5281/zenodo.7608146

**Keyword(s):**

social media twitter nlp covid-19 covid19

**Published in:**

Epidemiologia: 2 pp. 315-324 (3).

**Related identifiers:**Continued by  
<http://www.panacealab.org/covid19/> (Other)

## Supplement to

<https://arxiv.org/abs/2004.03688> (Preprint)**Alternate identifiers:**10.3390/epidemiologia2030024 (Journal article)  
[https://github.com/thepanacealab/covid19\\_twitter](https://github.com/thepanacealab/covid19_twitter)  
(Software)**Communities:**Bio-Hackathon  
Coronavirus Disease Research Community - COVID-19  
Zenodo**License (for files):**

Other (Public Domain)

**Cite as**

Banda, Juan M., Tekumalla, Ramya, Wang, Guanyu, Yu, Jingyuan, Liu, Tuo, Ding, Yuning, Artemova, Katya, Tutubalina, Elena, & Chowell, Gerardo. (2023). A large-scale COVID-19 Twitter chatter dataset for open scientific research - an international collaboration [Data set]. In *Epidemiologia* (Version 152, Vol. 2, Number 3, pp. 315–324). Zenodo. <https://doi.org/10.5281/zenodo.7608146>

Start typing a citation style...

1. Publicar en un repositorio público
2. Añadir metadatos mínimos
3. Añadir metadatos de dominio
4. Usar identificadores persistentes
5. Indicar cómo citar el recurso



February 5, 2023

Dataset Open Access

# A large-scale COVID-19 Twitter chatter dataset for open scientific research - an international collaboration

Bandu, Juan M.; Tekumalla, Ramya; Wang, Guanyu; Yu, Jingyuan; Liu, Tuo; Ding, Yuning; Artemova, Katya; Tutubalina, Elena; Chowell, Gerardo

Version 152 of the dataset. **MAJOR CHANGE NOTE:** The dataset files: `full_dataset.tsv.gz` and `full_dataset_clean.tsv.gz` have been split in 1 GB parts using the Linux utility called `Split`. So make sure to join the parts before unzipping. We had to make this change as we had huge issues uploading files larger than 2GB's (hence the delay in the dataset releases). The peer-reviewed publication for this dataset has now been published in *Epidemiologia an MDPI Journal*, and can be accessed here: <https://doi.org/10.3390/epidemiologia2030024>. Please cite this when using the dataset.

File Name	Size
2020-01-04_clean-emoji_char.tsv	11 Bytes
2020-01-04_clean-emoji_text.tsv	21 Bytes
2020-01-06_clean-emoji_char.tsv	1 Byte
2020-01-06_clean-emoji_text.tsv	1 Byte
2020-01-08_clean-emoji_char.tsv	1 Byte
2020-01-08_clean-emoji_text.tsv	1 Byte
2020-01-09_clean-emoji_char.tsv	1 Byte
2020-01-09_clean-emoji_text.tsv	1 Byte
2020-01-10_clean-emoji_char.tsv	1 Byte
2020-01-10_clean-emoji_text.tsv	1 Byte
2020-01-11_clean-emoji_char.tsv	29 Bytes
2020-01-11_clean-emoji_text.tsv	104 Bytes

**Publication date:**

February 5, 2023

**DOI:**

DOI: 10.5281/zenodo.7608146

**Keyword(s):**

social media twitter nlp covid-19 covid19

**Published in:**

Epidemiologia: 2 pp. 315-324 (3).

**Related identifiers:**Continued by  
<http://www.panacealab.org/covid19/> (Other)

## Supplement to

<https://arxiv.org/abs/2004.03688> (Preprint)**Alternate identifiers:**10.3390/epidemiologia2030024 (Journal article)  
[https://github.com/thepanacealab/covid19\\_twitter](https://github.com/thepanacealab/covid19_twitter)  
(Software)**Communities:**Bio-Hackathon  
Coronavirus Disease Research Community - COVID-19  
Zenodo**License (for files):**

