Curso gestión de datos ISCIII

Semana 1. Lunes y miércoles 28 y 30 de junio de 10 a 13 h (o de 9 a 12) y viernes 2 de julio de 11- 13 h repaso y dudas

Semana 2. Lunes y miércoles 5 y 7 de julio de 10 a 13 h (o de 9 a 12) y viernes 9 de julio de 11- 13 h repaso y dudas

Remedios Melero. IATA-CSIC, Valencia, España

CONTENIDOS / PROGRAMA:

Tema 1 : Fundamentos de la gestión de datos de investigación.

- Los datos de investigación: Definiciones y tipología.
- Ciencia abierta y datos abiertos. Datos FAIR.
- Ventajas y dificultades de poner datos en abierto.

Tema 2: Aspectos legales y éticos.

- Requerimientos en la gestión de datos de investigación.
- Normativa de GDPR.
- Licencias para compartir datos.

Tema 3: El ciclo de vida de los datos.

- Organización y descripción.
- Almacenaje y preservación.
- Publicación y difusión.

Tema 4: El plan de gestión de datos (PGD).

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

Práctica: Elaboración de un plan de gestión de datos.

CAMPUS

07 juni 2021 · 09:00 door Heather Montague

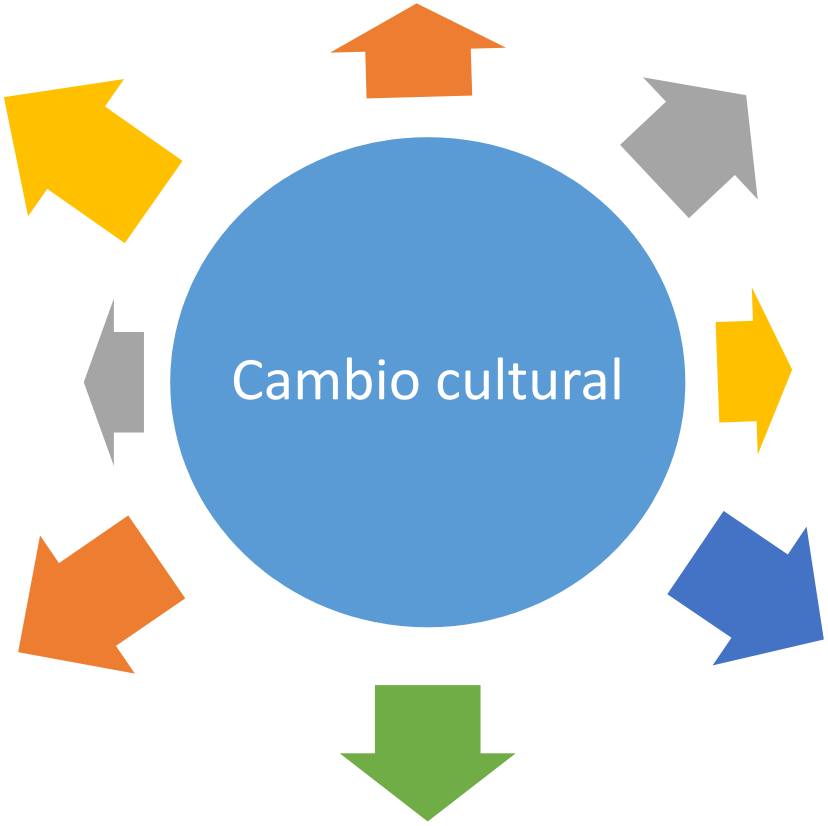
‘Open science is just science done right’

Anke Versteeg: “Open science is a means to work together with other disciplines more easily, to share your work with others, to work more in teams, to work more internationally.” (Photo: Nicole Will)

Open science is just science done right, according to Anke Versteeg. She says we should be proud to share what’s going on here at TU Delft.

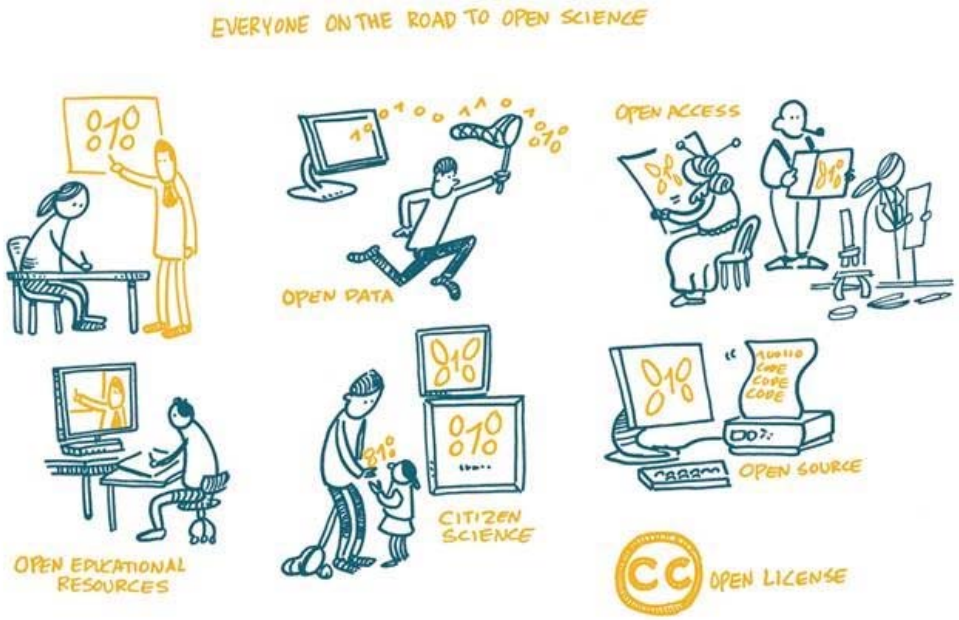
“That means that you are managing all of your work right from the start, managing your data, thinking about licences, your own rights, about connecting with others. You have to do a lot of research management, which helps you in the end.”

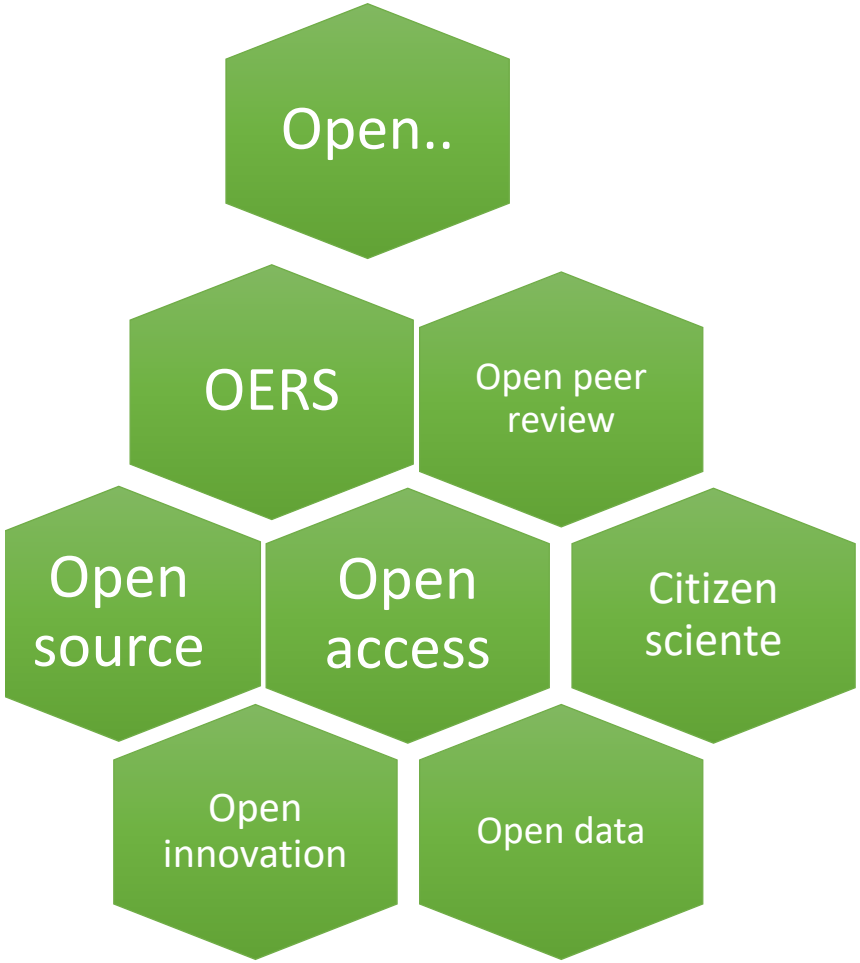
Open science… una visión holística

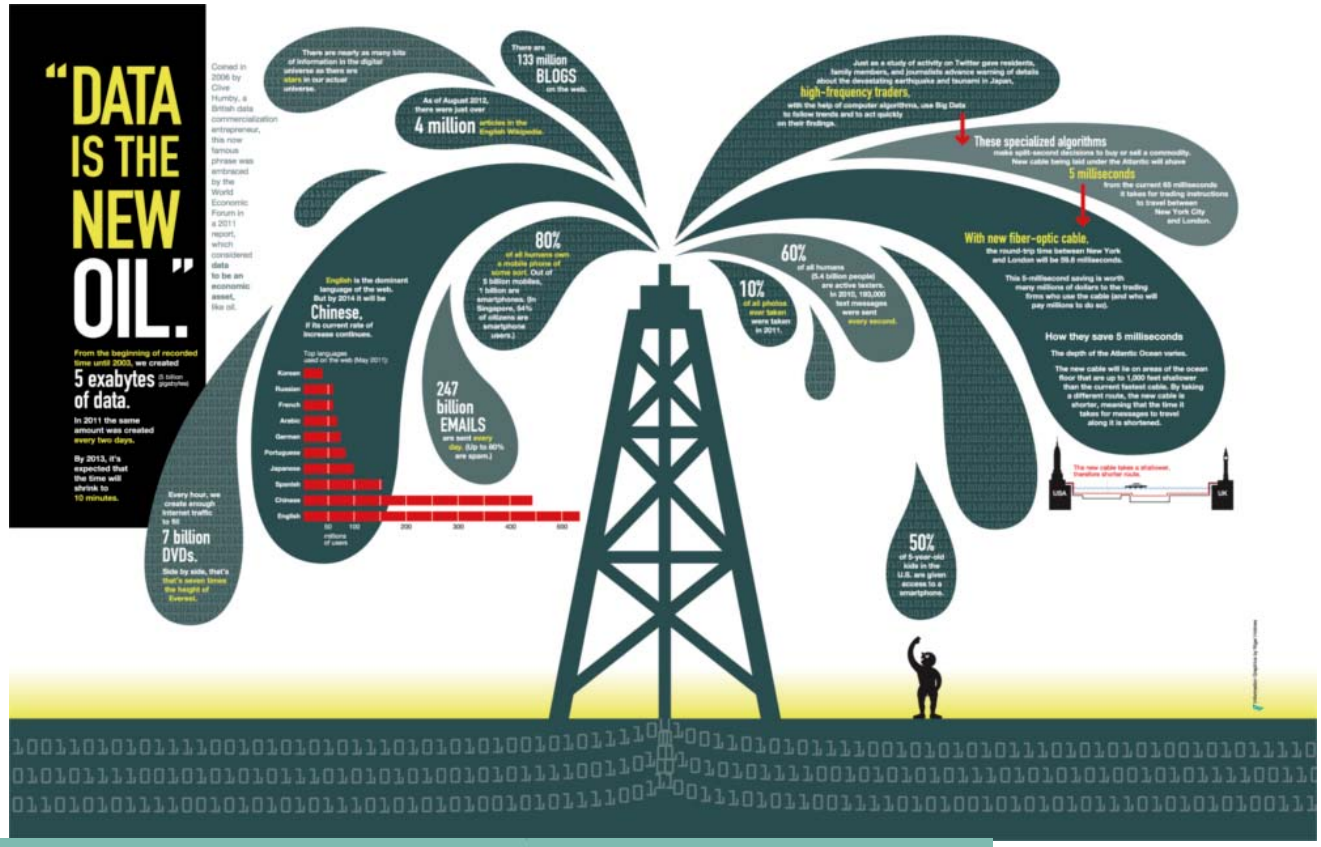


“Contribuye en tu institución o centro de investigación a implementar el paradigma de Open Science, desde una visión holística, integrando todas las partes y procesos que afecten a la producción, comunicación y preservación de los resultados de la investigación financiada con fondos públicos” (directrices Maredata)

Open science






"DATA IS THE NEW OIL."
Coined in 2006 by Clive Humby, a British data commercialization entrepreneur, this now famous phrase was embraced by the World Economic Forum in a 2011 report, which considered data to be an economic asset, like oil.
From the beginning of recorded time until 2003, we created 5 exabytes (5 billion gigabytes) of data.
In 2011 the same amount was created every two days.
By 2013, it's expected that the time will shrink to 10 minutes.
There are nearly as many bits of information in the digital universe as there are stars in our actual universe.
As of August 2012, there were just over 4 million articles in the English Wikipedia.
There are 133 million BLOGS on the web.
Just as a study of activity on Twitter gave residents, family members, and journalists advance warning of details about the devastating earthquake and tsunami in Japan,
high-frequency traders,
with the help of computer algorithms, use Big Data to follow trends and to act quickly on their findings.
These specialized algorithms
make split-second decisions to buy or sell a commodity.
New cable being laid under the Atlantic will shave
5 milliseconds
from the current 65 milliseconds it takes for trading instructions to travel between New York City and London.
With new fiber-optic cable,
the round-trip time between New York and London will be 59.6 milliseconds.
This 5-millisecond saving is worth many millions of dollars to the trading firms who use the cable (and who will pay millions to do so).
How they save 5 milliseconds
The depth of the Atlantic Ocean varies.
The new cable will lie on areas of the ocean floor that are up to 1,000 feet shallower than the current fastest cable. By taking a different route, the new cable is shorter, meaning that the time it takes for messages to travel along it is shortened.
The new cable takes a shallower, shorter route.
USA
UK
80%
of all humans own a mobile phone of some sort. Out of 5 billion mobiles, 1 billion are smartphones. (In Singapore, 94% of citizens are smartphone users.)
247 billion EMAILS
are sent every day. (Up to 90% are spam.)
English is the dominant language of the web. But by 2014 it will be Chinese, if its current rate of increase continues.
Top languages used on the web (May 2011)
Korean
Russian
French
Arabic
German
Portuguese
Japanese
Spanish
Chinese
English
0
50
100
200
300
400
600
millions of users
Every hour, we create enough Internet traffic to fill 7 billion DVDs. Side by side, that's that's seven times the height of Everest.
10% of all photos ever taken were taken in 2011.
60% of all humans (4.4 billion people) are active texters. In 2010, 193,000 text messages were sent every second.
50% of 5-year-old kids in the U.S. are given access to a smartphone.

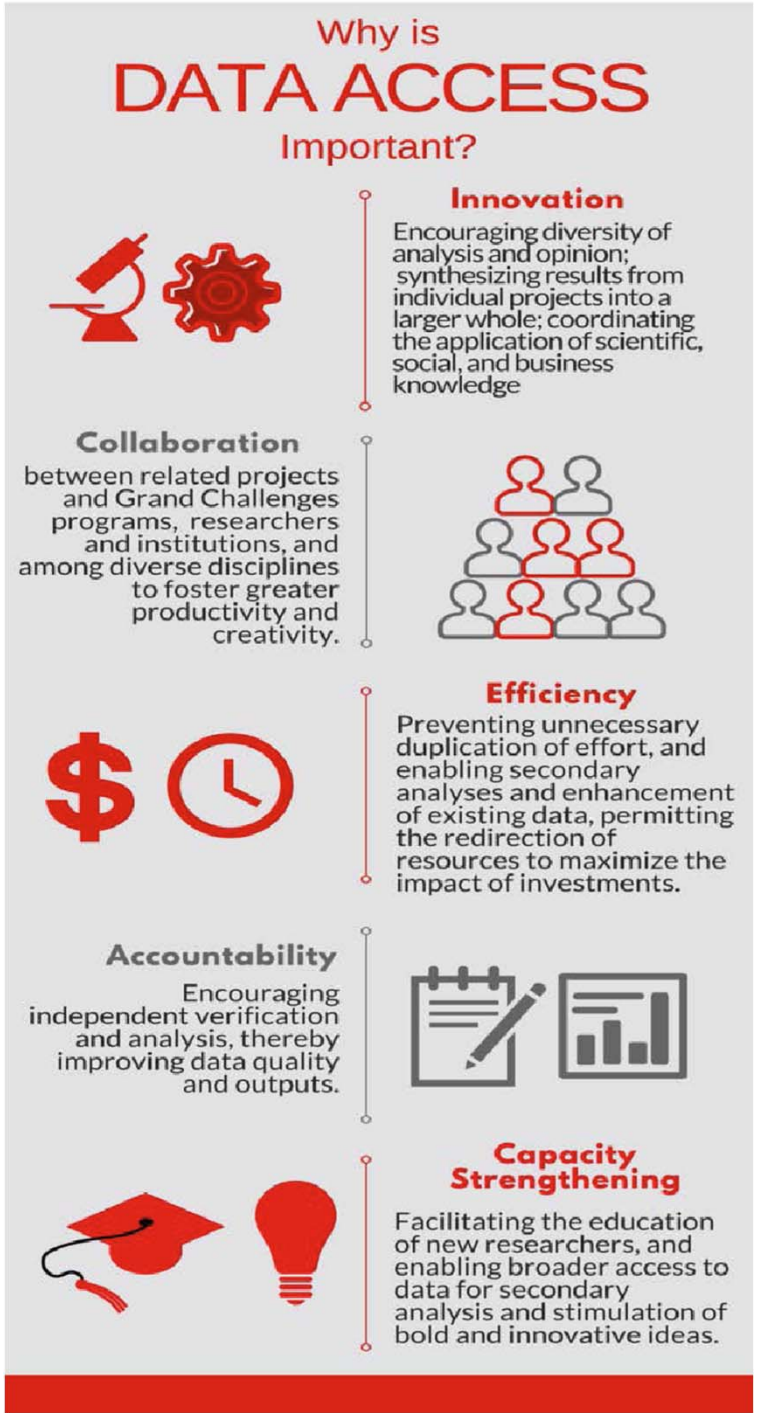

S
DATA
IS THE NEW OIL
but do you have the resource to refine it?

Algunas razones para compartir datos:

- Promueve la innovación y potenciales nuevos usos
- Conduce a la colaboración entre ususarios de datos y los creadores de datos
- Maximiza la transparencia y la fiabilidad
- Permite la verificación de los resultados de investigación
- Reduce costes al evitar duplicación de datos
- Aumenta el impacto y la visibilidad de la investigación
- Promueve la investigación de donde salieron los datos y sus publicaciones
- Puede generar un reconocimeinto directo a los investigadores como cualquier otro resultado de la investigación
- Genera nuevos datos a partir de los originales

10 Simple Rules for the Care and Feeding of Scientific Data
Alyssa Goodman et al. (2013)
http://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1003542

Regla 1. Amar tus datos y ayudar a los demás a amarlos también
Regla 2. Compartir tus datos online con un identificador permanente
Regla 3. Hacer ciencia con el concepto de reutilización en mente
Regla 4. Publicar el contexto de obtención de los datos
Regla 5. Vincular sus datos con sus publicaciones siempre que sea posible
Regla 6. Publicar el código fuente (incluso los pequeños *bits*)
Regla 7. Decidir cómo quiere que se le reconozca
Regla 8. Fomentar y utilizar repositorios de datos
Regla 9. Recompensar a los colegas que comparten sus datos adecuadamente
Regla 10. Ser impulsor de la ciencia de los datos

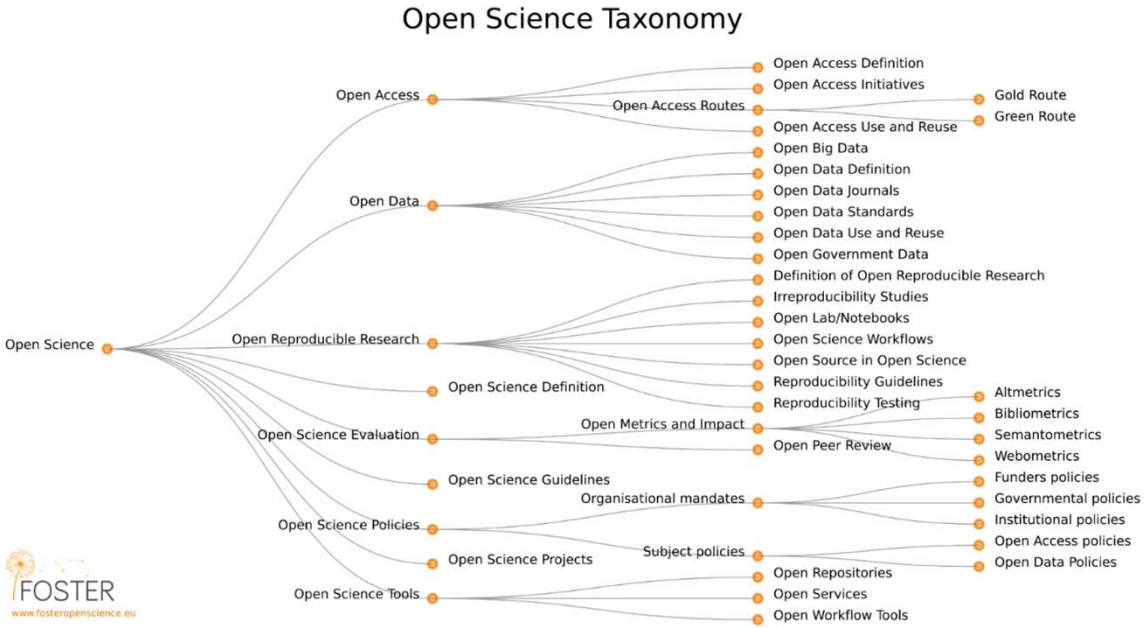


www.grandchallenges.ca/wp-content/uploads/20161006_GCC-Open-Access-Guide-for-Innovators_EN.pdf

Por qué……

- Innovación
- Colaboración
- Eficiencia
- Responsabilidad
- Transparencia
- Generación de nuevo conocimiento

Contexto



https://www.fosteropenscience.eu/taxonomy/term/7

Ley Propiedad Intelectual (LPI)

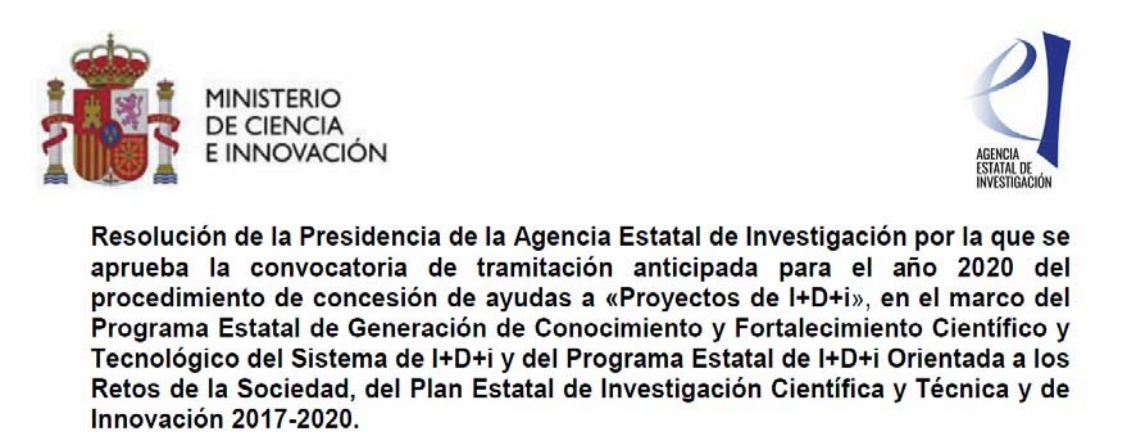
Artículo 10. **Obras y títulos originales**.

1. *Son objeto de propiedad intelectual todas las* ***creaciones originales literarias, artísticas o científicas*** *expresadas por cualquier medio o soporte, tangible o intangible, actualmente conocido o que se invente en el futuro, comprendiéndose entre ellas:……* ***no se habla de datos, sí se menciona las bases de datos (derecho sui generis)***

Motores para compartir y gestionar los dtos de investigación

2. TRANSPARENCIA Y RENDICIÓN DE CUENTAS de las ayudas concedidas y del proceso de concesión de las mismas, incluyendo:

- …
- ACCESO ABIERTO A RESULTADOS Y DATOS DE INVESTIGACIÓN de las actividades de investigación subvencionadas con recursos públicos. **Los trabajos publicados en revistas científicas financiados a través del Plan Estatal se depositarán en repositorios, institucionales y/o internacionales, en abierto** teniendo en cuenta las características específicas de las distintas materias, en cumplimiento de lo establecido en el Artículo 37 de **la Ley 14/2011, de 1 de junio, de la Ciencia, la Tecnología y la Innovación y de las recomendaciones vinculadas a la agenda europea en materia de acceso abierto y ciencia en abierto**. Con el fin de impulsar el acceso a datos de investigación, los proyectos de I+D+i financiados **podrán incluir, con carácter optativo, un plan de gestión de los datos de investigación** que se depositarán en repositorios institucionales, nacionales y/o internacionales tras la finalización del proyecto y trascurrido el plazo establecido en las correspondientes convocatorias. No obstante, se respetarán todas las situaciones en las que los mismos han de protegerse por razones de confidencialidad, seguridad, protección, etc., o cuando los mismos sean necesarios para la explotación comercial de los resultados obtenidos. Finalmente, **en la evaluación curricular de los investigadores así como en la evaluación ex post de las actuaciones financiadas se tendrán en cuenta los trabajos publicados en abierto en repositorios institucionales y temáticos, nacionales y/o internacionales, y la puesta de los datos de su investigación en abierto,** de modo que puedan ser utilizados para replicar y reproducir los análisis y resultados de investigación.


MINISTERIO
DE CIENCIA
E INNOVACIÓN

AGENCIA ESTATAL DE INVESTIGACIÓN

Resolución de la Presidencia de la Agencia Estatal de Investigación por la que se aprueba la convocatoria de tramitación anticipada para el año 2020 del procedimiento de concesión de ayudas a «Proyectos de I+D+i», en el marco del Programa Estatal de Generación de Conocimiento y Fortalecimiento Científico y Tecnológico del Sistema de I+D+i y del Programa Estatal de I+D+i Orientada a los Retos de la Sociedad, del Plan Estatal de Investigación Científica y Técnica y de Innovación 2017-2020.

Enlace (11-11-2020)

Obligaciones de las entidades beneficiarias

Artículo 19

…

Los datos de investigación se deberán depositar en repositorios institucionales, nacionales y/o internacionales **antes de que transcurran dos años desde la finalización del proyecto**, con el fin de impulsar el acceso a datos de investigación de las ayudas financiadas.

3. Toda referencia a alguno de los proyectos objeto de las presentes ayudas en cualquier medio de difusión deberá incluir que el mismo ha sido financiado por la Agencia Estatal de Investigación. En el caso de que existan limitaciones de espacio, particularmente en publicaciones, se mencionará de la forma indicada en el artículo 9.3.a) (REFERENCIA DEL PROYECTO / AEI / 10.13039/501100011033).

Aprox 500 palabras!!!

3. IMPACTO ESPERADO DE LOS RESULTADOS

En la aplicación informática de solicitud deberá introducir una descripción de un máximo de **3.500 caracteres** sobre el impacto esperado de los resultados del proyecto, cuyo contenido podrá ser publicado a efectos de difusión si el proyecto fuera financiado en esta convocatoria.

En este apartado se recomienda incluir:

a)

b) ..y/o resultados susceptibles de **transferencia**

Previsión del plan de gestión de datos de investigación en el que se indique qué datos se van a recoger o generar (tipologías y formatos), cómo será el acceso (quién, cómo y cuándo se podrá acceder a los datos) y en qué repositorio está previsto que se depositen. En el caso de datos que estén sometidos a la reglamentación de protección de los datos personales o de aspectos éticos, indicar cómo se gestionarán. En el caso de los proyectos que resulten financiados, durante la ejecución del proyecto y junto al informe final se podrá solicitar un plan de gestión de datos formal completo

ESTRATEGIA ESPAÑOLA DE CIENCIA, TECNOLOGÍA E INNOVACIÓN 2021-2027
https://www.ciencia.gob.es/stfls/MICINN/Ministerio/FICHEROS/EECTI-2021-2027.pdf

Principios…(p.24)
La Responsabilidad social y económica de la I+D+I a través de la **incorporación de la ciencia ciudadana y la aplicación de la co-creación y las políticas de acceso abierto**, así como el alineamiento de la I+D+I con los valores, necesidades y expectativas sociales.

(p.32) La ciencia excelente y abierta constituye uno de los pilares del *Objetivo 4* (Generación de conocimiento y liderazgo científico). **El impulso a un modelo de Ciencia Abierta favorecerá la generación de conocimiento de alta calidad e impacto, así como su transmisión a la sociedad**, elemento que se relaciona directamente con el *Eje de actuación 5* (fomentar y apoyar la generación de capacidades científicas e innovadoras en los agentes del SECTI).

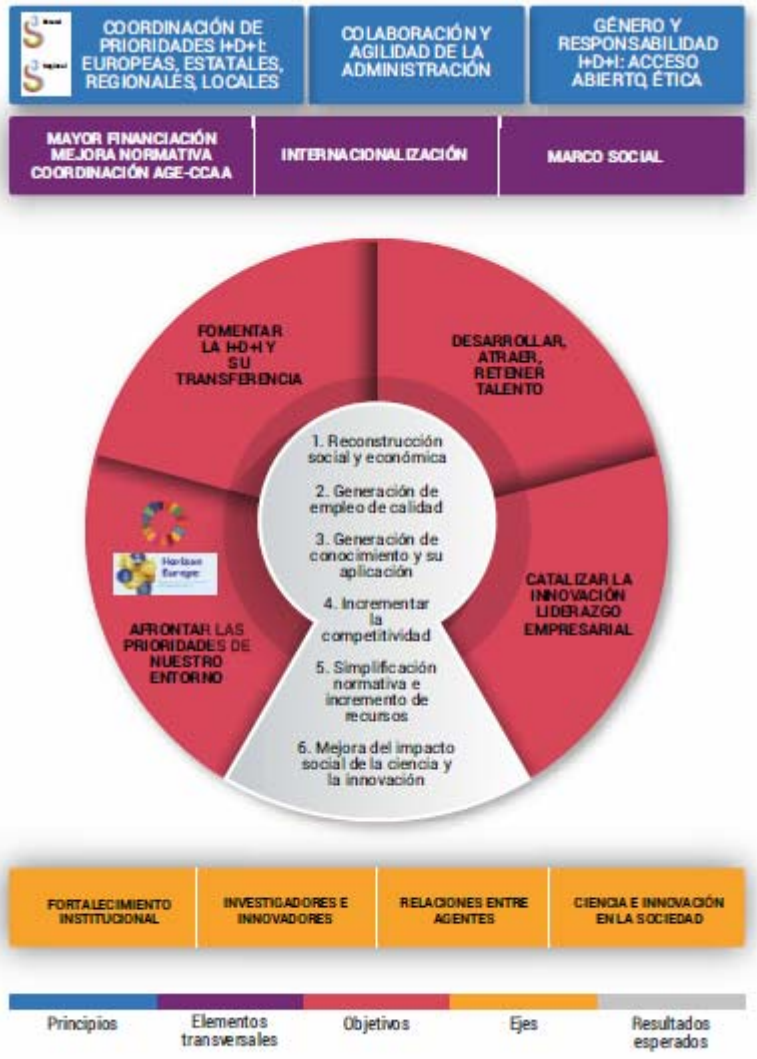


Figura 3. Modelo de actuación de la EECTI 2021-2027

(p.33) La coordinación de los agentes financiadores permitirá mejorar los instrumentos de financiación y los mecanismos de articulación de los PEICTI, **empleando para ello recursos humanos con formación en el acceso abierto a datos, microdatos, publicaciones, código (software) y, en general, a todos los resultados de la investigación financiada con fondos públicos**

(p. 36) El apoyo a la ciencia de excelencia, elemento básico del modelo de EECTI, fomentará, **en consonancia con las directrices de la UE, el acceso abierto a los resultados de investigación, permitiendo que los datos sean accesibles, interoperables y reutilizables** (su acrónimo en inglés *FAIR*). La difusión en el ámbito científico, junto al esfuerzo llevado a cabo por los repositorios abiertos, facilitará la accesibilidad a los avances científicos y fomentará la divulgación y comunicación científica hacia la sociedad, objetivo que se persigue en el *Eje de actuación 14……*

Este nuevo escenario, en donde **el dato científico adquiere un valor relevante**, estará acompañado de otras actuaciones que permitan, por ejemplo, desarrollar infraestructuras de datos y sus servicios, apoyando así la producción de datos FAIR y su integración en la Nube Europea,

Tres disposiciones a tener en cuenta de alta relevancia….

RECOMENDACIÓN (UE) 2018/790 DE LA COMISIÓN. de 25 de abril de 2018 relativa al acceso a la información científica y a su preservación.
https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32018H0790

- Acceso abierto a las publicaciones científicas
- **Gestión de los datos de investigación, incluyendo el acceso abierto**
- Preservación y reutilización de la información científica
- Infraestructuras para la ciencia abierta
- Capacidades y competencias
- Incentivos y recompensas
- Diálogo multilateral sobre la ciencia abierta a nivel nacional, europeo e internacional
- Coordinación estructurada de los Estados miembros a nivel de la UE y seguimiento de la presente Recomendación

"**Los Estados miembros deben establecer y aplicar políticas claras (tal y como se describe en los planes nacionales de acción) para la gestión de los datos resultantes de la investigación financiada con fondos públicos, incluido el acceso abierto**. "

Resolución legislativa del Parlamento Europeo, de 4 de abril de 2019, sobre la Directiva del Parlamento Europeo y del Consejo relativa a la reutilización de la información del sector público (versión refundida) (COM(2018)0234 – C8-0169/2018 – 2018/0111(COD))

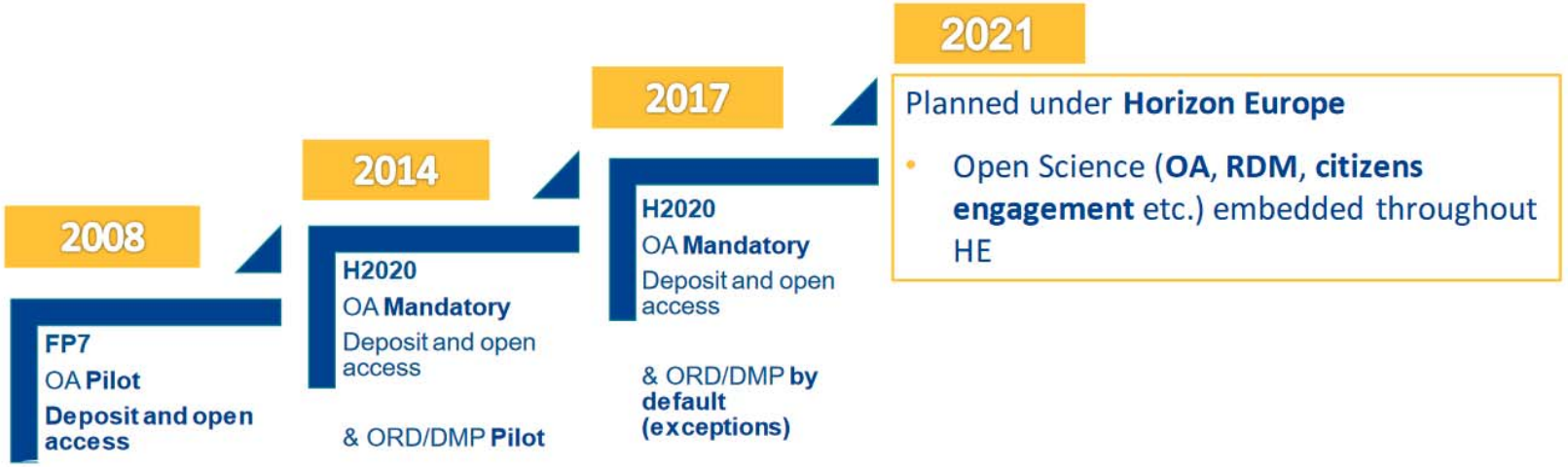
Tendrá que trasponerse en el plazo de 2 años y modificar la Ley 18/2015, de 9 de julio, sobre reutilización de la información del sector público (trasposición de la directiva 2013/37/UE).

"Con arreglo a las políticas nacionales de acceso abierto, los datos de la investigación financiada públicamente serán abiertos por defecto. Sin embargo, en este contexto, deben tenerse debidamente en cuenta las inquietudes relacionadas con la privacidad, la protección de datos personales, *la confidencialidad*, la seguridad nacional, los intereses comerciales legítimos, *como los secretos comerciales,* y los derechos de propiedad intelectual de terceros, *conforme al principio «tan abiertos como sea posible, tan cerrados como sea necesario». Por otra parte, los datos de investigación a los que no puede accederse por motivos de seguridad nacional, defensa o seguridad pública no deben estar cubiertos por la presente Directiva*."

……..

"Los Estados miembros pueden hacer extensiva la aplicación de la presente Directiva a los datos de investigación puestos a disposición del público a través de ***infraestructuras de datos distintas de los registros, a través de publicaciones de acceso abierto o en forma de fichero adjunto a un artículo, a un artículo de datos o a un artículo en una revista especializada en datos.*** *Los documentos distintos de los datos de investigación* deben seguir estando excluidos del ámbito de la presente Directiva."

Evolution of open science policy across FPs



European Commission

https://ec.europa.eu/research/participants/docs/h2020-funding-guide/other/event201009.htm

REGLAMENTO (UE) 2021/695 DEL PARLAMENTO EUROPEO Y DEL CONSEJO
de 28 de abril de 2021, por el que se crea el Programa Marco de Investigación e Innovación «Horizonte Europa», se establecen sus normas de participación y difusión, y se derogan los Reglamentos (UE) n.o 1290/2013 y (UE) n.o 1291/2013
https://eur-lex.europa.eu/legal-content/ES/TXT/HTML/?uri=CELEX:32021R0695&from=EN

Artículo 14
Ciencia abierta

1. El Programa fomentará la ciencia abierta como enfoque del proceso científico basado en el trabajo de cooperación y la difusión de conocimientos, en particular de conformidad con los siguientes elementos que se garantizarán conforme al artículo 39, apartado 3, del presente Reglamento:

a) acceso abierto a las publicaciones científicas derivadas de las investigaciones financiadas con cargo al Programa;

b) acceso abierto a los datos de investigación, incluidas las publicaciones científicas subyacentes, de conformidad con el principio «**tan abierto como sea posible y tan cerrado como sea necesario**».

2. El principio de reciprocidad en la ciencia abierta se fomentará y alentará en todos los acuerdos de asociación y cooperación con terceros países, incluidos los firmados por los organismos financiadores a los que se haya confiado la gestión indirecta del Programa.

3. Se garantizará la gestión responsable de los datos de investigación en consonancia con los principios según los **cuales los datos deben ser fáciles de encontrar, accesibles, interoperables y reutilizables (en lo sucesivo, «principios FAIR»). Se prestará también atención a la conservación a largo plazo de los datos**.

Horizon Europe will also introduce several new main features:

- A European Innovation Council (EIC) to help the EU become a frontrunner in market-creating innovation.
- New EU-wide research and innovation missions focusing on societal challenges and industrial competitiveness.
- Maximising the innovation potential across the EU.
- **The principle of 'open science' will become the modus operandi of Horizon Europe, requiring open access to publications and data**.
- A new generation of European Partnerships and increased collaboration with other EU programmes.

Main elements of Open Science in Horizon Europe

Horizon Europe Regulation

Ciencia abierta a lo largo del programa

Ciencia abierta

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

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

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

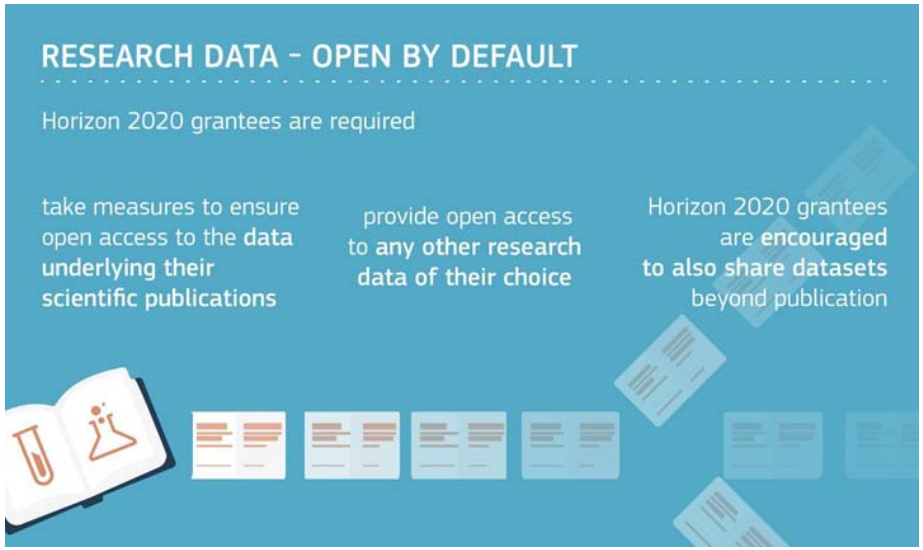
- Apoyo a las habilidades de los investigadores en materia de ciencia abierta, así como sistemas de recompensa.
- Uso de la Nube Europea de la Ciencia Abierta.

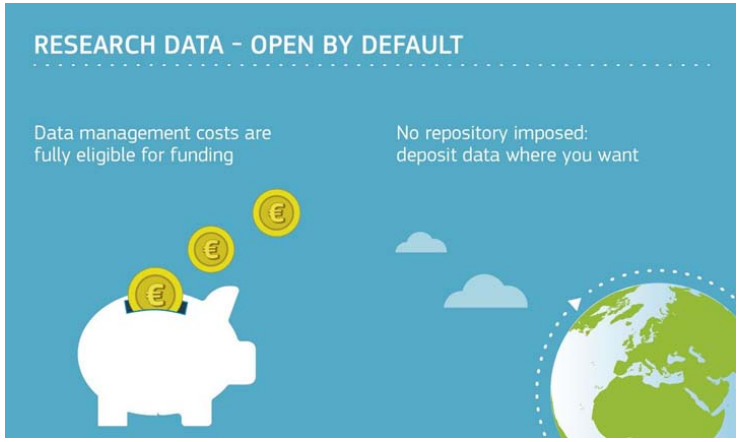
Comisión Europea

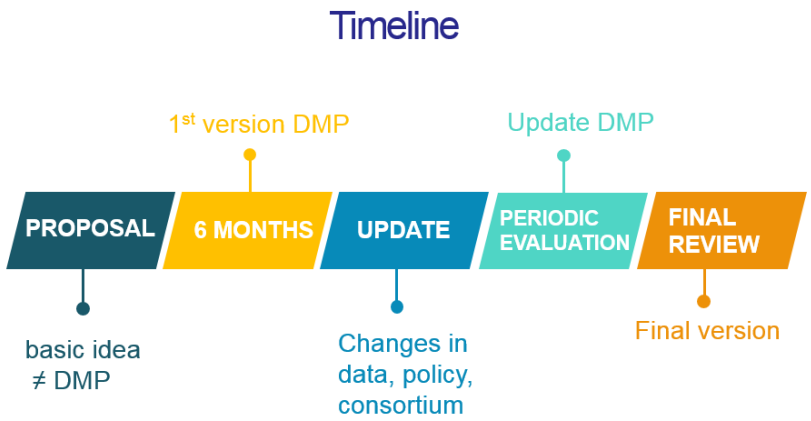
https://ec.europa.eu/info/sites/default/files/research_and_innovation/strategy_on_research_and_innovation/presentations/horizon_europe_es_invertir_para_dar_forma_a_nuestro_futuro.pdf

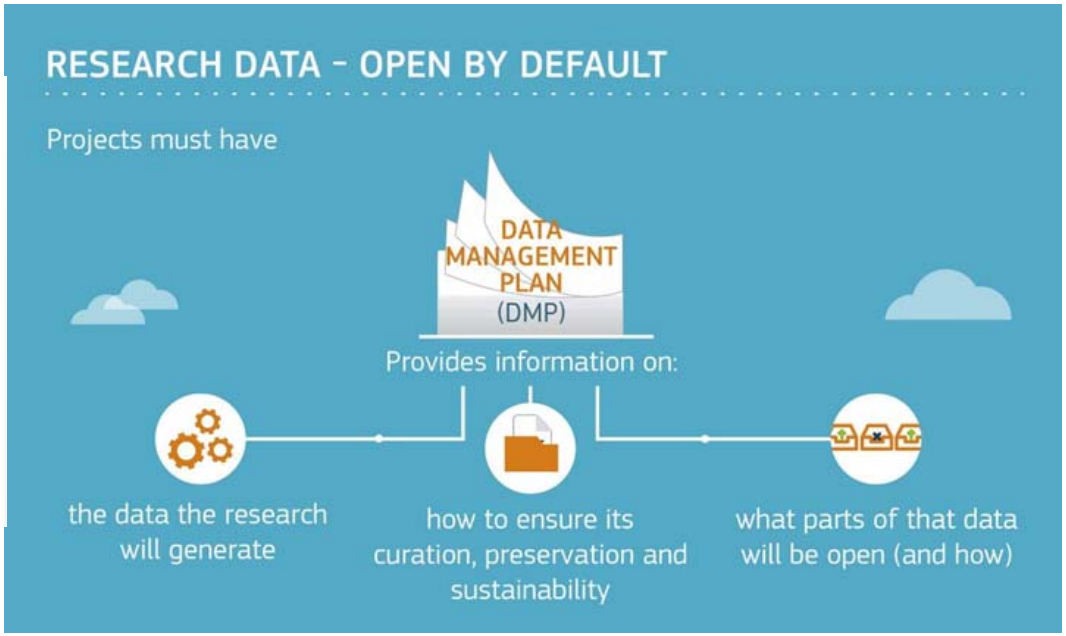
… Research Data Management (RDM)

- Establish and regularly update a **Data Management Plan**
- **Deposit data in a trusted repository** and provide **open access** through it
  - Deposit and open access **ASAP and per DMP**
  - For some actions, additional **obligation** to deposit in a repository that is **federated under EOSC**
- **CC BY** or **CC 0** (or equivalent) license required to open data
- **Exceptions to open access** (duly justified in the DMP; legitimate interests or constraints);
- **Information** via the repository about any other research output or any other tools and instruments needed to **re-use or validate the data;**
- **Metadata requirements** same as for publications (i.e. CC0 and PIDs)
- **Costs for RDM** (for example data storage, processing and preservation) are **eligible**