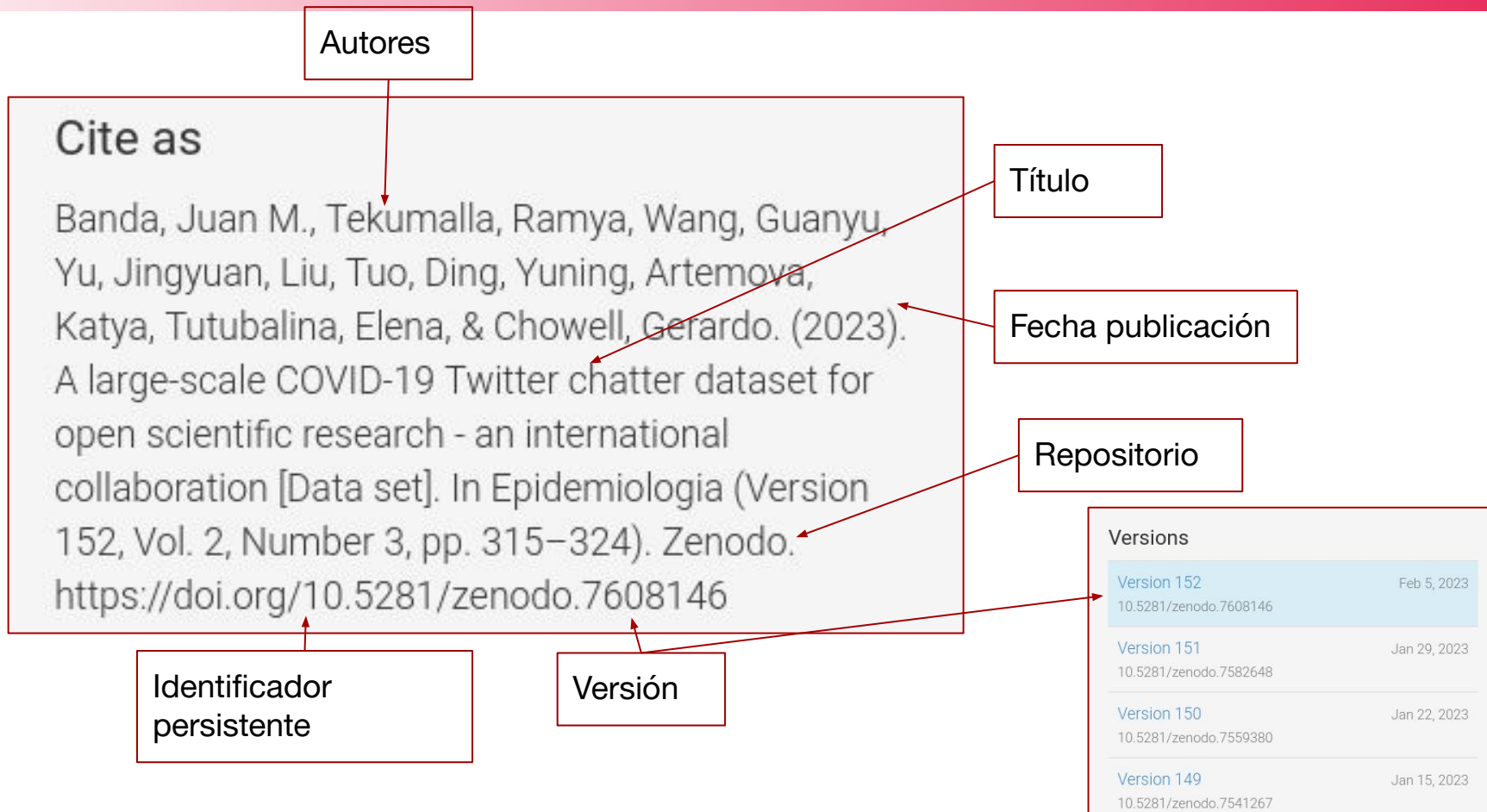
Other (Public Domain)

**Cite as**

Bandu, Juan M., Tekumalla, Ramya, Wang, Guanyu, Yu, Jingyuan, Liu, Tuo, Ding, Yuning, Artemova, Katya, Tutubalina, Elena, & Chowell, Gerardo. (2023). A large-scale COVID-19 Twitter chatter dataset for open scientific research - an international collaboration [Data set]. In *Epidemiologia* (Version 152, Vol. 2, Number 3, pp. 315–324). Zenodo. <https://doi.org/10.5281/zenodo.7608146>

Start typing a citation style...