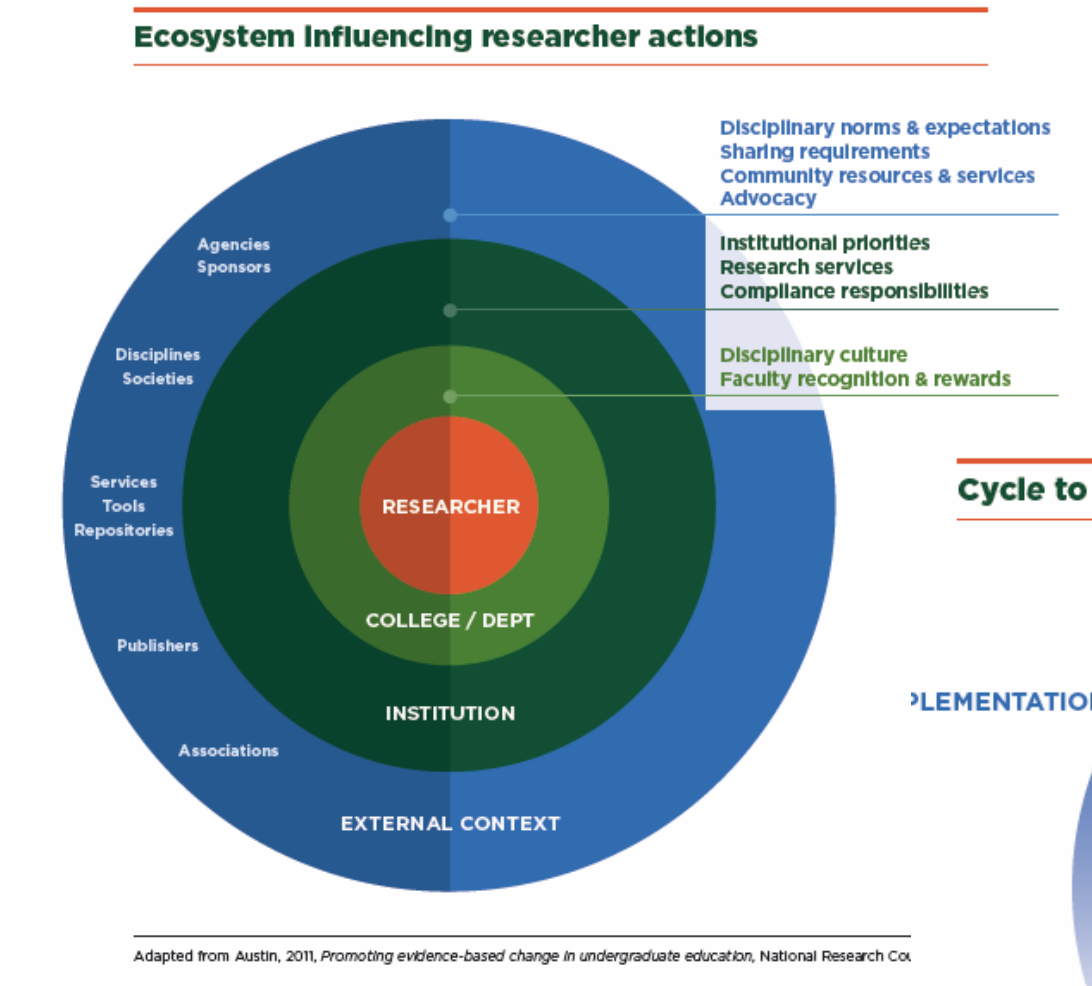






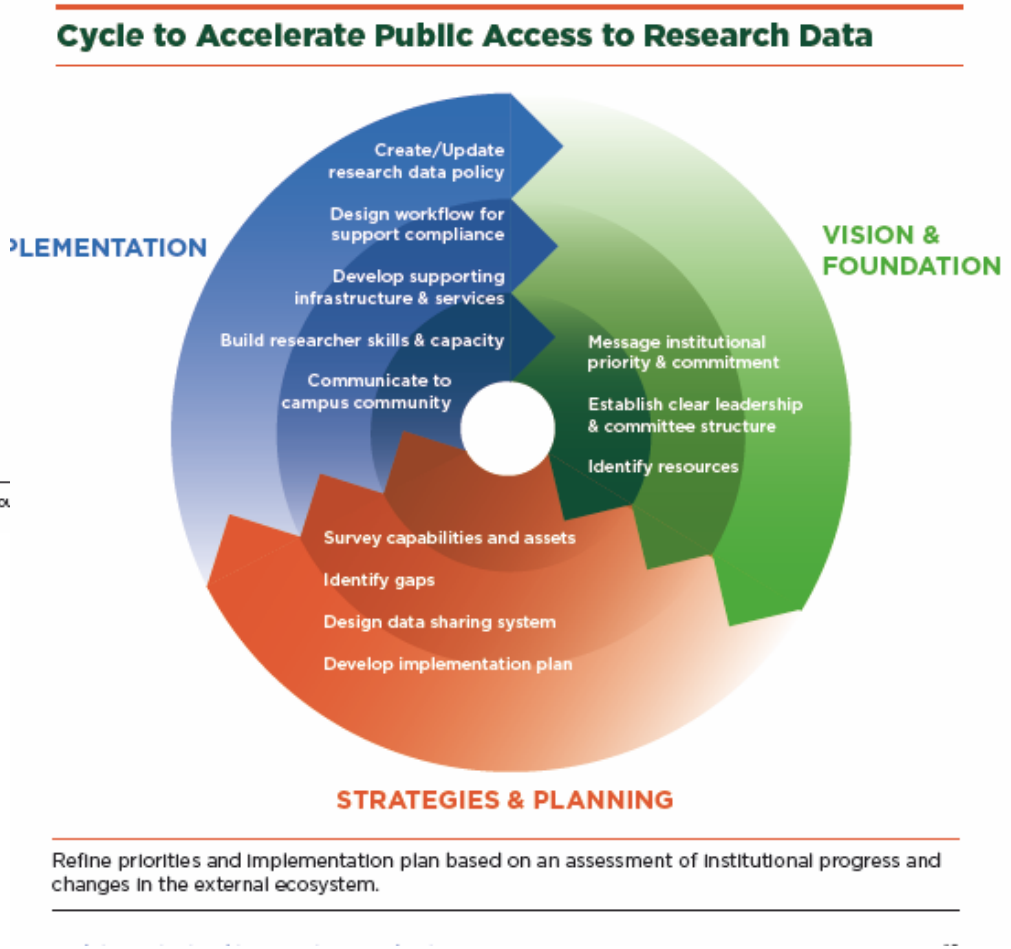


https://www.openaire.eu/how-to-comply-to-h2020-mandates-for-data



Guía de la Association of American Universities (AAU) and the Association of Public and Land-grant Universities https://www.aau.edu/accelerating-public-access-research-data

Factores y agentes que pueden influir en el acceso a los datos de investigación

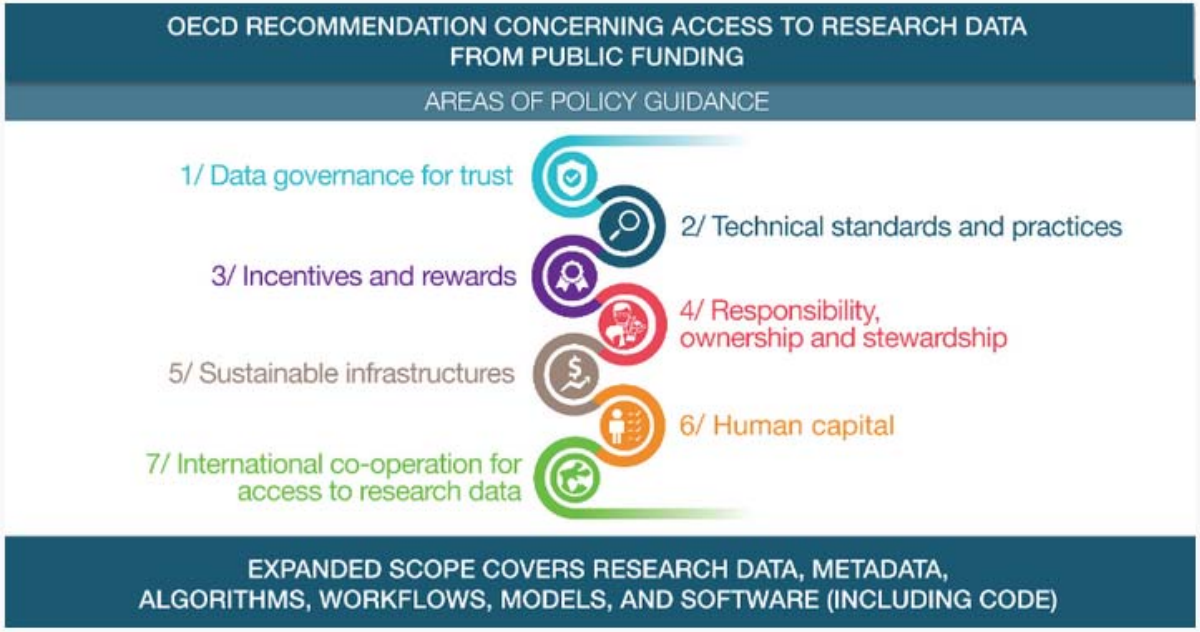


https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0347

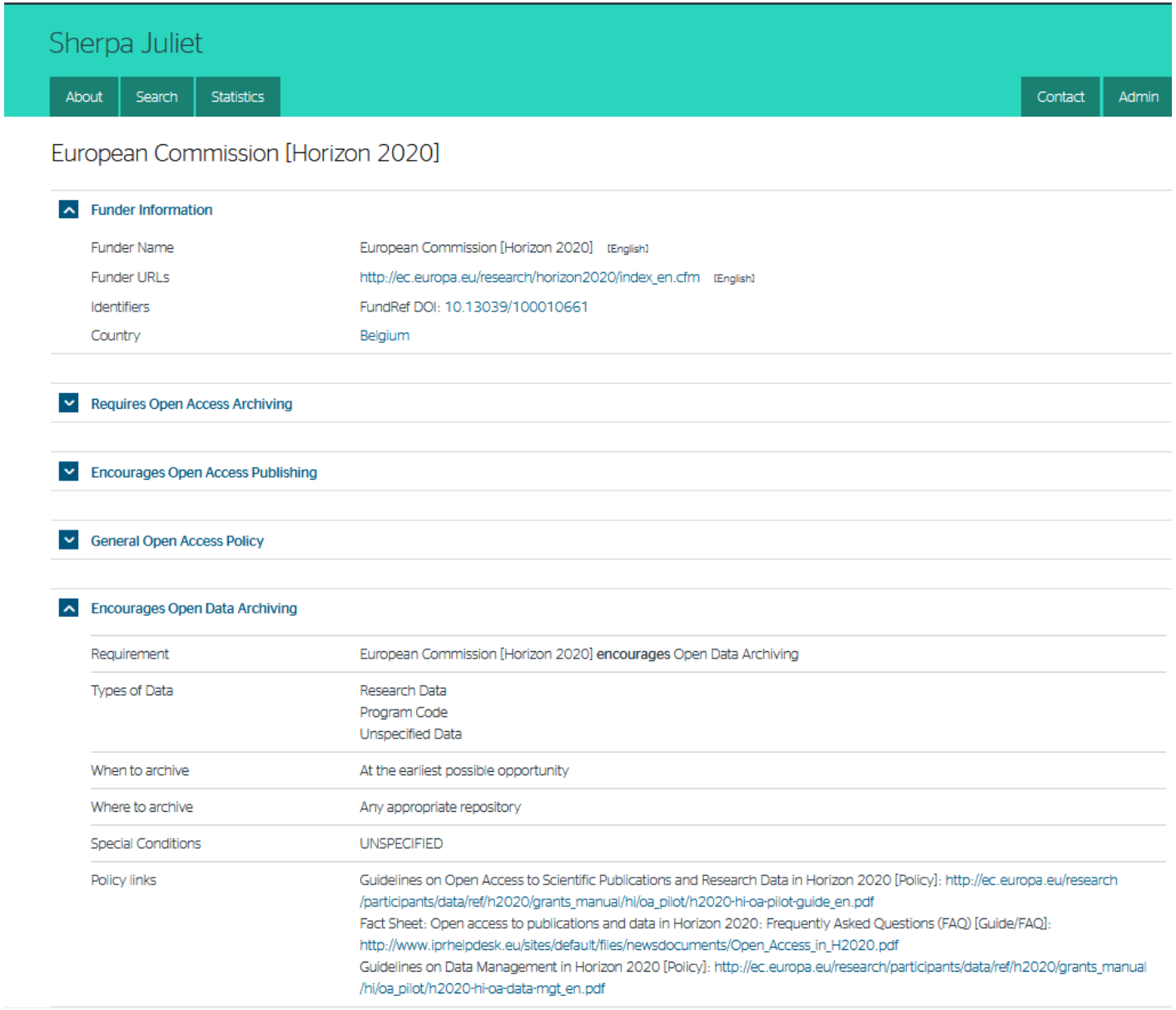
Recommendation of the OECD Council concerning Access to Research Data from Public Funding

On 20 January 2021, the OECD Council adopted a revised Council Recommendation on Access to Research Data from Public Funding. The legal instrument, in force since 2006, has been updated to address new technologies and policy developments, and provides policy guidance in seven areas shown in the figure below. In addition, the revision expands the scope to cover not only research data, but also related metadata, as well as bespoke algorithms, workflows, models, and software (including code), which are essential for their interpretation.

- Find out more in our blog: Making data for science as open as possible to address global challenges
- Policy brief: Why open science is critical to combatting COVID-19
- Analytical report: Enhanced Access to Publicly Funded Data for Science, Technology and Innovation
- Open science initiatives relating to COVID-19
- OECD work on open science



Para otras políticas: Sherpa Juliet. https://v2.sherpa.ac.uk/



Tipos de datos

Datos de investigación (*research data*)

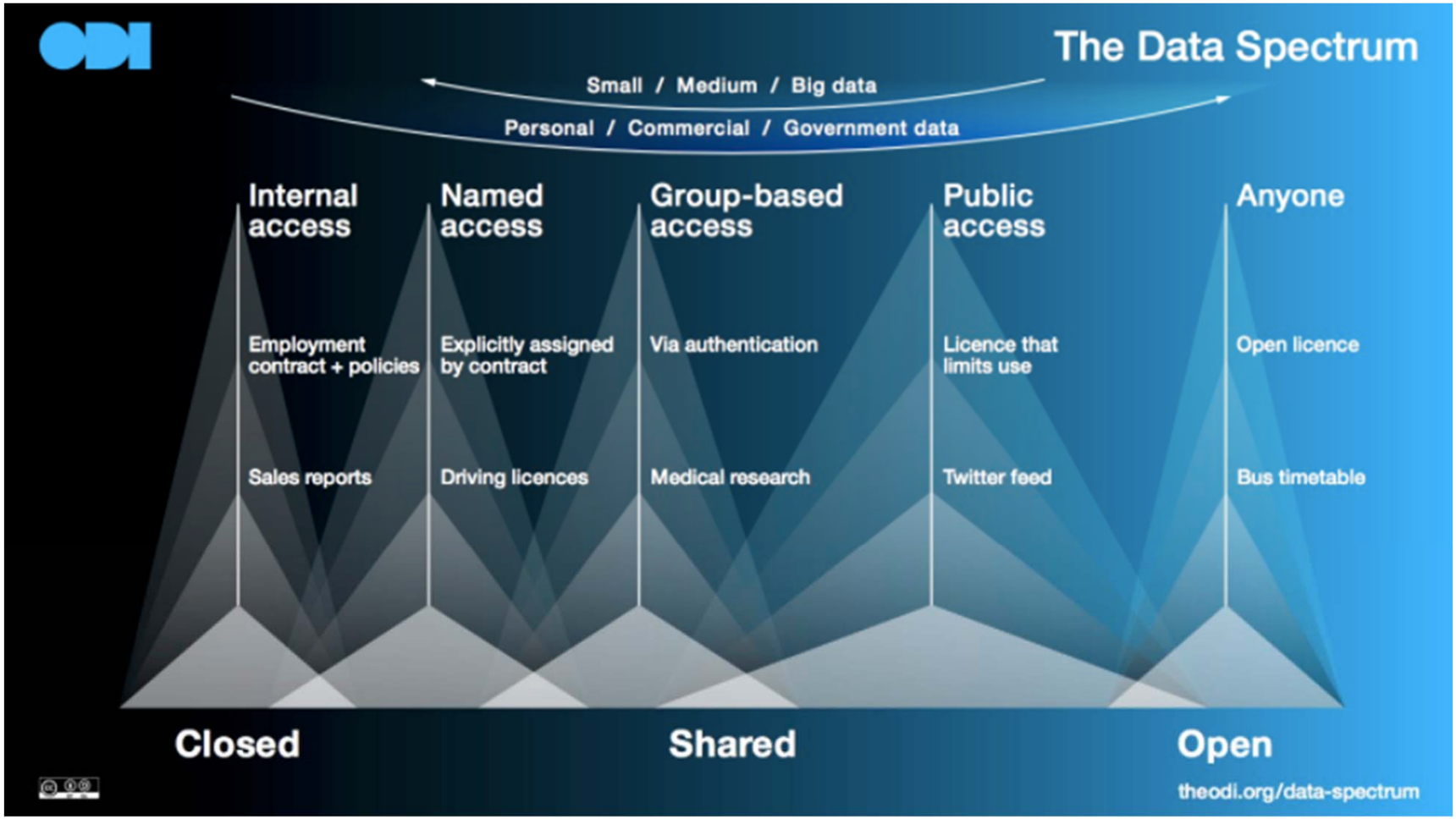
Información cuantitativa o cualitativa recogida por los investigadores en el curso de su trabajo obtenida de:

- La experimentación,
- La observación,
- La modelización,
- Por medio de encuestas o entrevistas, u otros medios
- ……

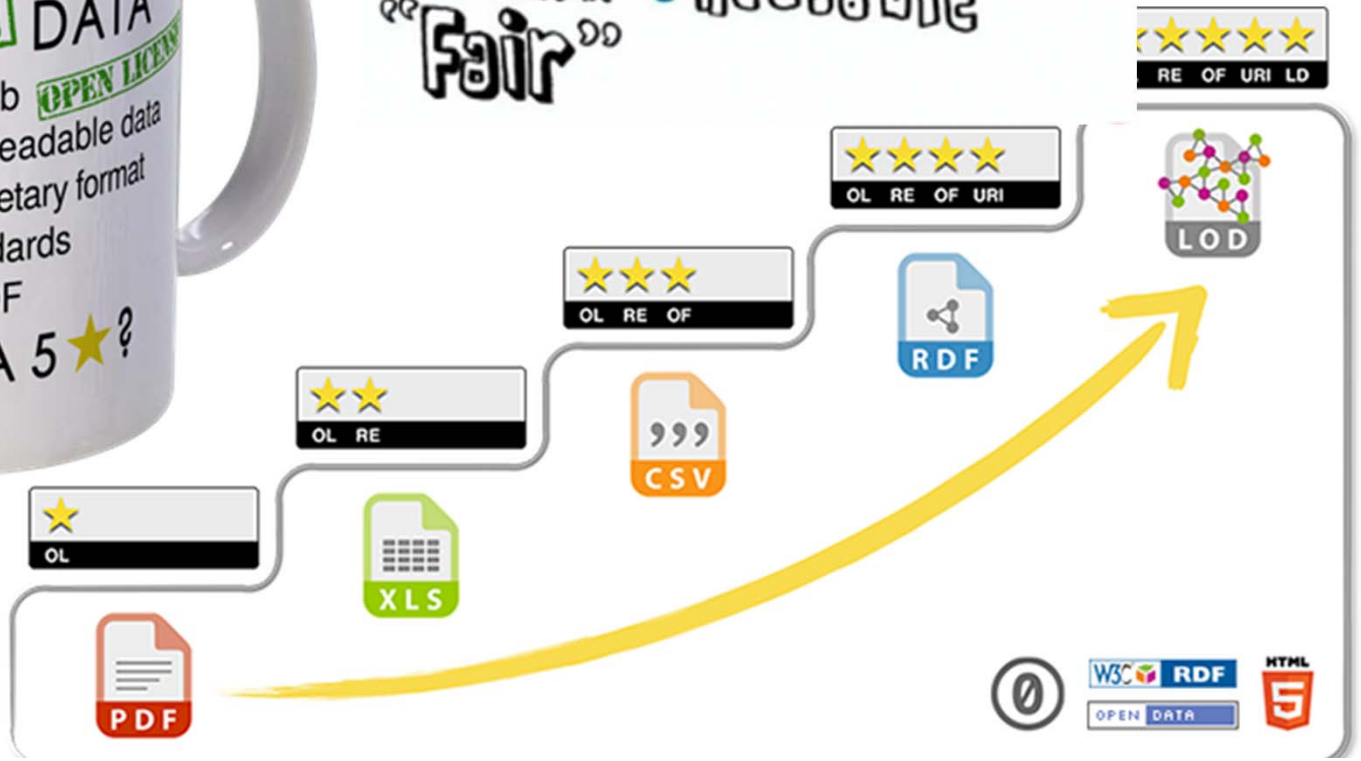
- También derivada de la ya existente

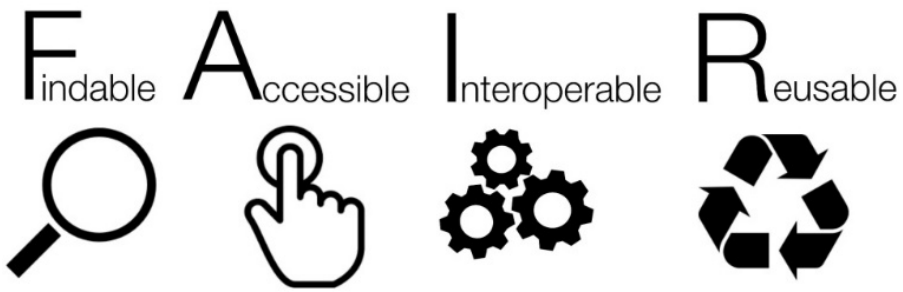
Facilitan la información necesaria para apoyar o validar los resultados o conclusiones de la investigación

https://theodi.org/about-the-odi/the-data-spectrum/



Open data must be accessible, useable, assessable and intelligible (definición de *Science as an Open Enterprise, 2012)*

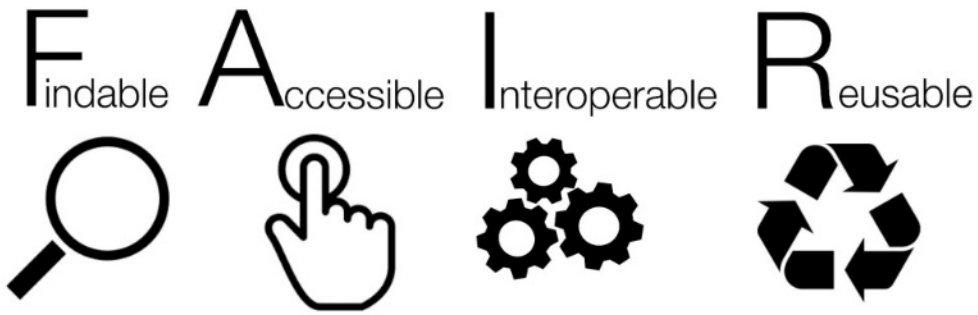




Principios FAIR (https://www.force11.org/group/fairgroup/fairprinciples)

Para ser localizable (*Findable*)
F1. A los (meta)datos se les debe asignar un identificador único y persistente.
F2. Los datos se describen con metadatos enriquecidos.
F3. (meta)datos se registran o indexan en un sistema que disponga de motor de búsqueda.
F4. Los metadatos contienen uno para el identificador del dataset.

Para ser accesible (*Accesible*):
A1. Los (meta)datos se pueden recuperar mediante su identificador utilizando un protocolo de comunicación estandarizado.
A1.1 El protocolo es abierto, gratuito y universalmente aplicable.
A1.2 El protocolo permite un procedimiento de autenticación y autorización, cuando sea necesario.
A2. Los metadatos son accesibles, incluso cuando los datos ya no están disponibles.