1. Publicar en un repositorio público
2. Añadir metadatos mínimos
3. Añadir metadatos de dominio
4. Usar identificadores persistentes
5. Indicar cómo citar el recurso



Indica cómo los datos de investigación **serán gestionados durante y tras la finalización de un proyecto**

Requisito de las agencias de financiación!

### Secciones de un plan de gestión de datos:

- Descripción/resumen
- Problemas éticos y legales (GDPR, etc)
- Documentación y metadatos (interoperabilidad)
- Recursos (costes)
- Responsabilidades (quién hace qué)
- Políticas para compartir datos (licencias)
- Almacenamiento de datos durante el proyecto
- Preservación de recursos a futuro





















source: <https://rdmkit.elixir-europe.org/>



## Funder Templates

Templates are provided by a funder.

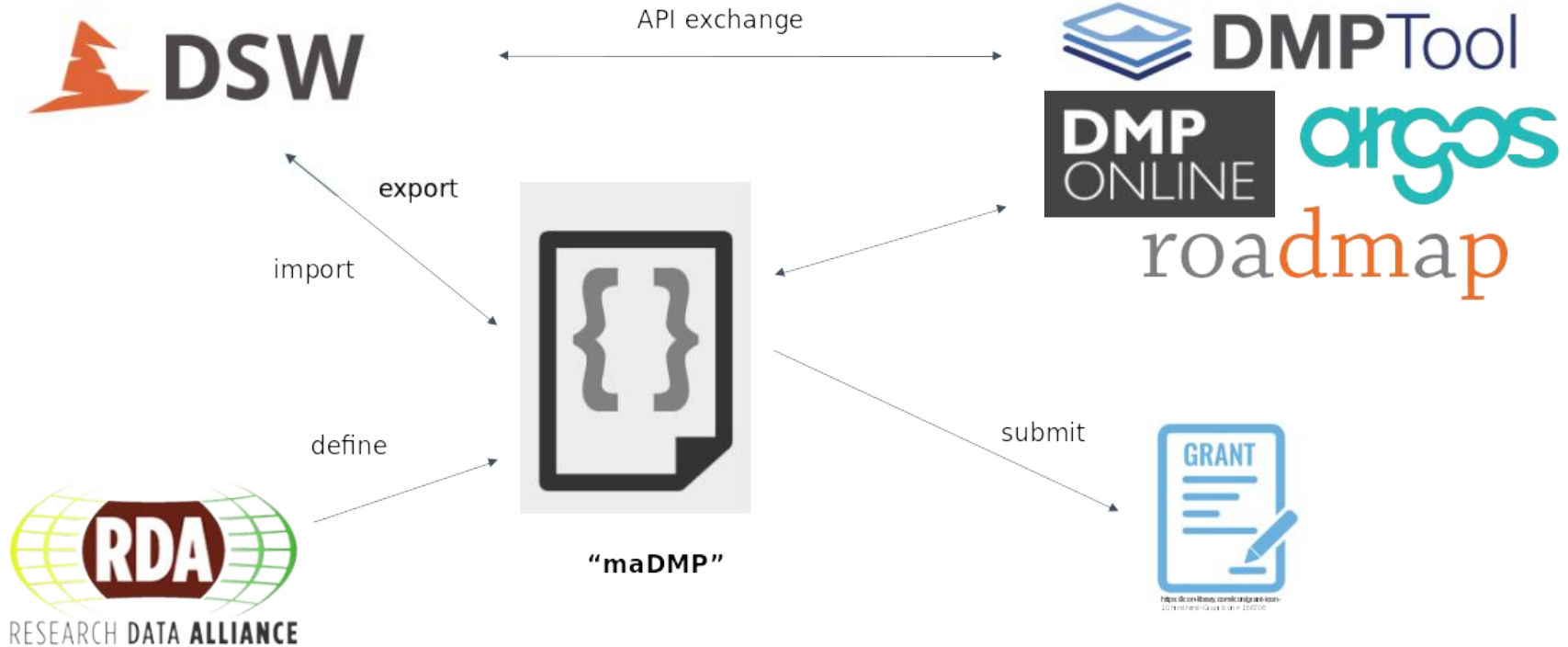
Templates for data management plans are based on the specific requirements listed in funder policy documents. DMPonline.be maintains these templates, however, researchers should always consult the funder guidelines directly for authoritative information.