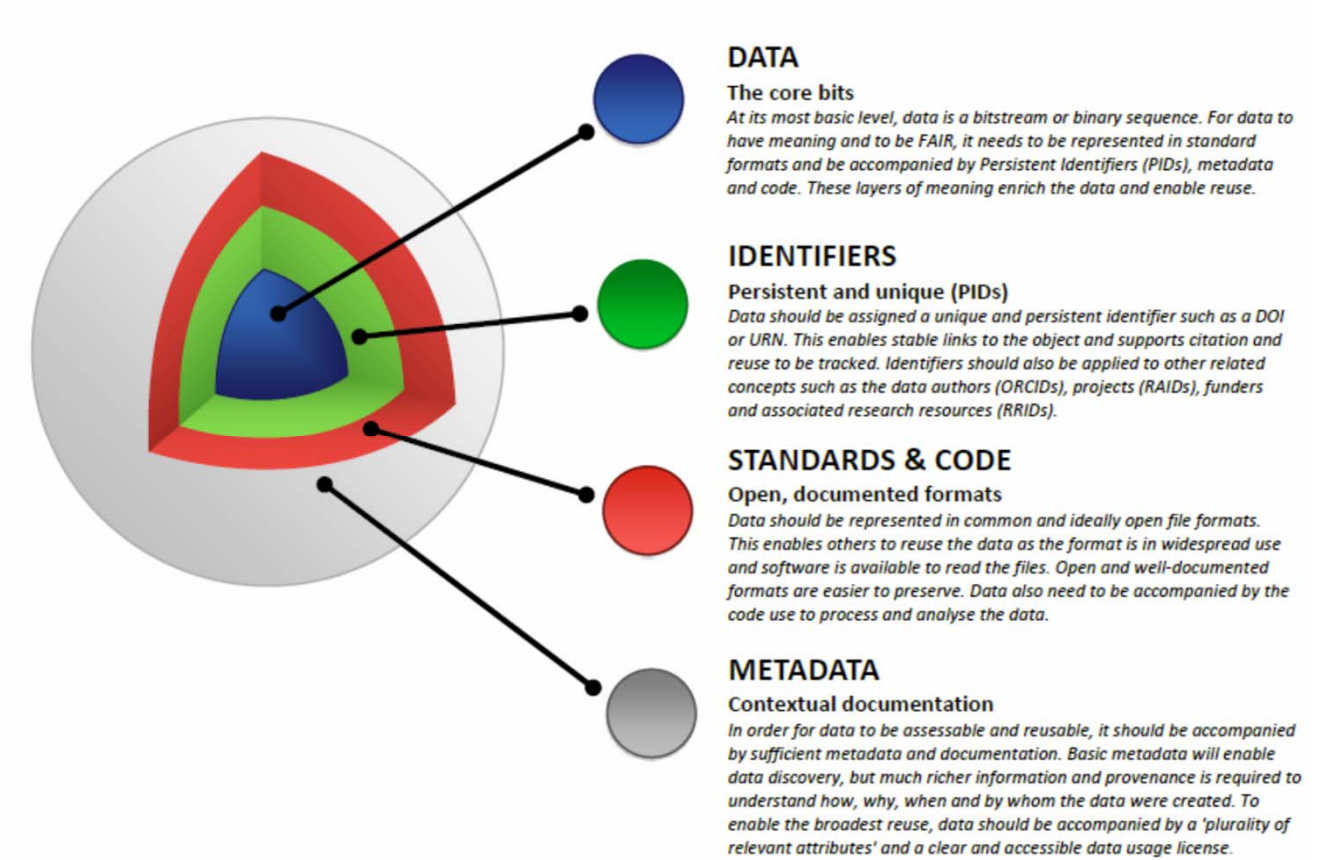
Principios FAIR (https://www.force11.org/group/fairgroup/fairprinciples)

Para ser interoperable (Interoperable):
I1. Los (meta)datos utilizan un lenguaje formal, accesible, compartido y ampliamente aplicable para la representación del conocimiento.
I2. Los (meta)datos usan vocabularios que siguen principios FAIR.
I3. Los (meta)datos incluyen referencias calificadas a otros (meta)datos.

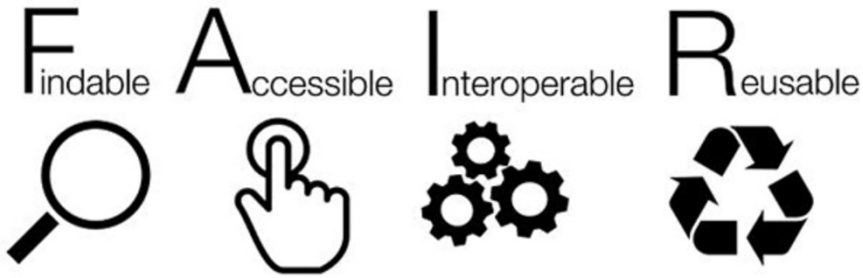
Para ser reutilizable (Reusable)

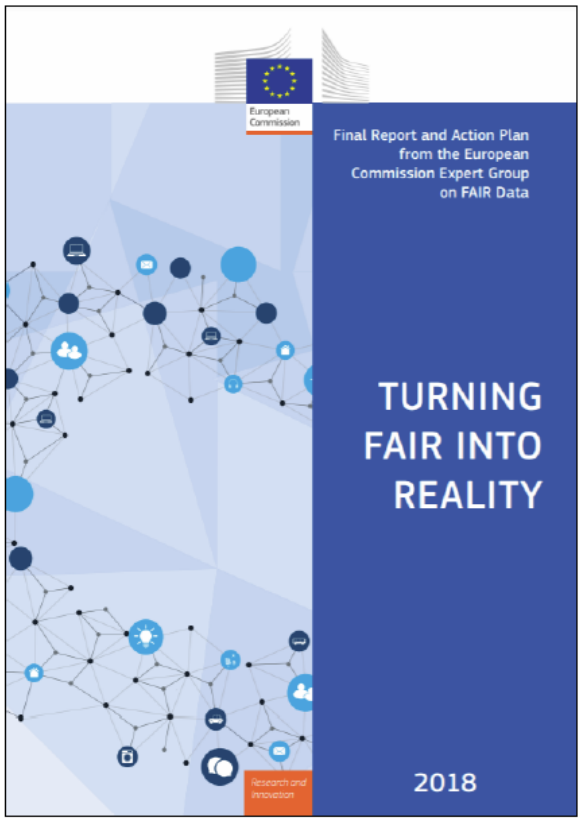
R1. Los (meta)datos tienen disponen de atributos precisos y pertinentes.
R1.1. Los (meta)datos se publican con una licencia de reutilización clara y accesible.
R1.2. Los (meta)datos están asociados con su procedencia.
R1.3. Los (meta)datos cumplen con los estándares propios de una comunidad de un área de conocimiento determinada.

Turning FAIR data into reality. Interim report from the European Commission Expert Group on FAIR data



Model for FAIR data objects



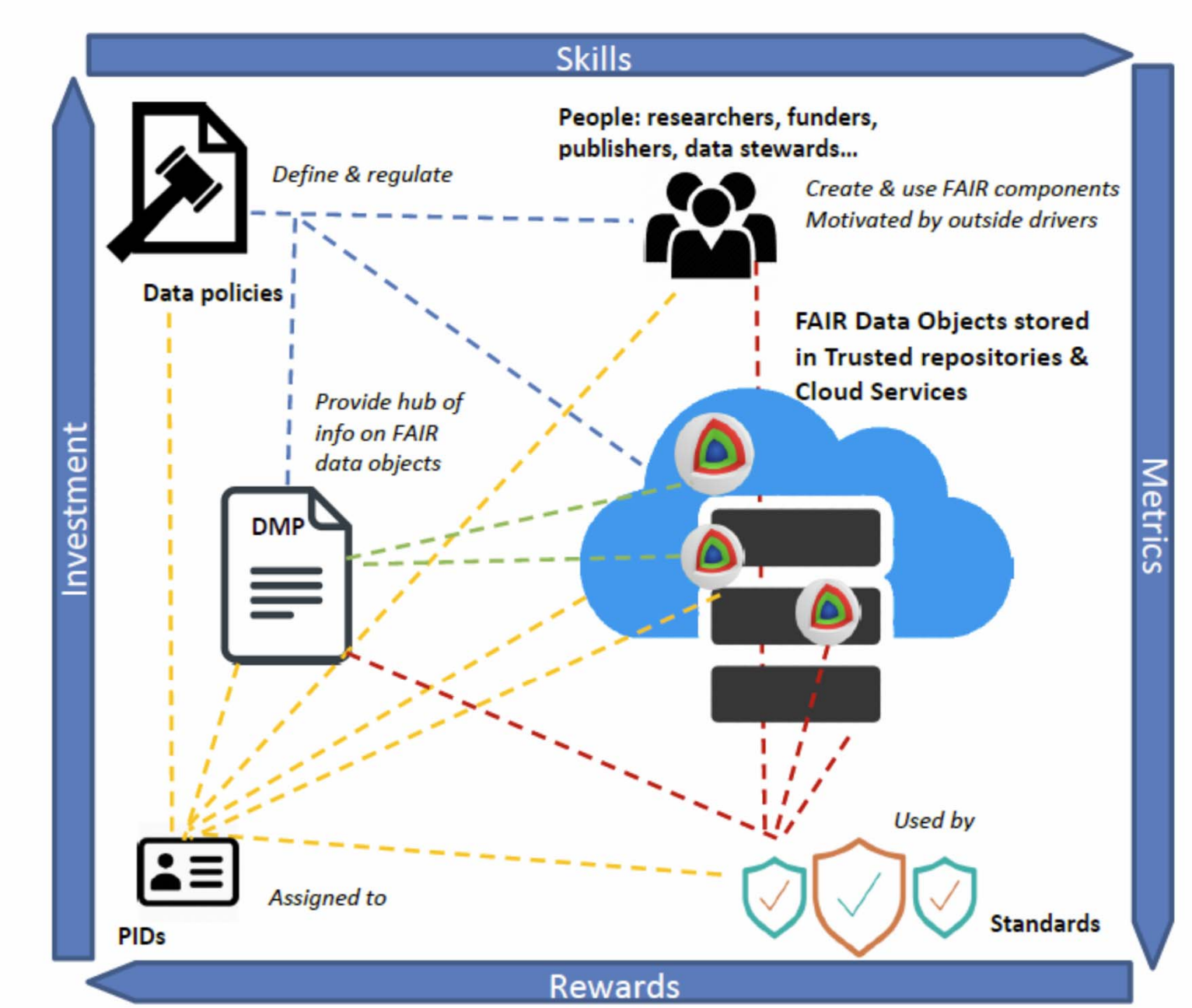


https://ec.europa.eu/info/sites/default/files/turning_fair_into_reality_1.pdf

Open ≠ FAIR

- FAIR and Open should not be conflated. Data can be FAIR or Open, both or neither
- Greatest potential reuse comes when data are both
- Even internal or restricted data will benefit from being FAIR, and there are legitimate reasons for restriction which vary by discipline

CONCEPTS

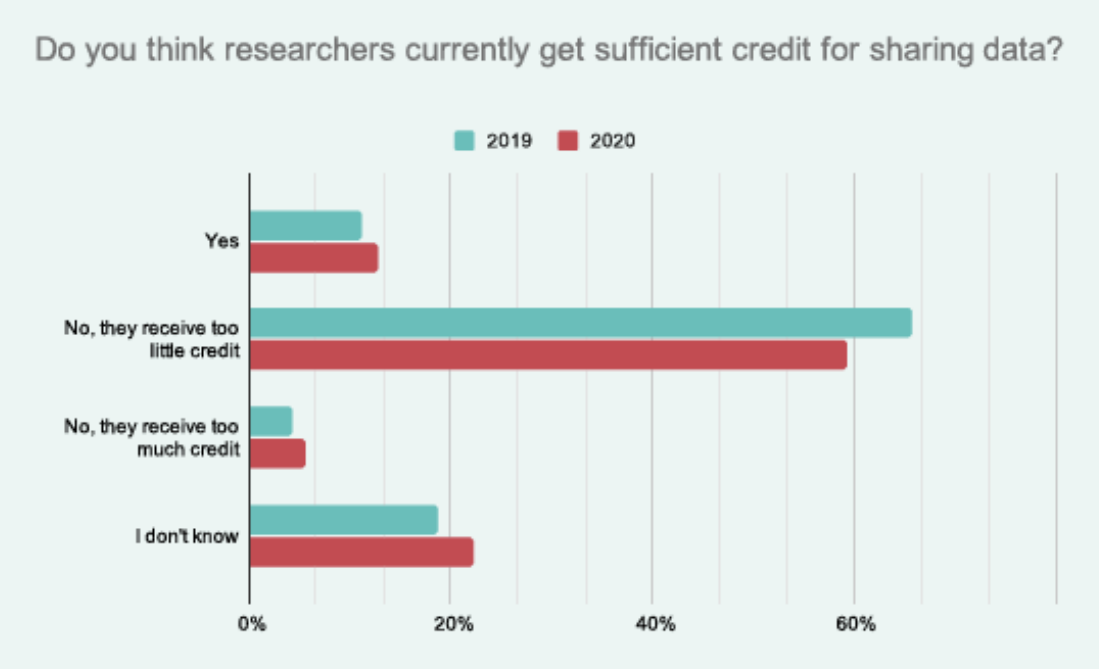


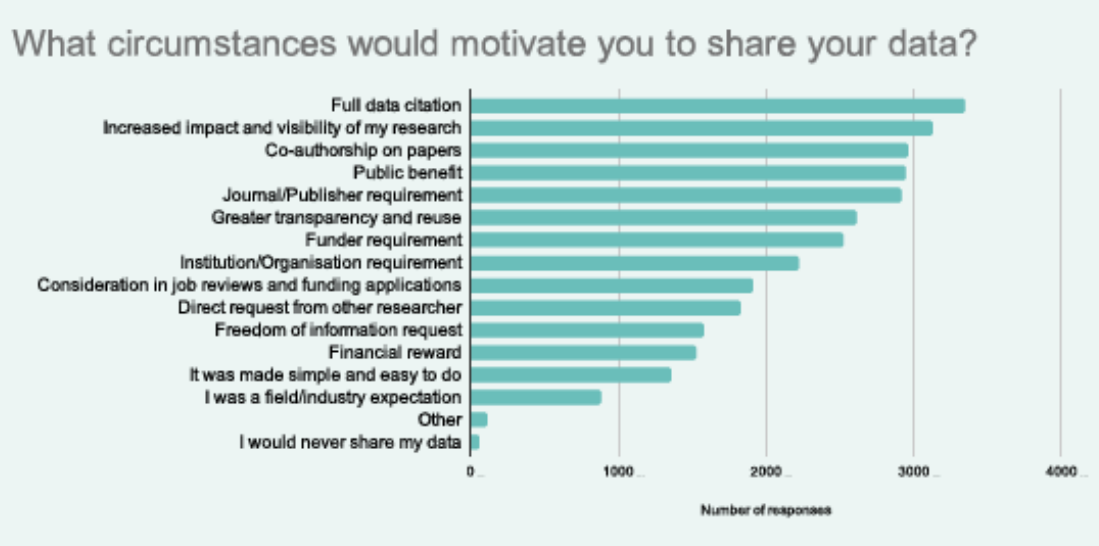
Interacciones entre los componentes del ecosistema de datos FAIR

Turning FAIR data into reality. Interim report from the European Commission Expert Group on FAIR data

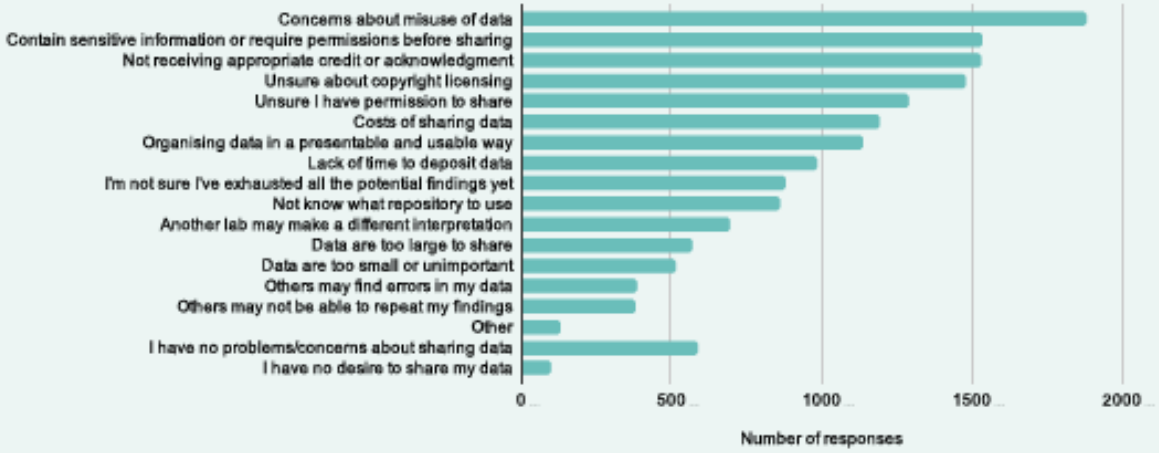
Sectores		Acciones
Research communities		Step 1: Define and apply FAIR appropriate
Data services		Step 2: Develop and support a sustainable FAIR data ecosystem
Data stewards		Step 3: Ensure FAIR data and certified services to support FAIR
Standards bodies		Step 4: Embed a culture of FAIR in research practice
Global coordination fora		Fair data policy
Policymakers		FAIR data culture
Research funders		Technology for FAIR
Institutions		Skills and roles for FAIR
Publishers		FAIR metrics
		Costs and investment in FAIR
		How fair DATA Action Plan supports EOSC

The State of Open Data 2020 (n= 4500 respuestas)
https://www.digital-science.com/resources/portfolio-reports/the-state-of-open-data-2020/

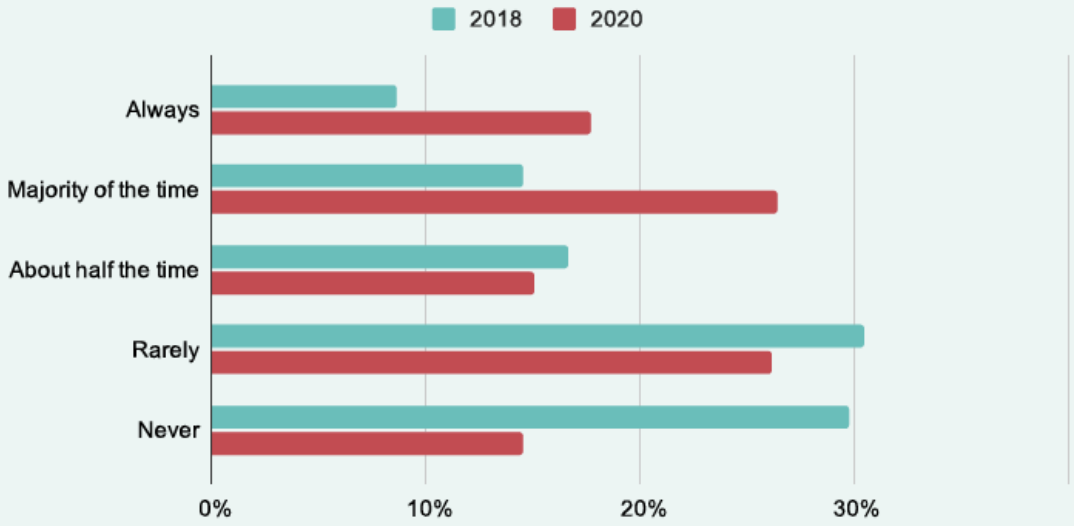




What problems/concerns, if any, do you have with sharing datasets?



How often do you create a data management plan for the research you carry out? by Year



Formatos de datos en función de su contenido

Las **convenciones de nomenclatura de archivos y carpetas** son claves para mantener organizados los datos de investigación.

Recomendaciones para nombrar archivos de datos:

- Definir nombres cortos y relevantes: 25 caracteres es una longitud suficiente para capturar información descriptiva de un archivo de datos.
- No utilizar caracteres especiales en un nombre de archivo como: & *% $ £] {! @ ya que se utilizan para tareas específicas en sistemas operativos.
- Usar guiones bajos en lugar de puntos o espacios completos porque, como los caracteres especiales, estos se analizan de manera diferente en diferentes sistemas.
- El nombre del archivo debe incluir información descriptiva que ayude a la identificación independientemente de dónde se almacena.
- Ser consistente en el uso de minúsculas/mayúsculas al nombrar archivos ya que algunos programas pueden considerar diferentes archivos con el mismo nombre.
- Cuando sea posible, usar extensiones de archivo (a menudo predeterminadas) para reflejar con precisión el entorno de software en el que se creó el archivo y el formato físico del archivo. Por ejemplo, use “por” para archivos portátiles de SPSS, .xls o .xlsx para archivos de Excel, “.txt” para archivos de texto.
- Nombrar archivos por cronología, usando el formato Año-Mes-Día: AAAA-MM-DD o AAAA-MM. Esto mantendrá el orden cronológico de los archivos.

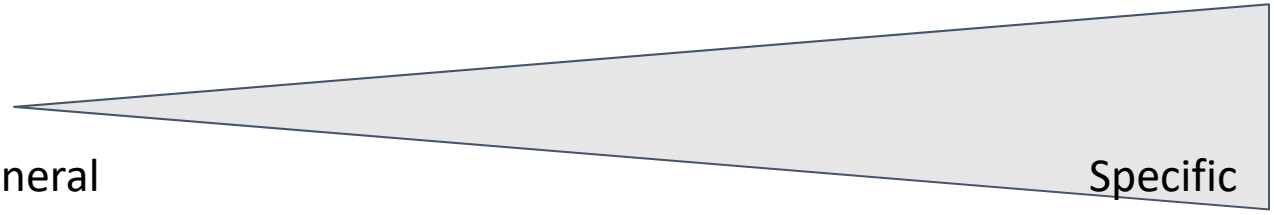
File naming convention (FNC)

- Include data in yyyy-mm-dd format
- Use meaningful abbreviations
- Have group identifiers
- Document your decisions
- Be consistent

20130825_DOEProject_Ex1Test1_Data_Gonzalez_v3-03.xlsx

Date **Project** **Type** **Version**

Experiment **ID**



General Specific

http://guides.lib.purdue.edu/c.php?g=353013&p=2378292

Recomendaciones respecto a los formatos de los ficheros

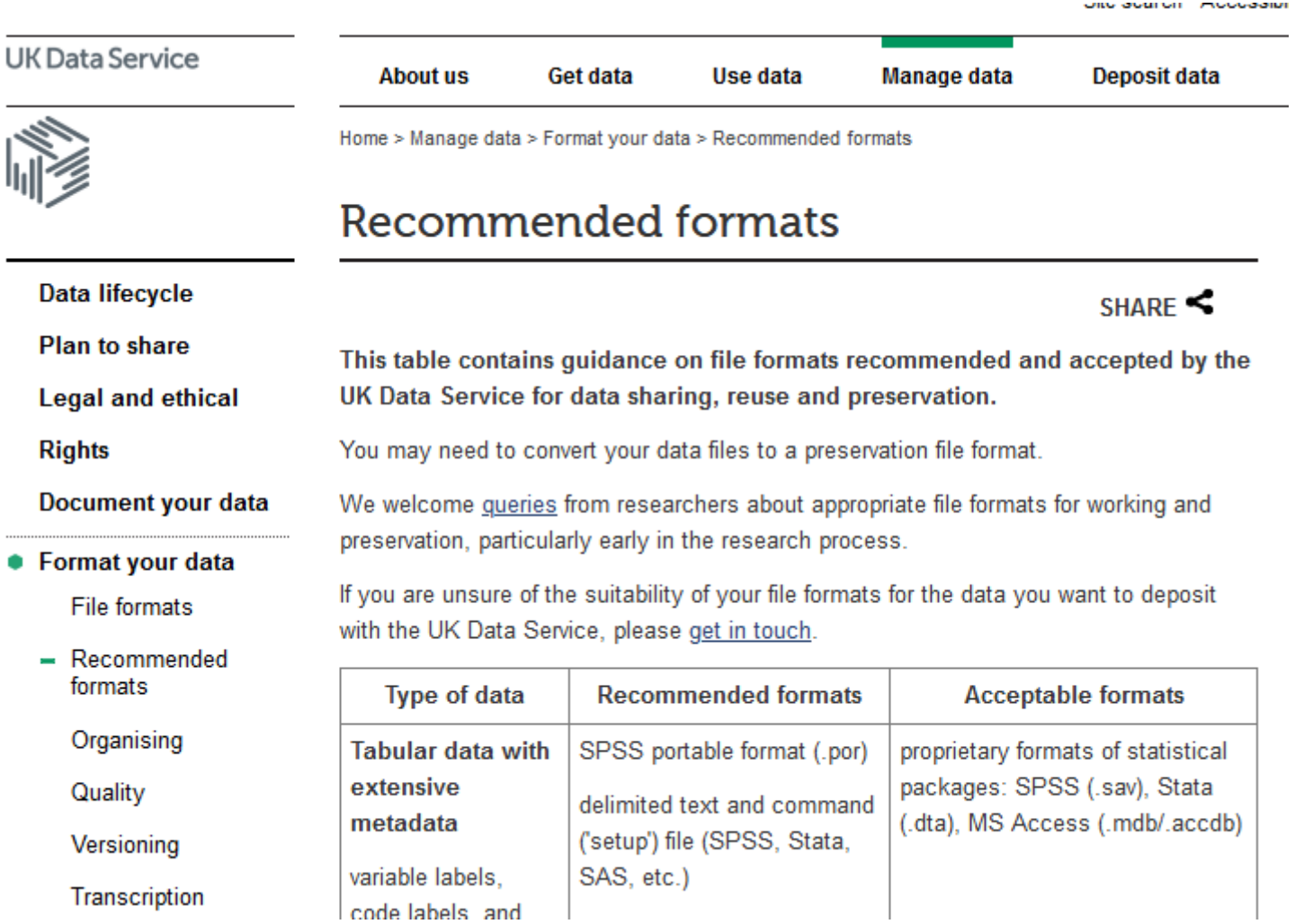
- Publicar los datos en formatos abiertos, no propietarios y legibles por máquina, si es posible.
- Guarde los archivos en formatos abiertos y propietarios (por ejemplo, data.xlsx y data.csv) para mejorar la usabilidad, si procede.
- Comprima los datos para facilitar el intercambio y la descarga de archivos grandes. Compruebe los formatos de archivos comprimidos (zip) en la investigación.
- Mantenga el mismo nombre de archivo para el mismo archivo en diferentes formatos (por ejemplo, data.doc y data.txt).
- Incluir la documentación de los datos, las transformaciones y las directrices del software para acceder al formato propietario.
- Considere la posibilidad de compartir el código de software alternativo a los ejecutables propietarios.
- Compruebe si hay errores u omisiones si convierte el archivo a un formato diferente.