Template Name	Download	Organisation Name	Last Updated	Funder Links	Sample Plans (if available)
BELSPO DMP +	 	Belgian Federal Science Policy Office (BELSPO)	27-09-2021		
ERC DMP +	 	European Research Council (ERC)	27-09-2021		
DCC Template	 	Digital Curation Centre	27-09-2021		
Horizon 2020 FAIR DMP +	 	European Commission (Horizon)	27-09-2021		
BRAIN 2.0	 	Belgian Federal Science Policy Office (BELSPO)	13-05-2022	<a href="http://www.belspo.be">www.belspo.be</a>	
FNRS DMP	 	Fonds National de la Recherche Scientifique (FNRS)	19-05-2022		
VLAIO cSBO DMP (Flemish Standard DMP)	 	Vlaams Agentschap Innoveren & Ondernemen (VLAIO)	02-09-2022		
Horizon Europe DMP +	 	European Commission (Horizon)	12-10-2022		
FWO DMP (Flemish Standard DMP)	 	Fonds voor Wetenschappelijk Onderzoek - Research Foundation Flanders (FWO)	24-10-2022		

**Fuente: DMPonline.be**

Inspirado en: Alexander Botzki et al. RltrainPlus: Data Management Plans and Co. Ack: Esteban Gonzalez Guardia.

Datos de investigación: Experiencias, retos y oportunidades ante la reforma de la evaluación de la investigación



Inspirado en: <https://doi.org/10.7490/f1000research.1117954.1>

## La Ciencia Abierta va más allá de “datos”:

- Buenas prácticas para **Software Científico**:
  - <https://fair-impact.github.io/RSMD-guidelines/>
- **Planes de gestión de Software Científico**:
  - [https://nfdi4ds.fokus.fraunhofer.de/news/202402\\_smp/](https://nfdi4ds.fokus.fraunhofer.de/news/202402_smp/)
- Buenas prácticas para **Ontologías, vocabularios**:
  - <https://zenodo.org/records/8305173>
- **Hardware**:
  - <https://archive.rd-alliance.org/groups/fair-principles-research-hardware>
- **Modelos** de Aprendizaje Automático
  - <https://www.rd-alliance.org/groups/fair-machine-learning-fair4ml-ig>
- **Reproducibilidad**:
  - <https://www.rd-alliance.org/groups/reproducibility-ig/work-statement/>
- ...

- **Evaluación científica:** ¿Cómo hemos llegado hasta aquí?
  - DORA, Leiden Manifesto, CoARA
  - Indicadores dan paso a narrativa
  - Reconocimiento más allá de artículos (comienzo)
- Datos de investigación: ¿Qué son y qué **buenas prácticas hay**?
  - Datos FAIR
  - Planes de gestión de datos
- ¿Cómo **adoptar buenas prácticas** invirtiendo el **mínimo de tiempo**?
  - Hay recursos más allá de los datos! (Software, modelos, Workflows)
  - Recursos FAIR equivale más visibilidad, cumplir con los requisitos de las agencias y facilitar aprendizaje de estudiantes.



# Datos de investigación: retos, oportunidades y buenas prácticas ante la reforma de la evaluación de la investigación

**Daniel Garijo**

Universidad Politécnica de Madrid

[daniel.garijo@upm.es](mailto:daniel.garijo@upm.es)

@dgarijov

**Madrid, 24 de octubre de 2024**

<https://doi.org/10.5281/zenodo.13971475>