https://www.ukdataservice.ac.uk/manage-data/format/recommended-formats.aspx

Type of data	Recommended formats	Acceptable formats
Image data	TIFF 6.0 uncompressed (.tif)	JPEG (.jpeg, .jpg, .jp2) if original created in this format GIF (.gif) TIFF other versions (.tif, .tiff) RAW image format (.raw) Photoshop files (.psd) BMP (.bmp) PNG (.png) Adobe Portable Document Format (PDF/A, PDF) (.pdf)
Audio data	Free Lossless Audio Codec (FLAC) (.flac)	MPEG-1 Audio Layer 3 (.mp3) if original created in this format Audio Interchange File Format (.aif) Waveform Audio Format (.wav)
Video data	MPEG-4 (.mp4) OGG video (.ogv, .ogg) motion JPEG 2000 (.mj2)	AVCHD video (.avchd)
Documentation and scripts	Rich Text Format (.rtf) PDF/UA, PDF/A or PDF (.pdf) XHTML or HTML (.xhtml, .htm) OpenDocument Text (.odt)	plain text (.txt) widely-used formats: MS Word (.doc/.docx), MS Excel (.xls/.xlsx) XML marked-up text (.xml) according to an appropriate DTD or schema, e.g. XHMTL 1.0

Type of data	Recommended formats	Acceptable formats
Tabular data with extensive metadata variable labels, code labels, and defined missing values	SPSS portable format (.por) delimited text and command ('setup') file (SPSS, Stata, SAS, etc.) structured text or mark-up file of metadata information, e.g. DDI XML file	proprietary formats of statistical packages: SPSS (.sav), Stata (.dta), MS Access (.mdb/.accdb)
Tabular data with minimal metadata column headings, variable names	comma-separated values (.csv) tab-delimited file (.tab) delimited text with SQL data definition statements	delimited text (.txt) with characters not present in data used as delimiters widely-used formats: MS Excel (.xls/.xlsx), MS Access (.mdb/.accdb), dBase (.dbf), OpenDocument Spreadsheet (.ods)
Geospatial data vector and raster data	ESRI Shapefile (.shp, .shx, .dbf, .prj, .sbx, .sbn optional) geo-referenced TIFF (.tif, .tfw) CAD data (.dwg) tabular GIS attribute data Geography Markup Language (.gml)	ESRI Geodatabase format (.mdb) MapInfo Interchange Format (.mif) for vector data Keyhole Mark-up Language (.kml) Adobe Illustrator (.ai), CAD data (.dxf or .svg) binary formats of GIS and CAD packages
Textual data	Rich Text Format (.rtf) plain text, ASCII (.txt) eXtensible Mark-up Language (.xml) text according to an appropriate Document Type Definition (DTD) or schema	Hypertext Mark-up Language (.html) widely-used formats: MS Word (.doc/.docx) some software-specific formats: NUD*IST, NVivo and ATLAS.ti

https://www.ukdataservice.ac.uk/manage-data/format/recommended-formats.aspx



UK Data Service

About us | Get data | Use data | Manage data | Deposit data

Home > Manage data > Format your data > Recommended formats

- Data lifecycle
- Plan to share
- Legal and ethical
- Rights
- Document your data
- Format your data
  - File formats
  - Recommended formats
  - Organising
  - Quality
  - Versioning
  - Transcription

Recommended formats

SHARE

This table contains guidance on file formats recommended and accepted by the UK Data Service for data sharing, reuse and preservation.

You may need to convert your data files to a preservation file format.

We welcome queries from researchers about appropriate file formats for working and preservation, particularly early in the research process.

If you are unsure of the suitability of your file formats for the data you want to deposit with the UK Data Service, please get in touch.

Type of data	Recommended formats	Acceptable formats
Tabular data with extensive metadata variable labels, code labels, and	SPSS portable format (.por) delimited text and command ('setup') file (SPSS, Stata, SAS, etc.)	proprietary formats of statistical packages: SPSS (.sav), Stata (.dta), MS Access (.mdb/.accdb)

Cómo licenciar y citar datasets

Data licensing allows knowing how to reuse your data

Open Data Licensing Animation - OERIPR Support
https://www.youtube.com/watch?v=Tvwp5LK_Wko

① Data

CC0

Creative Commons Zero is ideal for openly sharing data – it has no restrictions on reuse whatsoever. While CC0 contains no requirement for attribution, citing CC0 datasets is widely accepted and expected in research.

Other CC licenses

If a CC0 license is not suitable for your data, consider the following CC licenses - all of which require attribution and further restrictions:

CC-BY	Prevents others from applying legal restrictions beyond the terms of the license to the licensed dataset.
CC BY-SA	Requires outputs derived from licensed dataset to also be licensed as CC BY-SA.
CC BY-NC	Prevents the licensed data from being used for commercial purpose.
CC BY-ND	Prevents the licensed data from being modified.
CC BY-NC-ND	Prevents the licensed data from being used for commercial purposes or modified.
CC BY-NC-SA	Prevents the licensed data from being used for commercial purposes, and requires outputs derived from licensed dataset to also be licensed as CC BY-SA

Tres “Capas” de las licencias

Caution!

NC, ND and SA licenses have implications for reuse and interoperability.We suggest using a license that keeps your data as open as possible and as closed as necessary.

Licencias recomendadas por el OKF conformes a su definición de abierto (atribución y compartir igual) . Open Data Commons y Creative Commons

License	Domain	By	SA	Comments
Creative Commons CCZero (CC0)	Content, Data	N	N	Dedicate to the Public Domain (all rights waived)
Open Data Commons Public Domain Dedication and Licence (PDDL)	Data	N	N	Dedicate to the Public Domain (all rights waived)
Creative Commons Attribution 4.0 (CC-BY-4.0)	Content, Data	Y	N	
Open Data Commons Attribution License (ODC-BY)	Data	Y	N	Attribution for data(bases)
Creative Commons Attribution Share-Alike 4.0 (CC-BY-SA-4.0)	Content, Data	Y	Y	
Open Data Commons Open Database License (ODbL)	Data	Y	Y	Attribution-ShareAlike for data(bases)

http://opendefinition.org/licenses/

Cómo asignar estas licencias

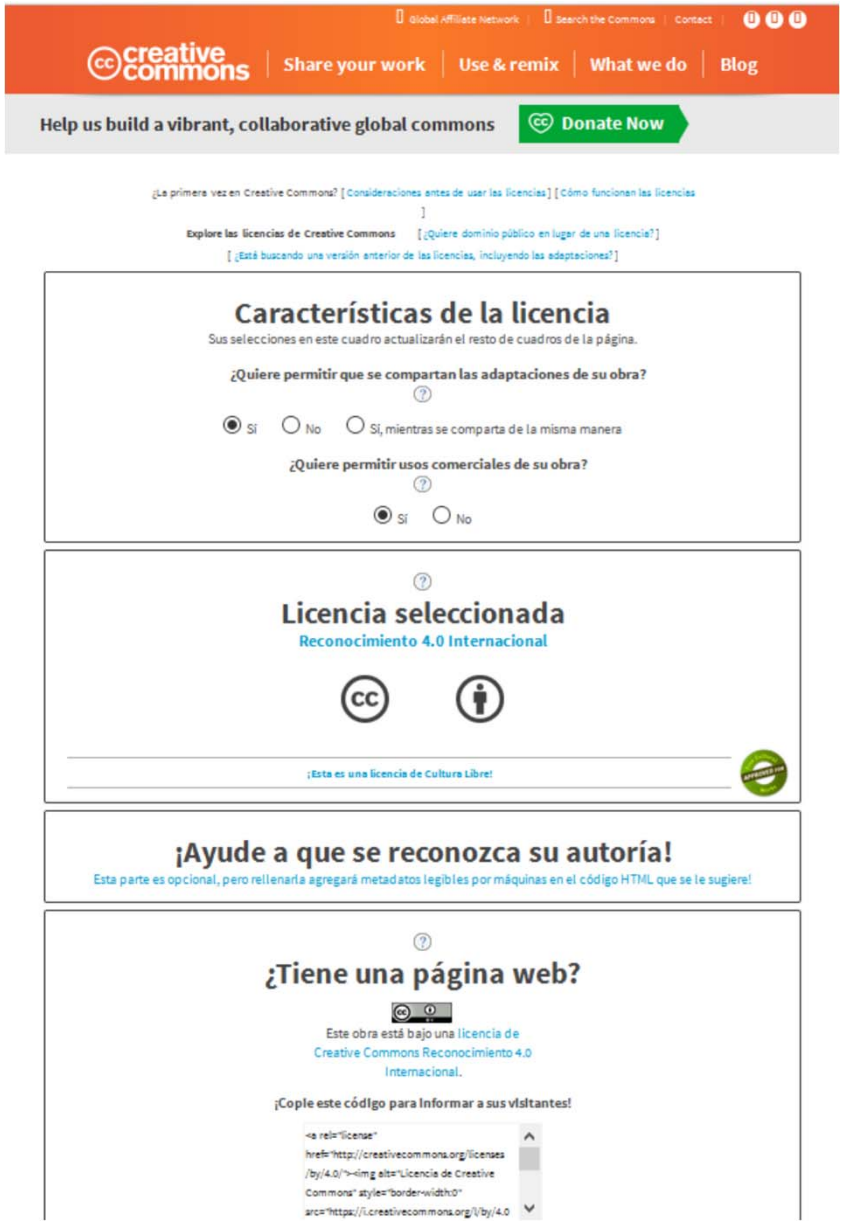
Public Domain Dedication and License (PDDL)

This {DATA(BASE)-NAME} is made available under the Public Domain Dedication and License version v1.0 whose full text can be found at http://opendatacommons.org/licenses/pddl/

NB: you may also wish to apply some complementary Community Norms
Open Database License

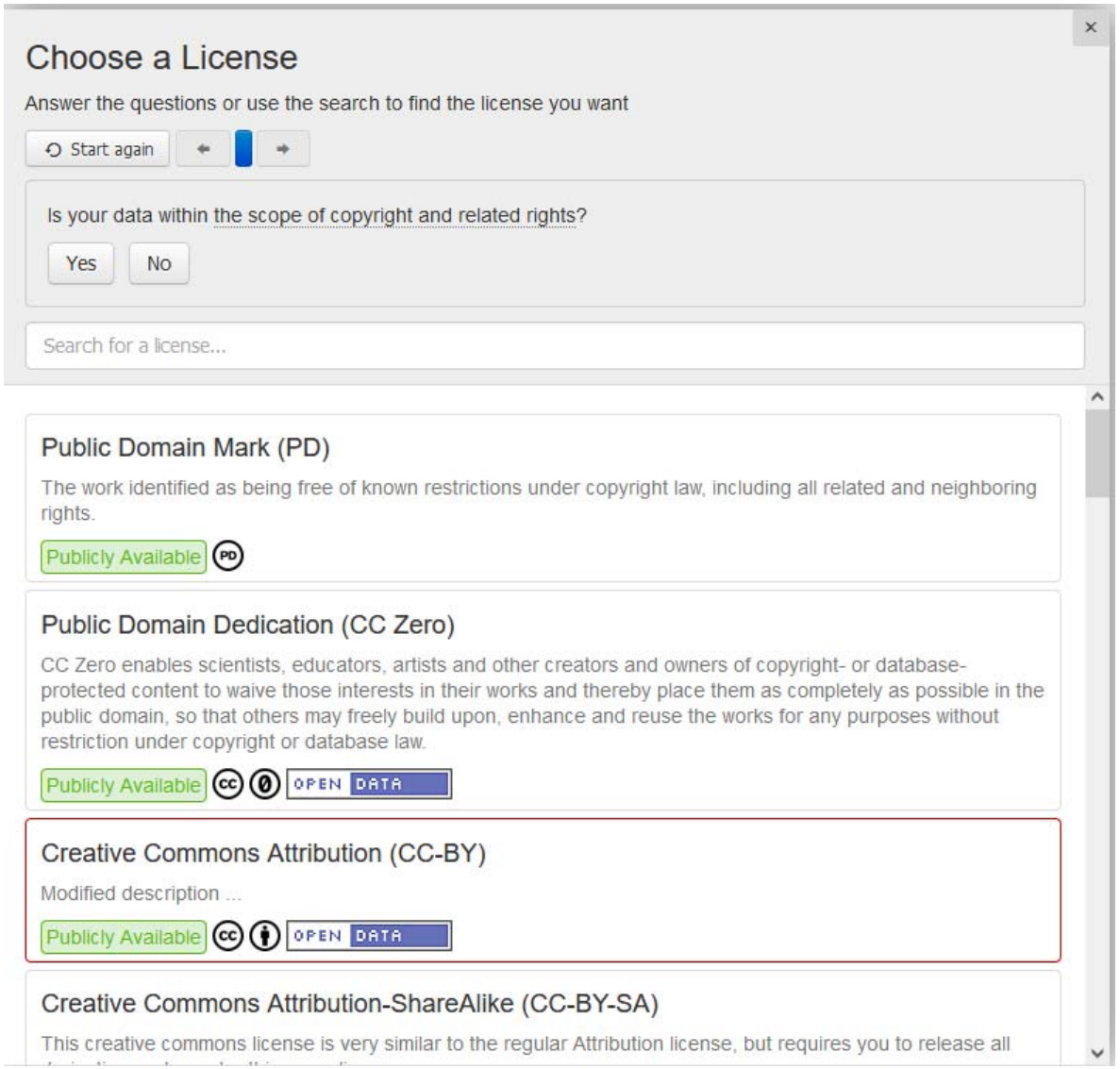
This {DATA(BASE)-NAME} is made available under Open Database License whose full text can be found at http://opendatacommons.org/licenses/odbl/. Any rights in individual contents of the database are licensed under the Database Contents License whose text can be found http://opendatacommons.org/licenses/dbcl/

https://opendatacommons.org/licenses/



https://creativecommons.org/choose/

https://ufal.github.io/public-license-selector/



Cómo citar

https://search.datacite.org/works/10.5281/zenodo.4478250

Por qué y cómo citar los datos (para autores)

¿Por qué?

1. Facilita la reproducibilidad y la validación de los resultados, permite la reutilización de los datos.
2. Proporciona crédito a los generadores de datos.
3. Las publicaciones vinculadas a datos disponibles públicamente se han asociado a un aumento de las citas.
4. Mejora la conectividad y el seguimiento de la procedencia de los datos descritos en las publicaciones.
5. Las agencias de financiación (ej Comisión Europea), los editores (Plos) y las instituciones (p.e. CSIC) exigen cada vez más que se compartan los datos.

¿Cómo?

1. En el caso de los datos primarios: determine un repositorio apropiado a largo plazo para archivar los datos. Su editor debe proporcionar acceso a una lista de repositorios de archivo aceptables.
2. Deposite sus datos y obtenga del repositorio un número de acceso o un DOI del conjunto de datos.
3. Para los datos secundarios: cite lo que utiliza. Cuando utilice los datos de otros, cite tanto la literatura revisada por pares como los conjuntos de datos utilizados
4. Incluya citas de datos "formales" siempre que sea posible: Cuando los conjuntos de datos tienen identificadores formales y estables o números de acceso, deben incluirse en la lista de referencias principal.
5. Sea lo más completo posible, pero no se invente los metadatos: Si un registro de datos no tiene un autor/creador o un título claros, no se invente uno.
6. Consulte la Guía para autores de su editor para dar formato a la referencia de su conjunto de datos.

Cousijn, H., Kenall, A., Ganley, E. *et al.* A data citation roadmap for scientific publishers. *Sci Data* **5,** 180259 (2018). https://doi.org/10.1038/sdata.2018.259

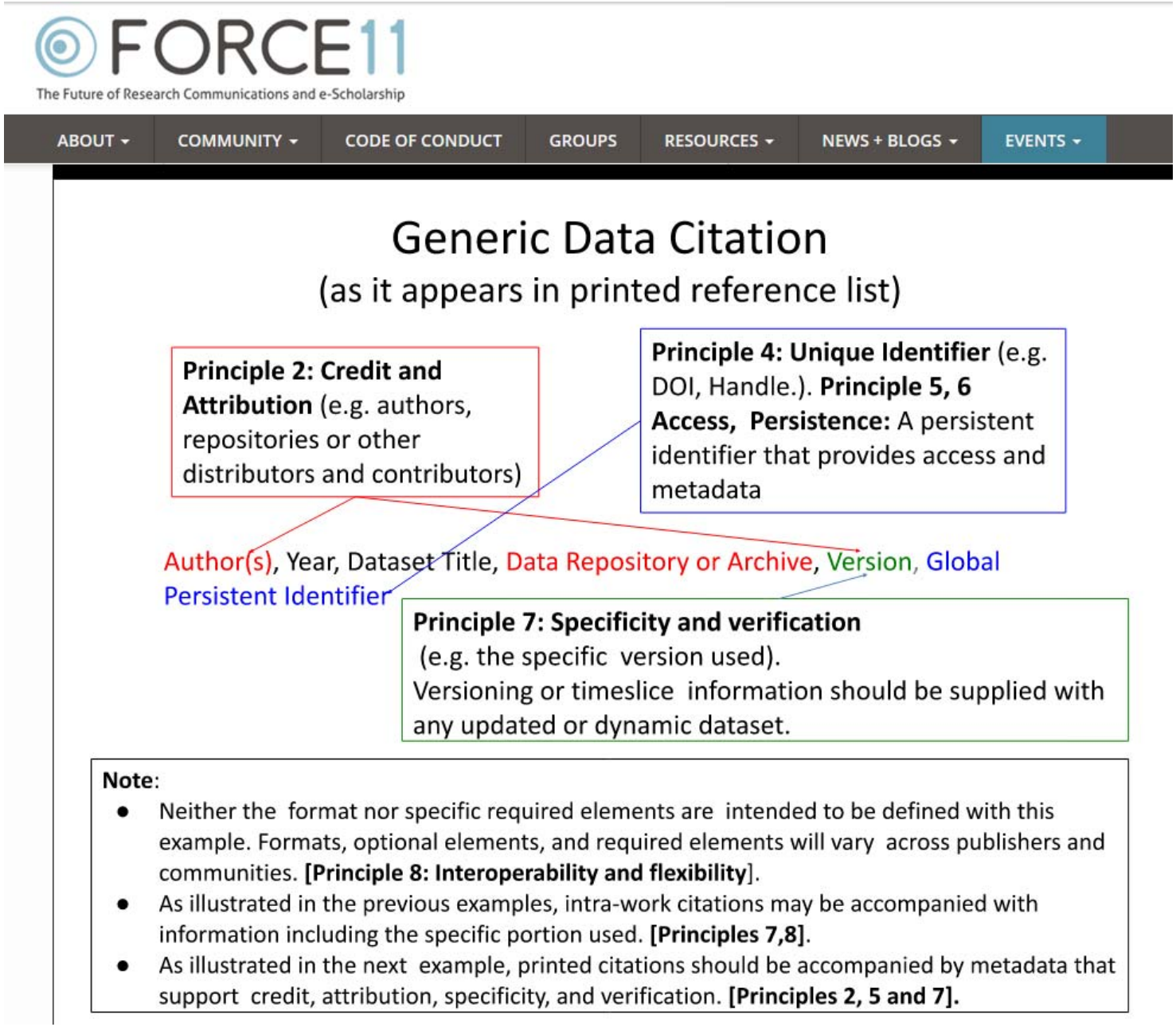
FORCE11
The Future of Research Communications and e-Scholarship

DC[1]
Data Citation Principles

1. Importancia. Los datos deben considerarse productos de investigación legítimos y citables
2. Crédito y atribución. Las citas de datos deben facilitar la concesión de crédito académico y atribución normativa y legal a todos los contribuyentes a los datos

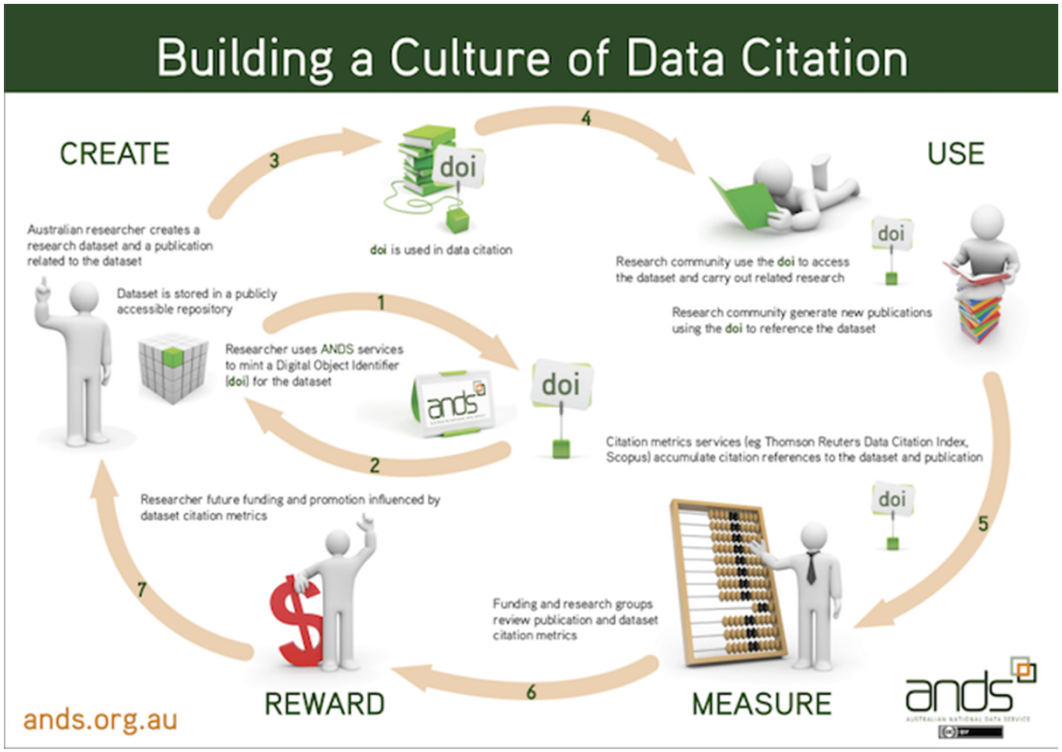
3. Evidencia. siempre y cuando una afirmación se base en datos, se deben citar los datos correspondientes
4. Identificación única. Una cita de datos debe incluir un identificador único y persistente y legible por máquinas
5. Acceso. Las citas de datos deben facilitar el acceso a los propios datos y a los metadatos, tanto a las personas como a las máquinas
6. Persistencia. Los identificadores únicos y los metadatos que describen los datos y su disposición deben persistir, incluso más allá de la vida útil de los datos que describen **7.**
7. Especificidad y verificabilidad. Las citas de datos deben facilitar la identificación, el acceso y la verificación de los datos específicos que respaldan una afirmación
8. Interoperabilidad y flexibilidad. Los métodos de citación de datos deben ser lo suficientemente flexibles para adaptarse a las prácticas de comunidades científicas, sin comprometer la interoperabilidad de las prácticas de citación de datos entre estas comunidades

https://www.force11.org/datacitationprinciples



https://www.force11.org/node/4771

http://www.ands.org.au/working-with-data/citation-and-identifiers/data-citation



También si no tienes DOI

A standard citation includes the following elements:

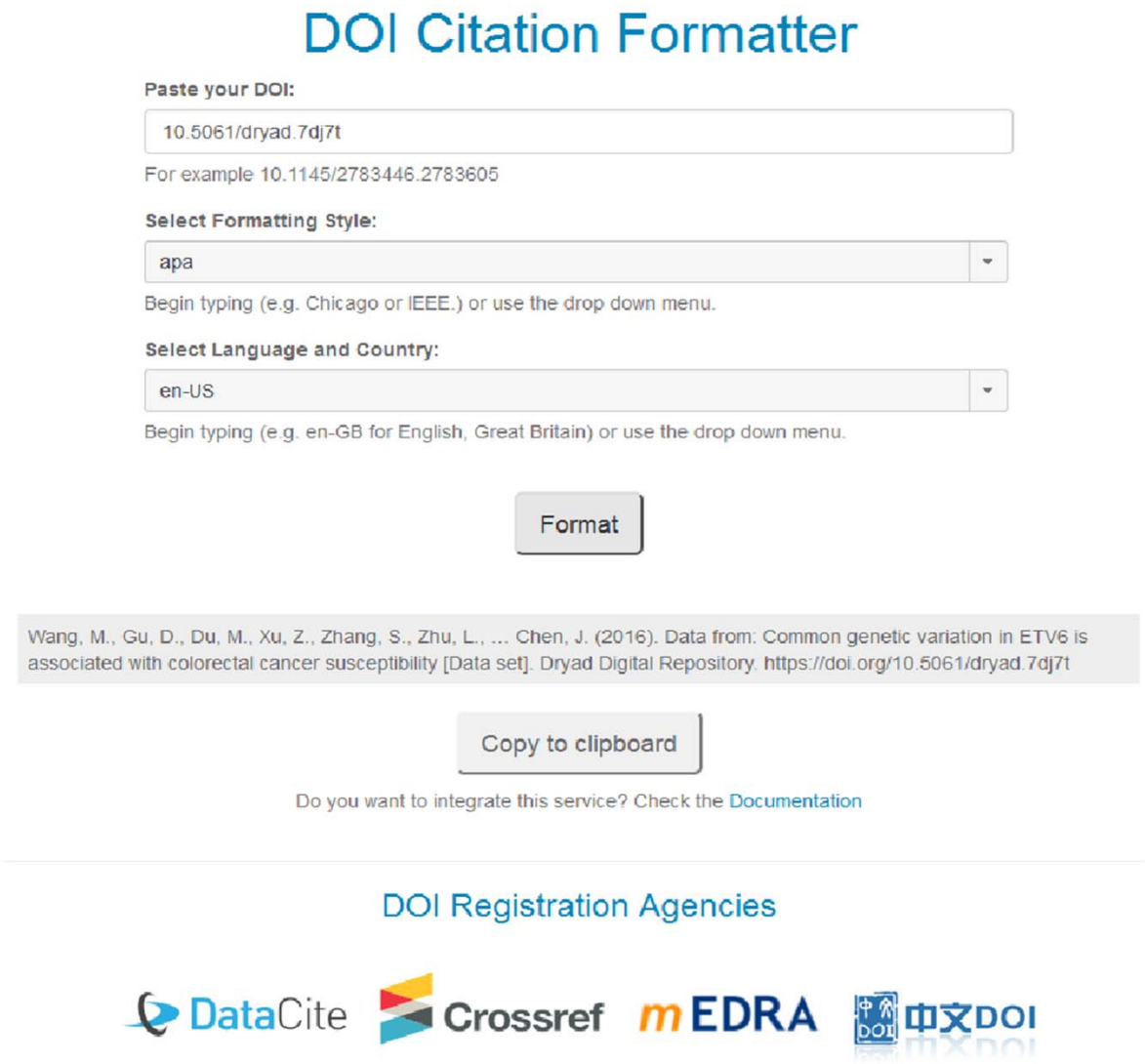
Autores (año): Título. Lugar de publicación. DOI (if used)

Ejemplo

Hanigan, Ivan (2012): Monthly drought data for Australia 1890-2008 using the Hutchinson Drought Index. The Australian National University Australian Data Archive. http://doi.org/10.4225/13/50BBFD7E6727A

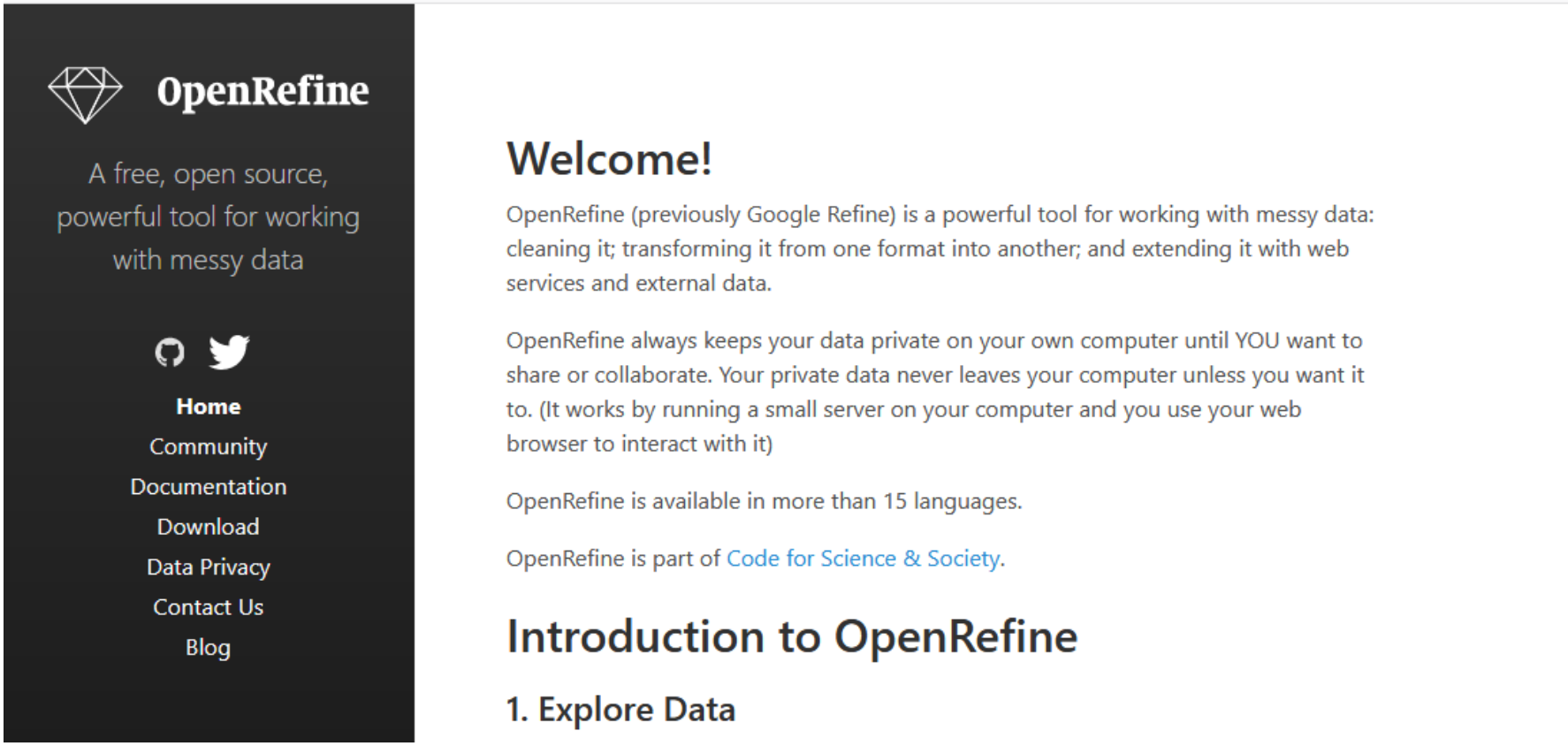
Ejemplo de formato de cita al usar el DOI Citation formatter

http://citation.crosscite.org/



https://data.research.cornell.edu/content/data-citation

https://openrefine.org/



Metadatos

DCC because good research needs good data

Search

Home | Digital curation | About us | News | Events | Resources | Training | Projects | Community | Tailored support

Home > Resources > Metadata Standards

Metadatos por disciplina

http://www.dcc.ac.uk/resources/metadata-standards

In this section

- Briefing Papers
- How-to Guides & Checklists
- Developing RDM Services
- Curation Lifecycle Model
- Curation Reference Manual
- Policy and legal
- Data Management Plans
- Tools
- Case studies
- Repository audit and assessment
- **Standards**
  - Disciplinary Metadata
  - DIFFUSE
- Publications and presentations
- Roles
- Curation journals
- Informatics research
- External resources
- Online Store

Disciplinary Metadata

While data curators, and increasingly researchers, know that good metadata is key for research data access and re-use, figuring out precisely what metadata to capture and how to capture it is a complex task. Fortunately, many academic disciplines have supported initiatives to formalise the metadata specifications the community deems to be required for data re-use. This page provides links to information about these disciplinary metadata standards, including profiles, tools to implement the standards, and use cases of data repositories currently implementing them.

For those disciplines that have not yet settled on a metadata standard, and for those repositories that work with data across disciplines, the General Research Data section links to information about broader metadata standards that have been adapted to suit the needs of research data.

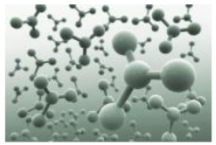
Please note that a community-maintained version of this directory has been set up under the auspices of the Research Data Alliance.

Search by Discipline

Biology

Earth Science

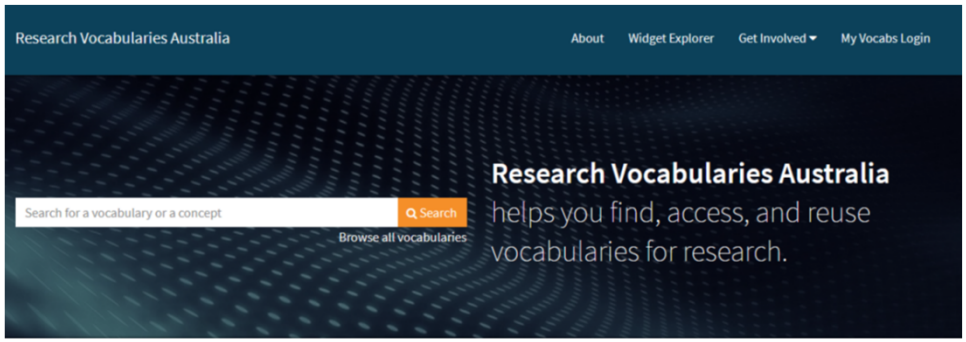
General Research Data



Physical Science

Social Science & Humanities

Vocabularios controlados por materias, p.e., https://vocabs.ands.org.au/



FAIRsharing.org
standards, databases, policies

Search all of FAIRsharing | Standards | Databases | Policies | Collections | Add/Claim Content | Stats | Log in or Regist

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

Recommendatic

Standards and/or databases recommended by j policies.

FAIRsharing.org
standards, databases, policies

Search all of FAIRsharing | Standards | Databases | Policies | Collections | Add/Claim Content

View as Table | View as Grid

Sort by: Name

Recommended Records: Recommended

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

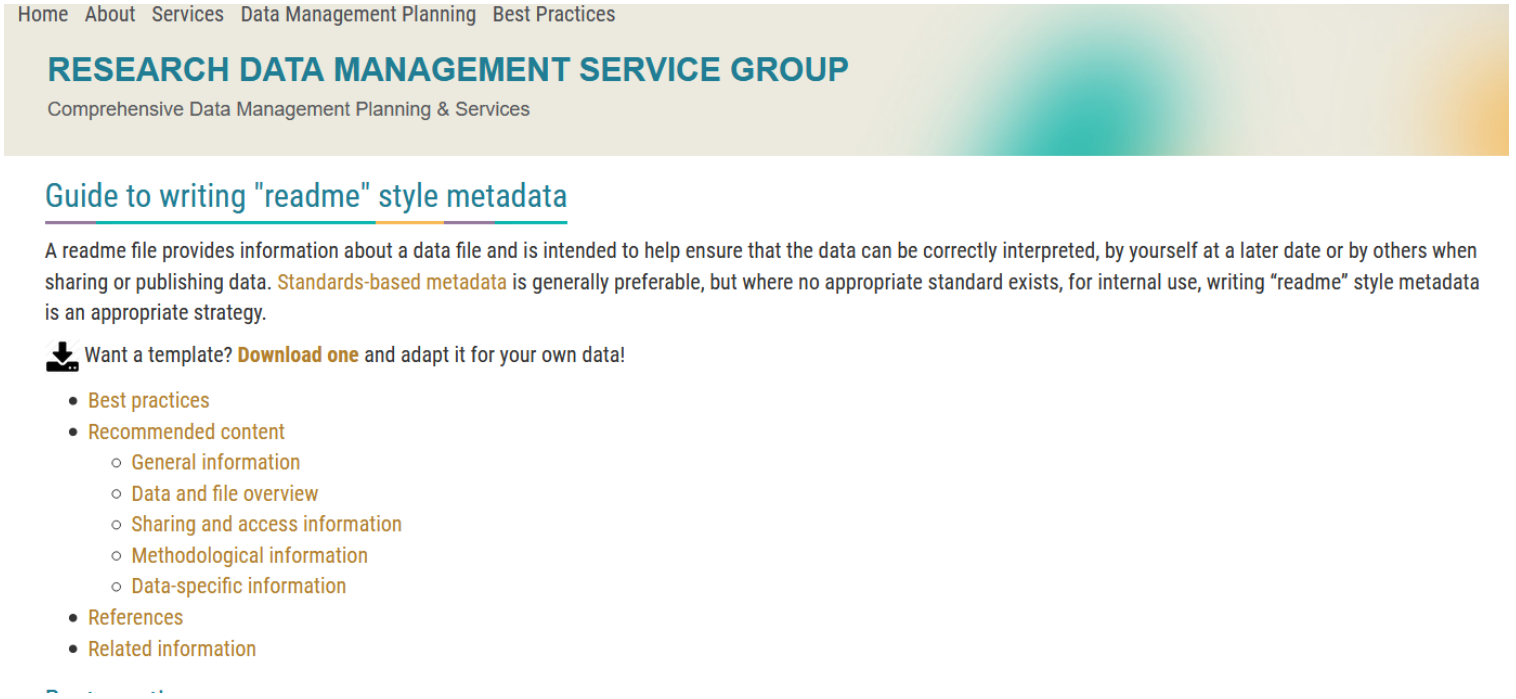
Domains
- Bibliography 52
- Nucleic Acid Sequence 3

Showing records 1 - 50 of 140.

« 1 2 3 »

Registry	Name	Abbreviation	Type	Subject	Domain	Taxonomy	Related Database	Related Standard	Related Policy	In Collectio
	National Child Development Study 1958BC data deposition policy	NCDS Policy	Project	Biomedical Science, Preclinical Studies	None	Homo sapiens	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)	National
	African Academy of Sciences Open Research Data Guidelines for Authors	AAS Policy	Journal	Biomedical Science, Life Science	Journal Article	All	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	Biology, Biomedical Science, Comparative Genomics, Epigenomics, Functional Genomics, Plus 3 more...	Annotation, Knowledge Representation, Protocol	All	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	Biomedical Science, Life Science	None	All	ClinicalTrials.gov	None	None	None
	American Association for the Advancement of Science - Science - Editorial Policies; Data Materials Availability	AAAS Science Policy	Journal	Biomedical Science, Environmental Science, Life Science	Bibliography, Journal Article, Publication	All	GenBank, CSD, ArrayExpress, ENA, PDBe, Plus 7 more...	CONSORT, MIAME	None	None

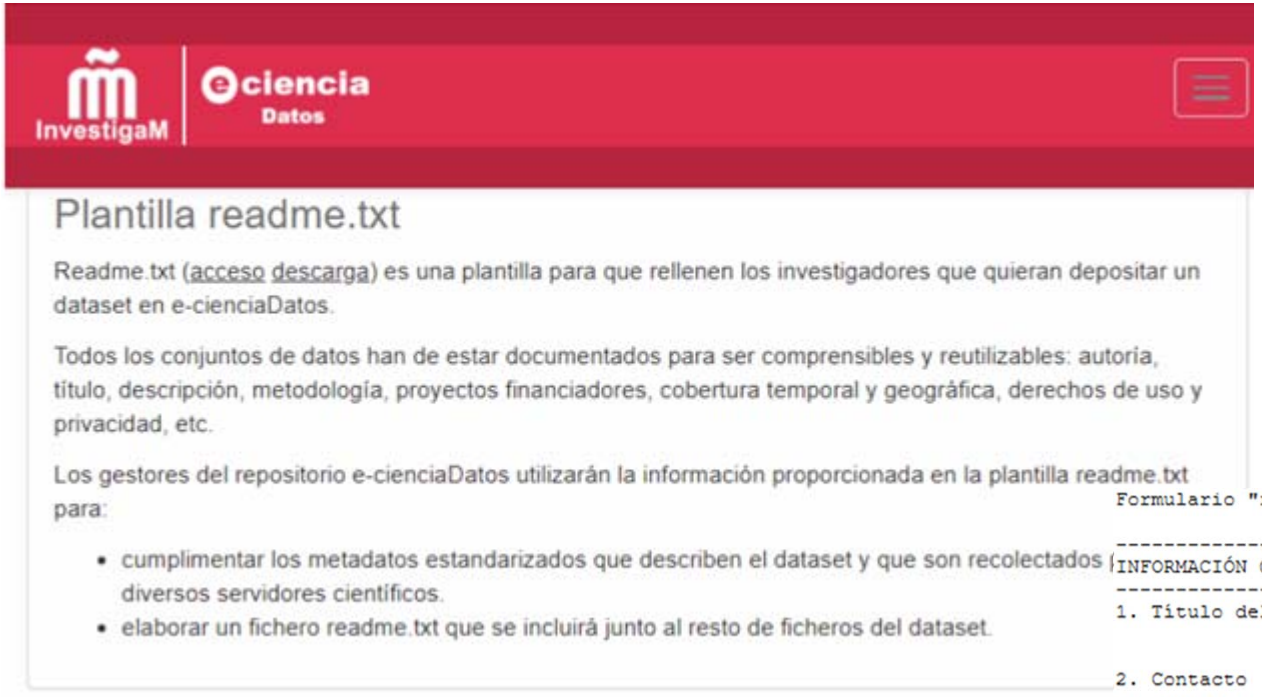
https://data.research.cornell.edu/content/readme




```
This DATSETNAMEreadme.txt file was generated on YYYY-MM-DD by NAME
<help text is included in angle brackets, and can be deleted before saving>


GENERAL INFORMATION

1. Title of Dataset:

2. Author Information
        A. Principal Investigator Contact Information
                Name:
                Institution:
                Address:
                Email:

        B. Associate or Co-investigator Contact Information
                Name:
                Institution:
                Address:
                Email:

        C. Alternate Contact Information
                Name:
                Institution:
                Address:
                Email:

3. Date of data collection (single date, range, approximate date) <suggested format YYYY-MM-DD>:

4. Geographic location of data collection <latitude, longiute, or city/region, State, Country, as appropriate>:

5. Information about funding sources that supported the collection of the data:
```

https://data.research.cornell.edu/content/readme#introductory

https://edatos.consorciomadrono.es/readme.xhtml



https://edatos.consorciomadrono.es/resources/txt/readme-es.txt

```
Formulario "readme" en e-cienciaDatos

--------------------
INFORMACIÓN GENERAL
--------------------
1. Título del dataset


2. Contacto

  Investigador/a de contacto
        Nombre:
        Filiación:
        Correo electrónico:
        ORCID:

3. Descripción del proyecto


4. Descripción del dataset


5. Notas


6. Fecha de depósito de los ficheros


7. Fecha de creación de los ficheros


8. Idioma
```

Tratamiento de datos personales y confidenciales

Data Ethics Canvas

2019-06

Data sources

Name/describe your project's key data sources, whether you're collecting data yourself or accessing via third parties.

Is any personal data involved, or data that is otherwise sensitive?

Limitations in data sources

Are there limitations that could influence your project's outcomes?

Consider:

bias in data collection, inclusion/exclusion, analysis, algorithms

gaps or omissions in data

provenance and data quality

other issues affecting decisions, such as team composition

Sharing data with others

Are you going to be sharing data with other organisations? If so, who?

Are you planning to publish any of the data? Under what conditions?

Ethical and legislative context

What existing ethical codes apply to your sector or project? What legislation, policies, or other regulation shape how you use data? What requirements do they introduce?

Consider: the rule of law; human rights; data protection; IP and database rights; anti-discrimination laws; and data sharing, policies, regulation and ethics codes/ frameworks specific to sectors (eg health, employment, taxation).

Rights around data sources

Where did you get the data from? Is it produced by an organisation or collected directly from individuals?

Was the data collected for this project or for another purpose? Do you have permission to use this data, or another basis on which you're allowed to use it? What ongoing rights will the data source have?

Your reason for using data

What is your primary purpose for collecting and using data in this project?

What are your main use cases? What is your business model?

Are you making things better for society? How and for whom?

Are you replacing another product or service as a result of this project?

Communicating your purpose

Do people understand your purpose – especially people who the data is about or who are impacted by its use?

How have you been communicating your purpose? Has this communication been clear?

How are you ensuring more vulnerable individuals or groups understand?

Positive effects on people

Which individuals, groups, demographics or organisations will be positively affected by this project? How?

How are you measuring and communicating positive impact? How could you increase it?

Negative effects on people

Who could be negatively affected by this project?

Could the way that data is collected, used or shared cause harm or expose individuals to risk of being re-identified? Could it be used to target, profile or prejudice people, or unfairly restrict access (eg exclusive arrangements)?

How are limitations and risks communicated to people? Consider: people who the data is about, people impacted by its use and organisations using the data.

Minimising negative impact

What steps can you take to minimise harm?

How could you reduce any limitations in your data sources? How are you keeping personal and other sensitive information secure?

How are you measuring, reporting and acting on potential negative impacts of your project?

What benefits will these actions bring to your project?

Engaging with people

How can people engage with you about the project?

How can people correct information, appeal or request changes to the product/service? To what extent?

Are appeal mechanisms reasonable and well understood?

Openness and transparency

How open can you be about this project? Could you publish your methodology, metadata, datasets, code or impact measurements?

Can you ask peers for feedback on the project? How will you communicate it internally?

Will you publish your actions and answers to this canvas openly?

Ongoing implementation

Are you routinely building in thoughts, ideas and considerations of people affected in your project? How?

What information or training might be needed to help people understand data issues?

Are systems, processes and resources available for responding to data issues that arise in the long-term?

Review and iteration

How will ongoing data ethics issues be measured, monitored, discussed and actioned?

How often will your responses to this canvas be reviewed or updated? When?

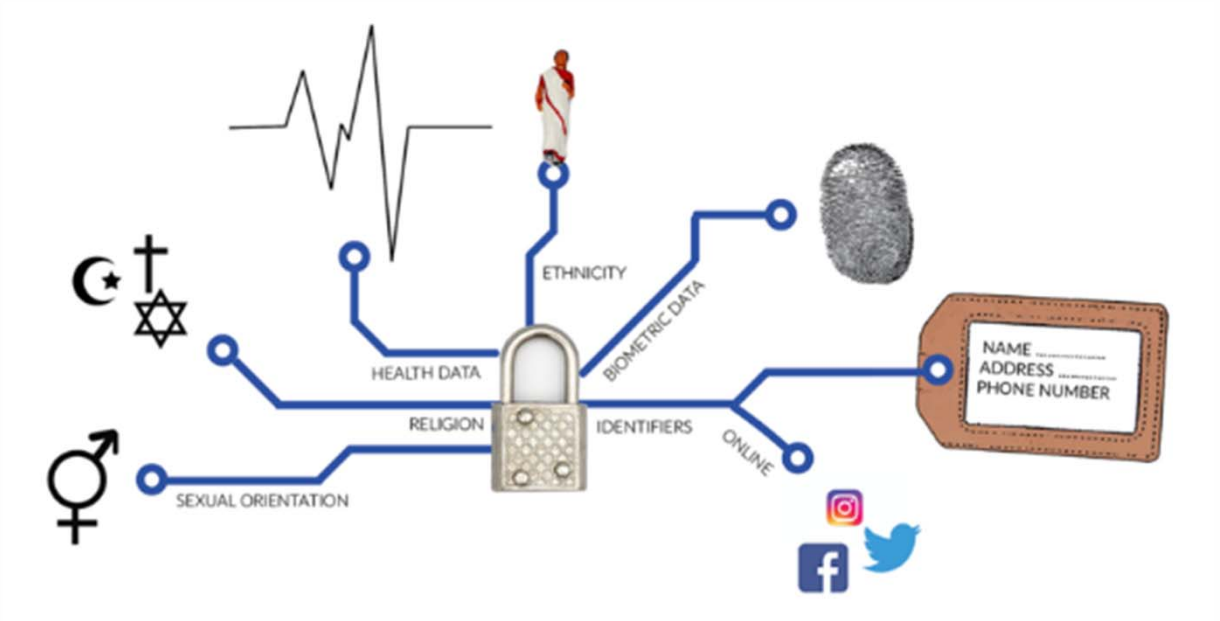
Your actions

What actions will you take before moving forward with this project? Which should take priority?

Who will be responsible for these actions, and who must be involved?

Will you openly publish your actions and answers to this canvas?

El término **datos personales** se refiere a cualquier información que pueda usarse para identificar individuos vivos (y, a veces, incluso fallecidos). Por ejemplo, su nombre o fecha de nacimiento. Algunos datos personales, como el origen, la opinión política, las creencias religiosas, la salud, la afiliación sindical o la orientación sexual de una persona, se clasifican como **datos personales confidenciales**. Si se manejan datos personales o confidenciales como parte de su investigación, debemos asegurarnos de que están protegidos.



https://www.cessda.eu/Training/Training-Resources/Library/Data-Management-Expert-Guide/5.-Protect/Ethics-and-data-protection

Cuestiones previas a la recopilación de datos personales

¿Qué datos personales recopilará?
Identifique qué datos personales se recopilarán y utilizarán para su investigación. Por ejemplo, ¿Reunirá los nombres y direcciones de los participantes de la investigación?

¿Cómo se cubrirán los costes de anonimización?
Discuta si realmente necesita recopilar datos personales para llevar a cabo su investigación. Tenga en cuenta que cualquier información personal innecesaria recopilada y agregada a su conjunto de datos probablemente deberá eliminarse o anonimizarse más tarde si planea compartirla. Esto puede ser costoso. En algunos casos, la limpieza de datos puede costar más que la recolección de datos en sí.

¿Realmente necesita recopilar datos personales?
Tómese siempre el tiempo necesario para volver a considerar si hay formas en que puede recopilar datos de forma anónima. Esto podría ahorrarle muchos dolores de cabeza más tarde. Así que solo se compromete a recopilar datos personales si es esencial para su investigación.

Al recopilar y / o manejar datos personales, los investigadores deben seguir una serie de principios que incluyen:

Transparencia: procesamiento de datos personales "de manera legal, justa y transparente"

Minimización de datos: el uso de los datos se limitará al propósito de la investigación respectiva.

Precisión: los datos inexactos deben ser "borrados o rectificados sin demora". Integridad y confidencialidad: los datos deben estar protegidos por medidas de seguridad adecuadas (técnicas y organizativas).

Con el nuevo Reglamento General de Protección de Datos (GDPR), la Unión Europea proporciona un marco legal para la protección de datos dentro de la UE, así como para la exportación de estos datos fuera de sus fronteras.

En cuanto al tratamiento de datos personales, el GDPR incluye una exención para el ulterior uso con fines de investigación:

Si se trata de "aspectos de interés público, investigación científica o histórica o fines estadísticos" (Art. 5.1 2016/679 / UE):

Artículo 5 **Principios relativos al tratamiento**

1. *Los datos personales serán: a) tratados de manera lícita, leal y transparente en relación con el interesado («licitud, lealtad y transparencia»); b) recogidos con fines determinados, explícitos y legítimos, y no serán tratados ulteriormente de manera incompatible con dichos fines; de acuerdo con el artículo 89, apartado 1, el tratamiento ulterior de los datos personales con fines de archivo en interés público, fines de investigación científica e e histórica o fines estadísticos no se considerará incompatible con los fines iniciales («limitación de la finalidad»);*

En resumen, la recolección de datos de investigación que incluyan datos sensibles ***con tratamientos de poca complejidad***

Requieren:

- **Consentimiento inequívoco (no vale genérico)**
- **Base legal**
- **Finalidad (¿para qué?)**
- **Análisis de riesgos**
- **Transparencia en la información a los interesados**
- **Medidas de seguridad y preservación (en caso necesario)**

Reglamento General de Protección de Datos (RGPD) de 28 de mayo de 2018

Ley Orgánica 3/2018, de 5 de diciembre, de Protección de Datos Personales y garantía de los derechos digitales.

https://www.aepd.es/infografias/index.html

El **consentimiento informado** es el procedimiento mediante el cual se garantiza que el sujeto ha expresado voluntariamente su intención de participar en la investigación, después de haber comprendido la información que se le ha dado, acerca de los objetivos del estudio, los beneficios, las molestias, los posibles riesgos y las alternativas, sus derechos y responsabilidades

(https://es.m.wikipedia.org/wiki/Consentimiento_informado)

Formulario de consentimiento

El formulario de consentimiento debe estar escrito en un lenguaje sencillo, sin jerga, y debe permitir que el participante responda claramente a puntos como:

- El participante ha leído y entendido la información sobre el proyecto.
- El participante ha tenido la oportunidad de hacer preguntas.
- El participante acepta voluntariamente participar en el proyecto.
- El participante entiende que puede retirarse en cualquier momento sin dar razones y sin penalización.
- Cómo se protegerá la confidencialidad, p. si se usarán nombres reales o seudónimos, cómo se anonimizarán los datos, etc.
- Términos de consentimiento separados para datos que pueden tener diferentes riesgos de divulgación, p. transcripciones de entrevistas anónimas, grabaciones de audio, videos, fotos
- Cómo se utilizarán los datos en las publicaciones
- Si los participantes están de acuerdo con el archivo y la reutilización de datos
- Firmas y fechas de firma para el participante y el investigador.
- El participante debe recibir una copia del formulario y el investigador debe conservar el original firmado.

https://www.ukdataservice.ac.uk/manage-data/legal-ethical/consent-data-sharing/consent-forms

https://www.ukdataservice.ac.uk/manage-data/tools-and-templates.aspx

Consent Form for [name of project]

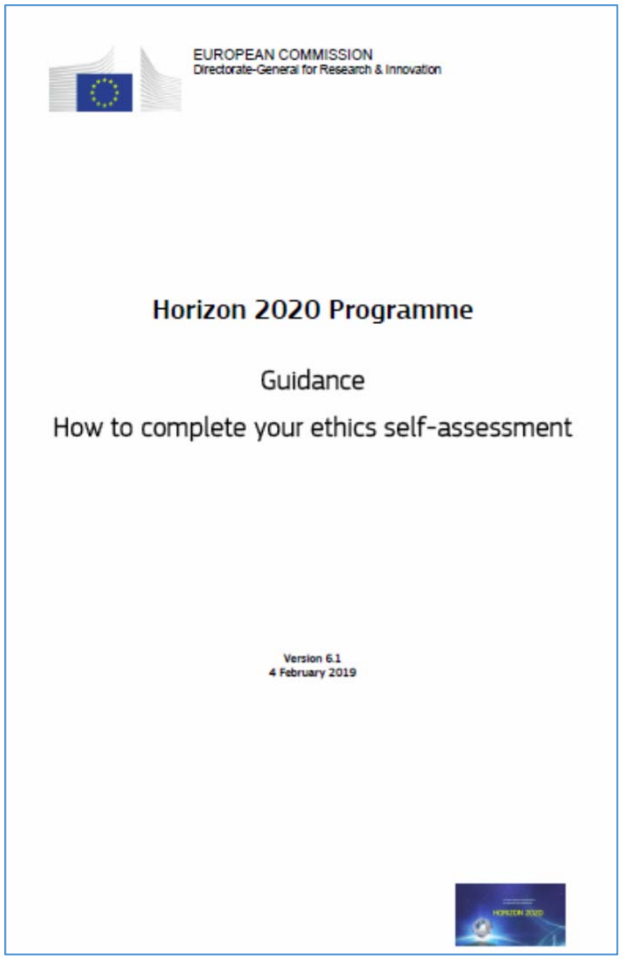
Please tick the appropriate boxes	Yes	No
Taking Part		
I have read and understood the project information sheet dated DD/MM/YYYY.	☐	☐
I have been given the opportunity to ask questions about the project.	☐	☐
I agree to take part in the project. Taking part in the project will include being interviewed and recorded (audio or video).[1]	☐	☐
I understand that my taking part is voluntary; I can withdraw from the study at any time and I do not have to give any reasons for why I no longer want to take part.	☐	☐
Use of the information I provide for this project only		
I understand my personal details such as phone number and address will not be revealed to people outside the project.	☐	☐
I understand that my words may be quoted in publications, reports, web pages, and other research outputs.	☐	☐
*Please choose **one** of the following two options:*		
I would like my real name used in the above	☐	
I would **not** like my real name to be used in the above.	☐	
Use of the information I provide beyond this project		
I agree for the data I provide to be archived at the UK Data Archive.[2]	☐	☐
I understand that other authenticated researchers will have access to this data only if they agree to preserve the confidentiality of the information as requested in this form.	☐	☐
I understand that other authenticated researchers may use my words in publications, reports, web pages, and other research outputs, only if they agree to preserve the confidentiality of the information as requested in this form.	☐	☐
So we can use the information you provide legally		
I agree to assign the copyright I hold in any materials related to this project to [name of researcher].	☐	☐

Name of participant [printed] _____ Signature _____ Date _____

Researcher [printed] _____ Signature _____ Date _____

Project contact details for further information: Names, phone, email addresses, etc.

Notes:
1. Other forms of participation can be listed.
2. More detail can be provided here so that decisions can be made separately about audio, video, transcripts, etc.



EUROPEAN COMMISSION
Directorate-General for Research & Innovation

Horizon 2020 Programme

Guidance

How to complete your ethics self-assessment

Version 6.1
4 February 2019

https://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/ethics/h2020_hi_ethics-self-assess_en.pdf

4.1 Ethics issues checklist

Section 4: PROTECTION OF PERSONAL DATA		YES/NO		Page	Information to be provided	Documents to be provided/kept on file
Does your research involve processing of personal data?		☐	☐		1) Details of the technical and organisational measures to safeguard the rights of the research participants. For instance: For organisations that must appoint a DPO under the GDPR: Involvement of the data protection officer (DPO) and disclosure of the contact details to the research participants. For all other organisations: Details of the data protection policy for the project (i.e. project-specific, not general). 2) Details of the informed consent procedures. 3) Details of the security measures to prevent unauthorised access to personal data. 4) How is all of the processed data relevant and limited to the purposes of the project ('data minimisation' principle)? Explain. 5) Details of the anonymisation /pseudonymisation techniques. 6) Justification of why research data will not be anonymised/ pseudonymised (if relevant). 7) Details of the data transfers (type of data transferred and country to which it is transferred – for both EU and non-EU countries).	1) Informed Consent Forms + Information Sheets used (if relevant).
If YES:	- Does it involve the processing of special categories of personal data *(e.g. genetic, health, sexual lifestyle, ethnicity, political opinion, religious or philosophical conviction)*?	☐	☐		1) Justification for the processing of special categories of personal data. 2) Why can the research objectives not be reached by processing anonymised/ pseudonymised data (if applicable)?	
	- Does it involve processing of genetic, biometric or health data?	☐	☐			1) Declaration confirming compliance with the laws of the country where the data was collected.
	- Does it involve profiling, systematic monitoring of individuals or processing of large scale of special	☐	☐		1) Details of the methods used for tracking, surveillance or observation of participants. 2) Details of the methods used for profiling. 3) Risk assessment for the data processing activities.	1) Opinion of the data controller on the need for a data protection impact assessment (art.35 GDPR) (if relevant).

17

https://www.ukdataservice.ac.uk/manage-data/tools-and-templates.aspx

UK Data Service

About us | Get data | Use data | Manage data | Deposit data

Home > Manage data > Tools and templates

- Data lifecycle
- Plan to share
- Legal and ethical
- Rights
- Document your data
- Format your data
- Store your data
- Collaborative research
- Training
- Tools and templates
- Handbook

Tools and templates

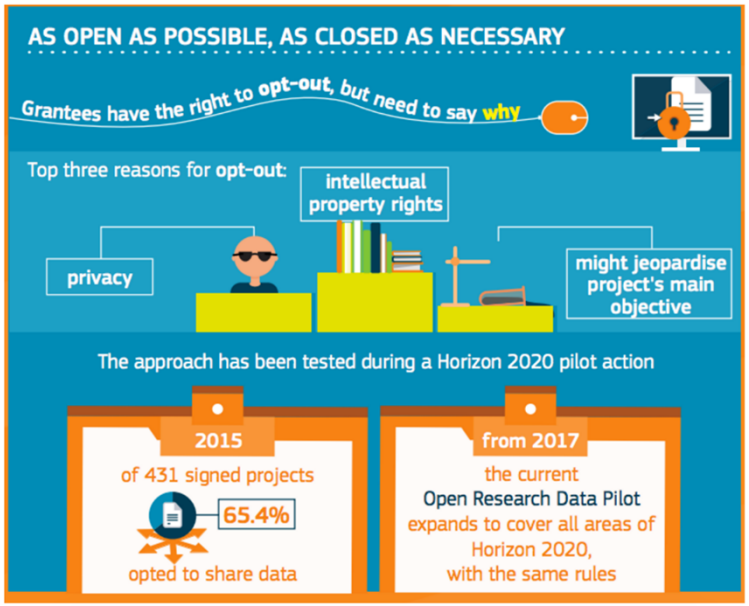
"Providing researchers with access to the tools they need"

SHARE

This is our selection of tools and templates that researchers may find useful for various data management tasks in social sciences research:

- Model consent form (doc) that takes into account consent for data sharing and future data reuse
- Sample survey consent statement (doc) that considers consent for data sharing and future data reuse

https://amnesia.openaire.eu/



Anonimización de datos

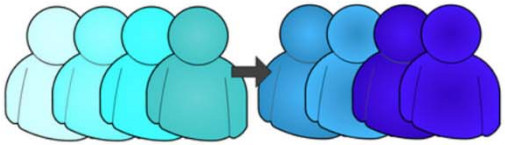
amnesia

Home Get Amnesia! What is Amnesia? Documentation On-line version About

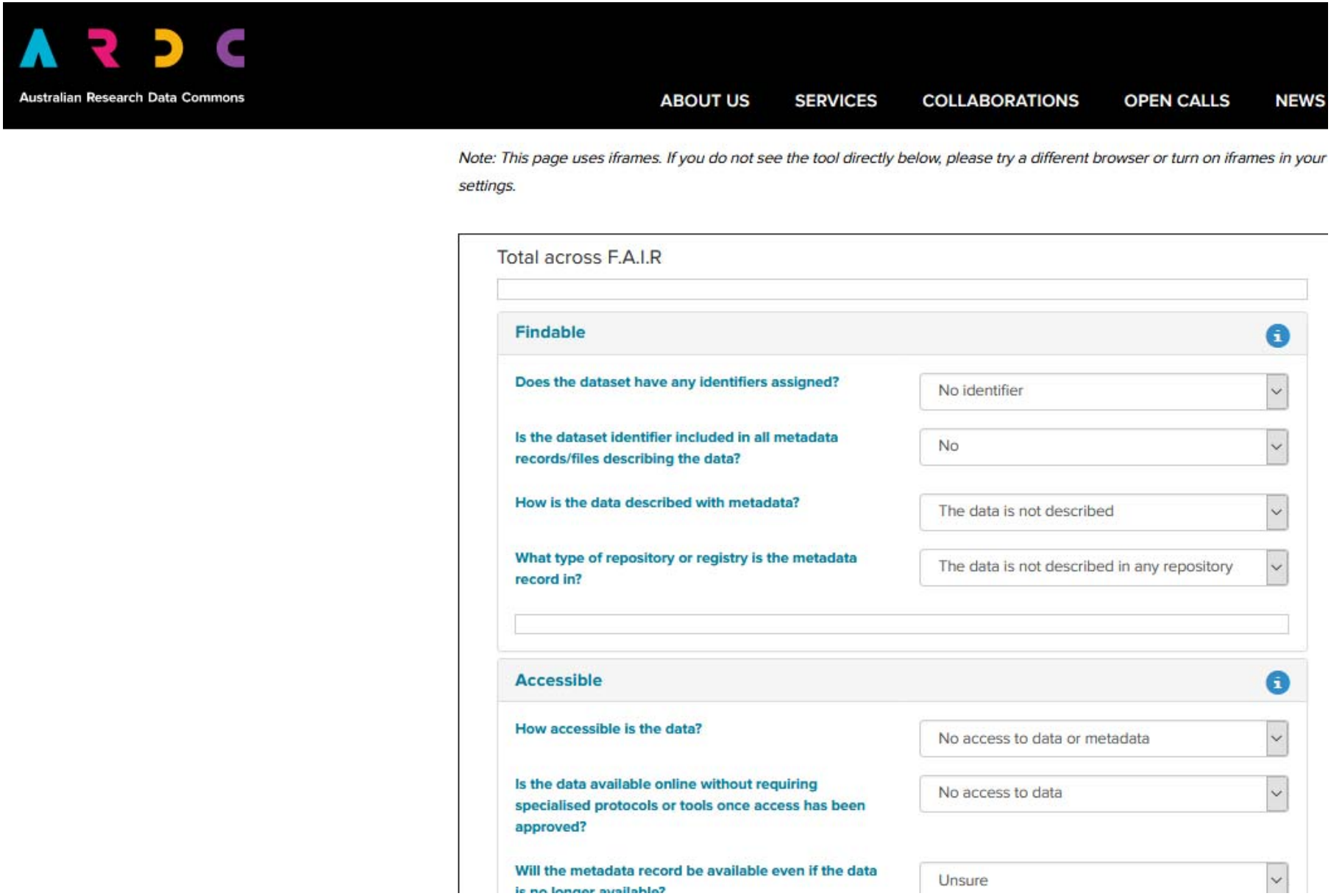
Amnesia

Amnesia is a data anonymization tool, that allows to remove identifying information from data. Amnesia not only removes direct identifiers like names, SSNs etc but also transforms secondary identifiers like birth date and zip code so that individuals cannot be identified in the data. Amnesia supports k-anonymity and k^m-anonymity.

amnesia



¿Cumplen mis datos con los principios FAIR?



https://ardc.edu.au/resources/working-with-data/fair-data/fair-self-assessment-tool/

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://satifyd.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.

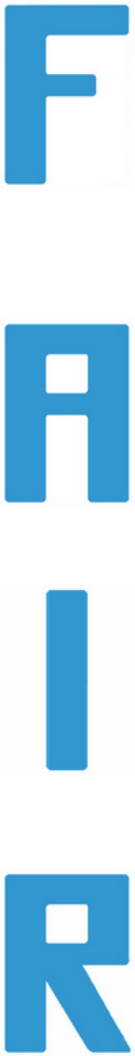
The 12 questions touch upon the FAIR data principles but do not strictly follow them. While answering the questions, the score per letter will be displayed underneath each letter. The more 'blue' the letters get, the more FAIR your dataset is. An overall score is provided at the end of the page.

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

Formulario de FAIR AWARE



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

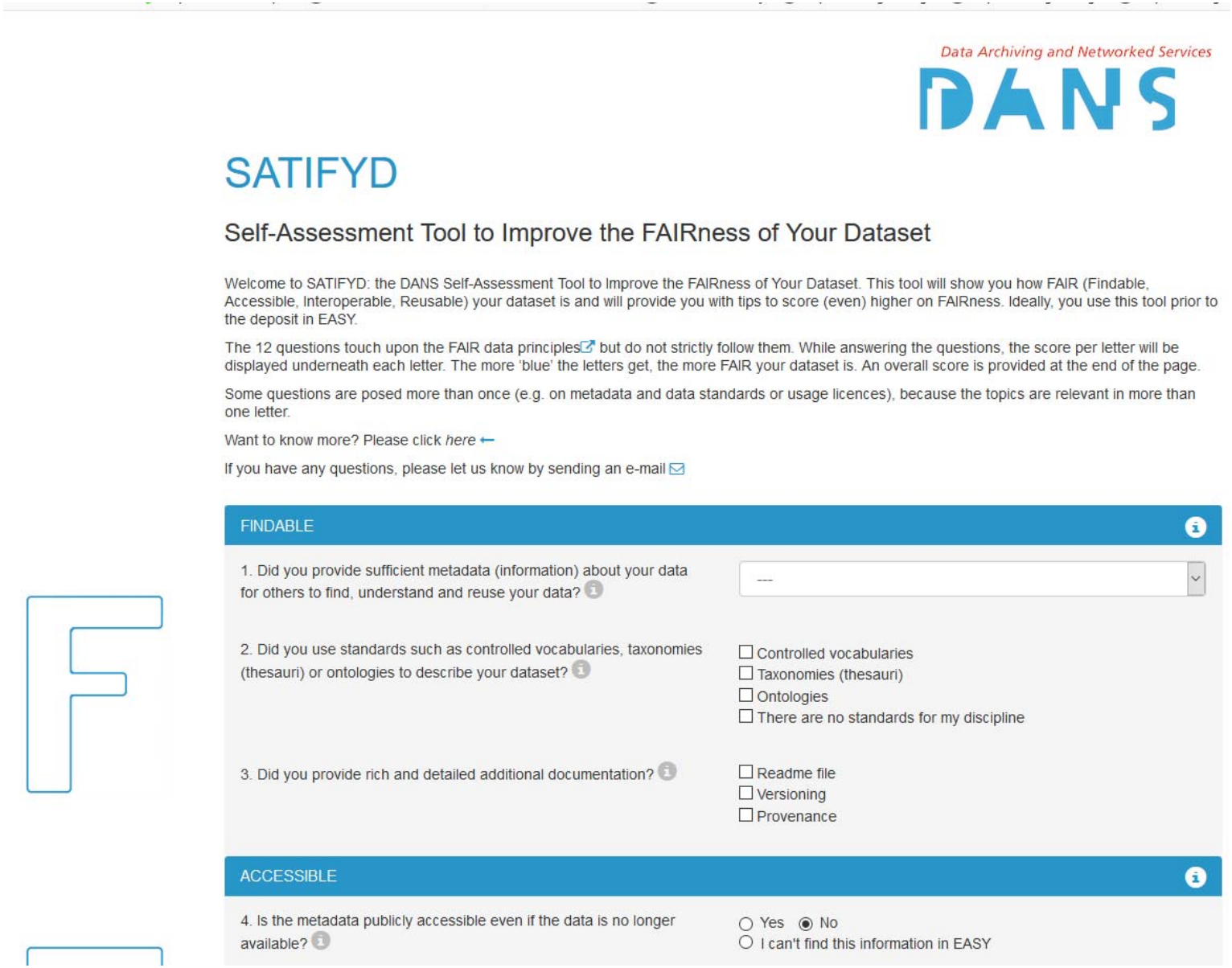
7. Are you aware that provenance information about the collection and/or generation of data should be included in the metadata? — Yes / No
8. Are you aware that metadata describing your data should follow the specifications of a community-endorsed standard? — Yes / No
9. Are you aware that data should 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

Feedback

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

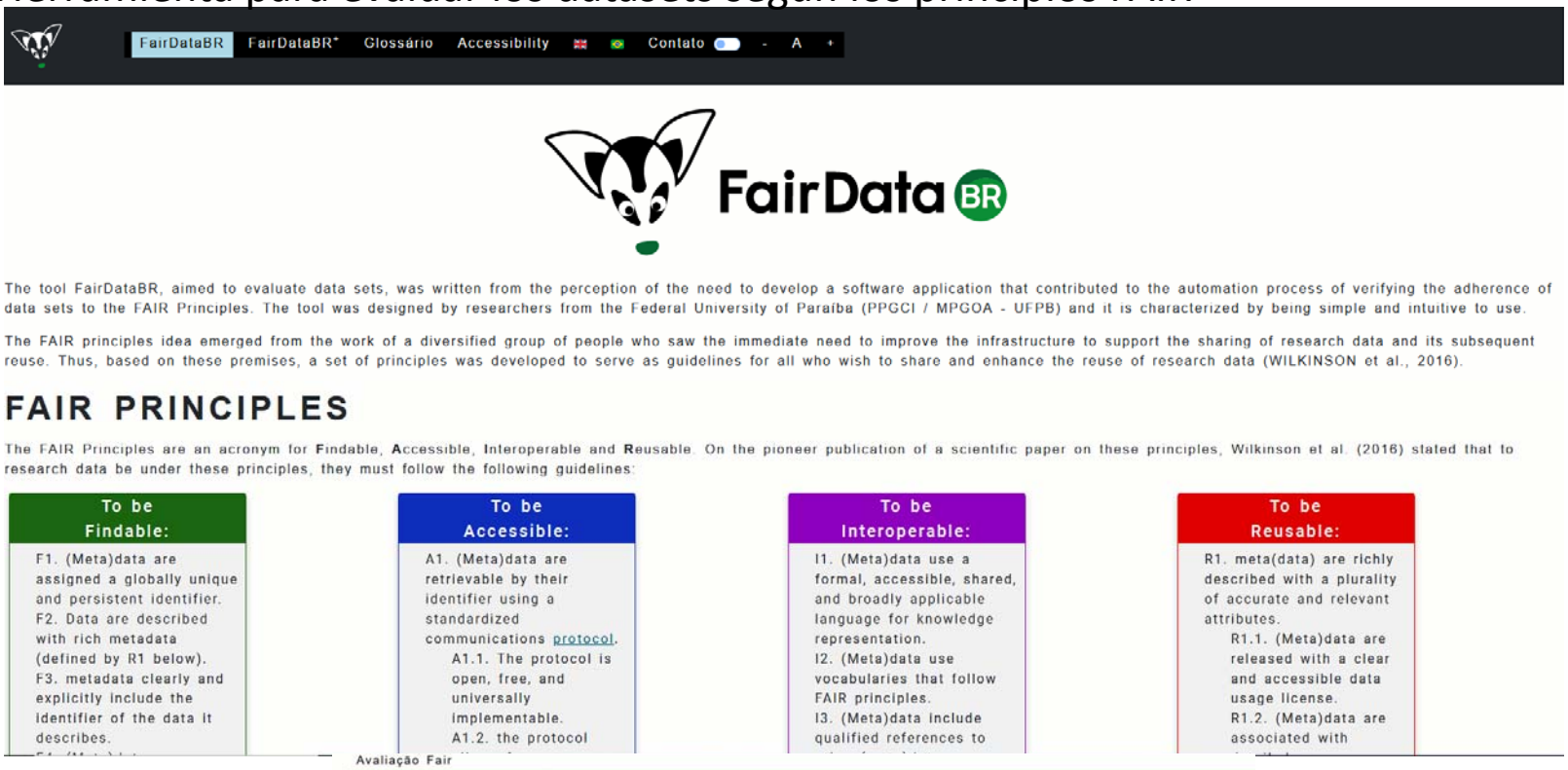
- Which of the following issues do you find hard to understand and/or to answer?
  1. Globally unique persistent identifier (PID) ☐
  2. Metadata for citation and discovery including PID ☐
  3. Metadata available online is readable by humans and machines ☐
  4. Metadata includes licence, level of access 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.)

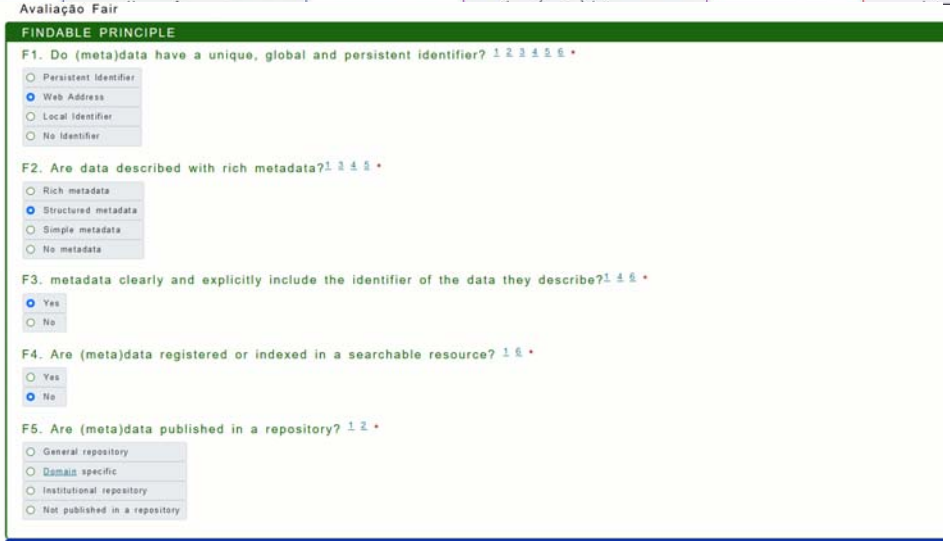
https://satifyd.dans.knaw.nl/



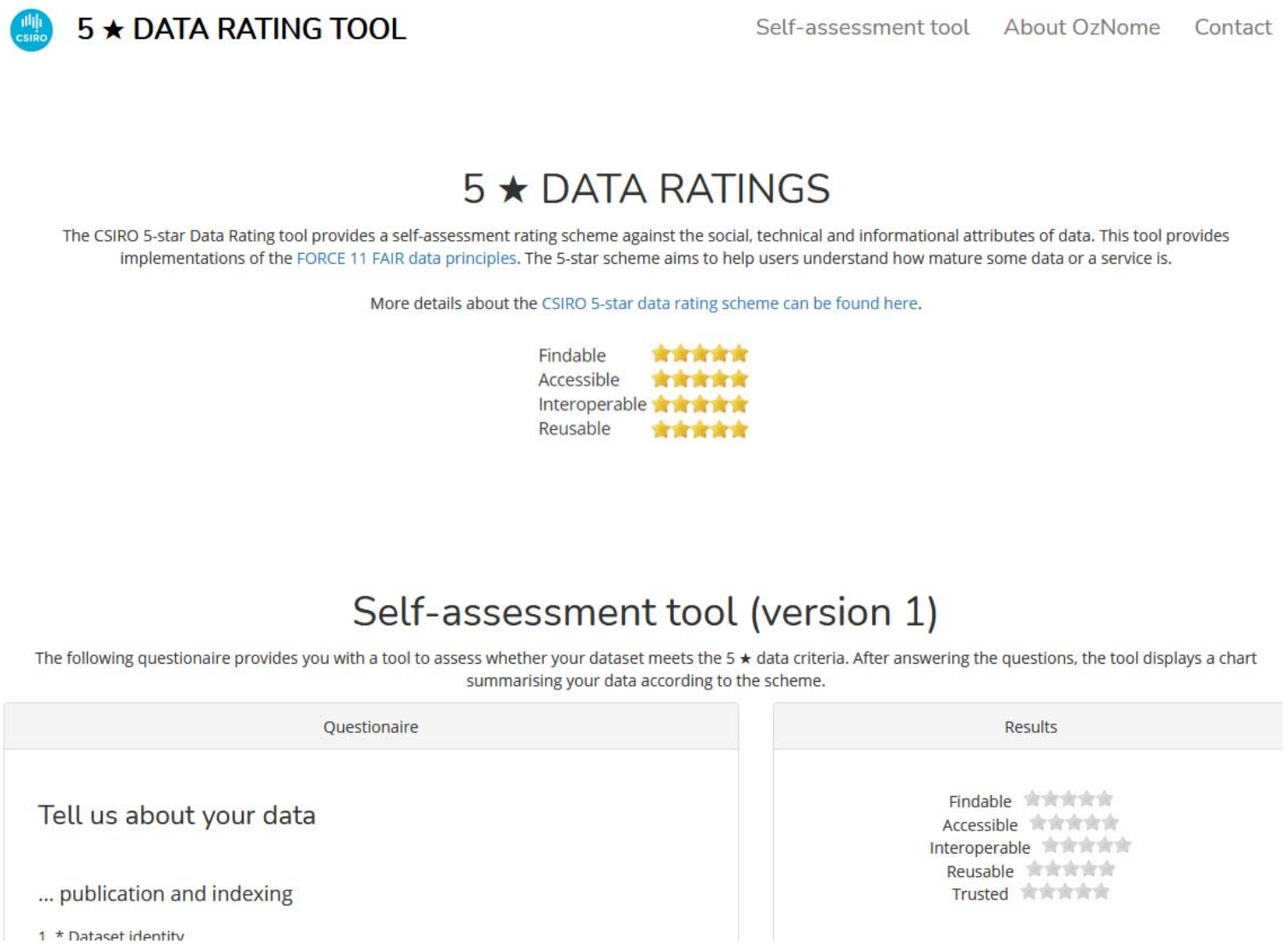
https://wrco.ufpb.br/fair/

Herramienta para evaluar los datasets según los principios FAIR





http://5stardata.csiro.au/

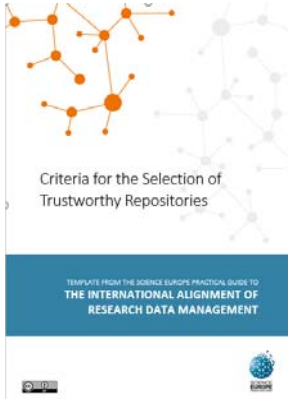


https://fairassist.org/

FAIRassist is the new, under development, educational component of the well established FAIRsharing resource.
FAIRassist is being designed to **offer personalised guidance to all stakeholders to discover standards and repositories in FAIRsharing**, which should be used to make data FAIR, as well as signpost other resources that enable FAIRness. FAIRassist is under development, and it will implement a phased rollout of its content, working with and for the community. As an initial step, we are collecting and describing existing resources for the assessment and/or evaluation of digital objects against the FAIR principles, which are aspirational. The focus is on manual *questionnaires*, *checklists* and automated *tests* that help users understand how to achieve a state of "FAIRness", and how this can be measured and improved. This is not intended to be a comprehensive list of all groups, projects and organizations that tackle FAIRness or FAIRification. If your resource is missing, you can submit it via the "Register resource" button below.

- Dónde localizar y depositar datasets

Criteria for the Selection of Trustworthy Repositories. Trustworthy repositories should meet the following minimum criteria:



Provision of Persistent and Unique Identifiers (PID)

- Allow data discovery and identification
- Enable searching, citing, and retrieval of data
- Provide support for data versioning

Metadata

- Enable finding of data
- Enable referencing to related relevant information, such as other data and publications
- Provide information that is publicly available and maintained, even for non-published, protected, retracted, or deleted data
- Use metadata standards that are broadly accepted (by the scientific community)
- Ensure that metadata are machine-retrievable

Data Access and Usage Licences

- Enable access to data under well-specified conditions
- Ensure data authenticity and integrity
- Enable retrieval of data
- Provide information about licensing and permissions (in ideally machine-readable form)
- Ensure confidentiality and respect rights of data subjects and creators

Preservation

- Ensure persistence of metadata and data
- Be transparent about mission, scope, preservation policies, and plans (including governance, financial sustainability, retention period, and continuity plan)

https://www.scienceeurope.org/media/b0bhp2ro/se-rdm-template-4-researcher-guidance-for-the-selection-of-trustworthy-repositories.docx

The Dataverse® Project

Open source research data repository software

Researchers

Enjoy full control over your data. Receive *web visibility, academic credit,* and *increased citation counts.* A personal Dataverse collection is easy to set up, allows you to display your data on your personal website, can be branded uniquely as your research program, makes your data more discoverable to the research community, and satisfies data management plans. Want to set up your personal Dataverse collection?

Journals

Seamlessly manage the submission, review, and publication of data associated with published articles. Establish an *unbreakable link* between *articles in your journal* and *associated data.* Participate in the open data movement by using a Dataverse collection as part of your journal data policy or list of repository recommendations. Want to find out more about journal Dataverse collections?

Institutions

Establish a research data management solution for your commu Federate with a growing list of Dataverse repositories worldwide for increased discoverability c in the drive to set norms for sharing, preserving, citing, explorin to install a Dataverse repository?

Developers

Participate in a vibrant and growing community that is helping preserving, citing, exploring, and analyzing research data. Contr documentation, testing, and/or standards. *Integrate research ar tools,* or other research and data archival systems with the Dat

http://dataverse.org/

PKP-Dataverse Integration Project

Dataverse collaboration with the Public Knowledge Project (PKP)
Contact

IQS
The Institute for Quantitative Social S
HARVARD UNIVERSITY

BLOG DOCUMENTS LINKS PEOPLE PRESENTATIONS

HOME / DOCUMENTS /

Project Documentation

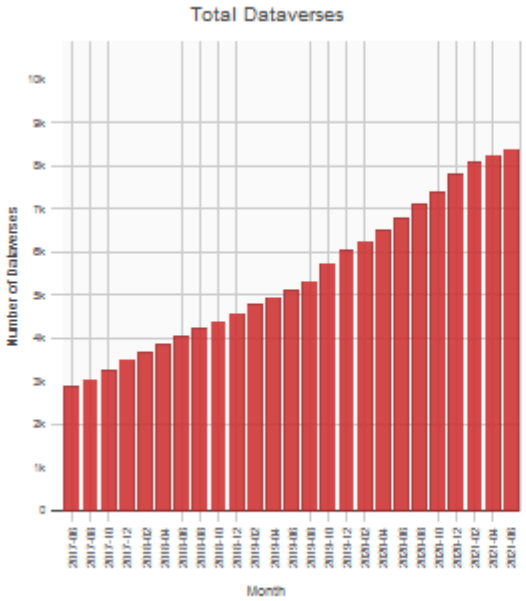
General Project Documentation

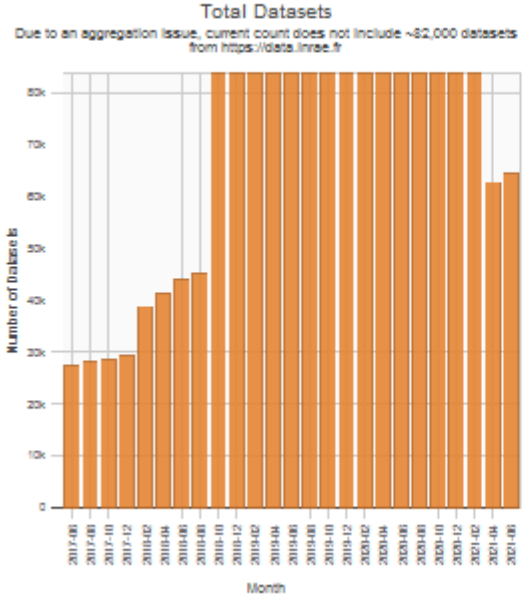
- Project FAQ
- High Level: OJS-Dataverse plugin workflow (visual)
- Plugin Use Cases Workflow (text, more in-depth)
- Video: OJS-Dataverse Plugin Demo
- OJS-Dataverse Plugin Guide: step by step instructions for admins, editors, reviewers, and authors.
- Boilerplate Data Sharing, Citation & Review Policies
- Publications:
  - Article on the project in IJDC (2014 Vol. 9 Issue 1)

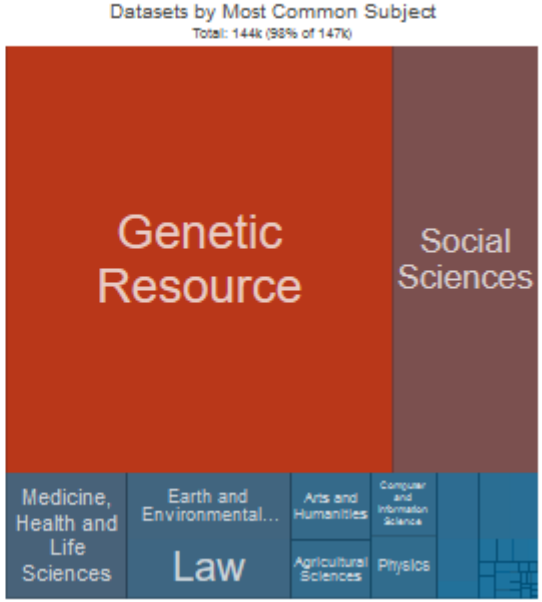
Plugin & API Development Documentation (For Developers)

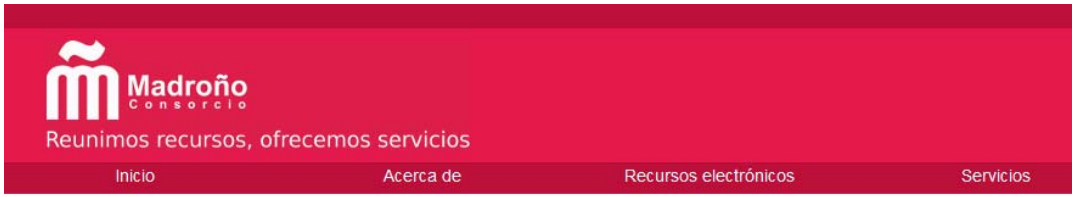
- OJS Dataverse Plugin: on GitHub
  - Plugin setup guide
- Dataverse Network Data Deposit API - (SWORDv2-based)

https://dataversemetrics.odum.unc.edu/

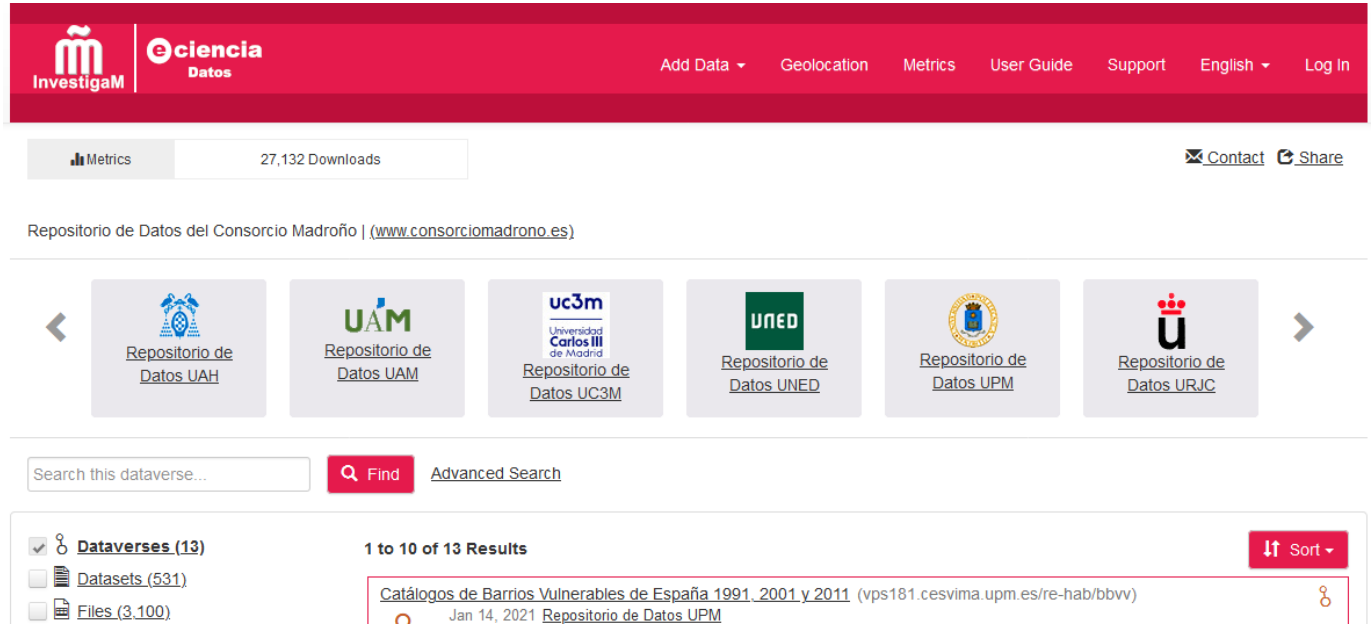




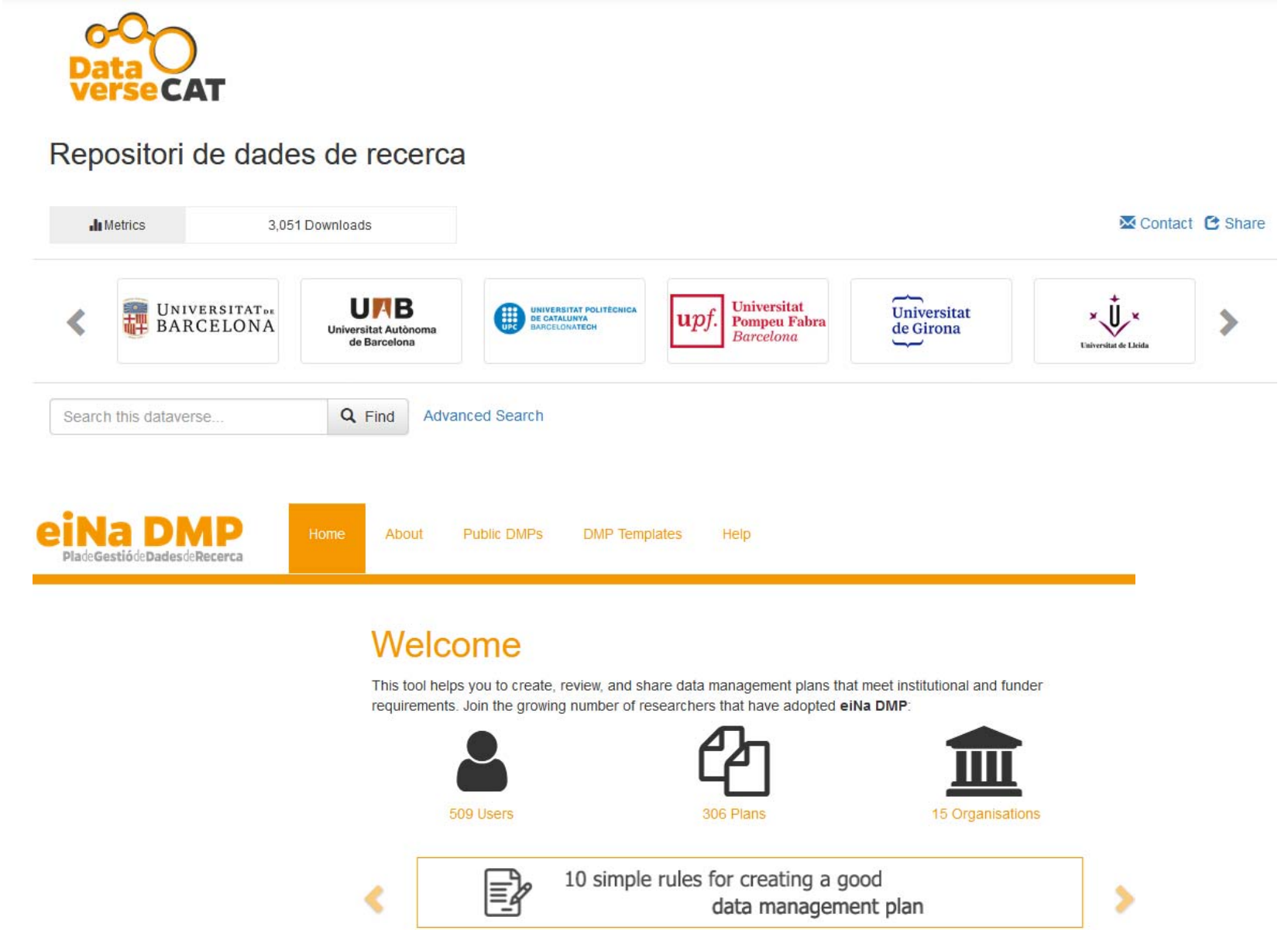




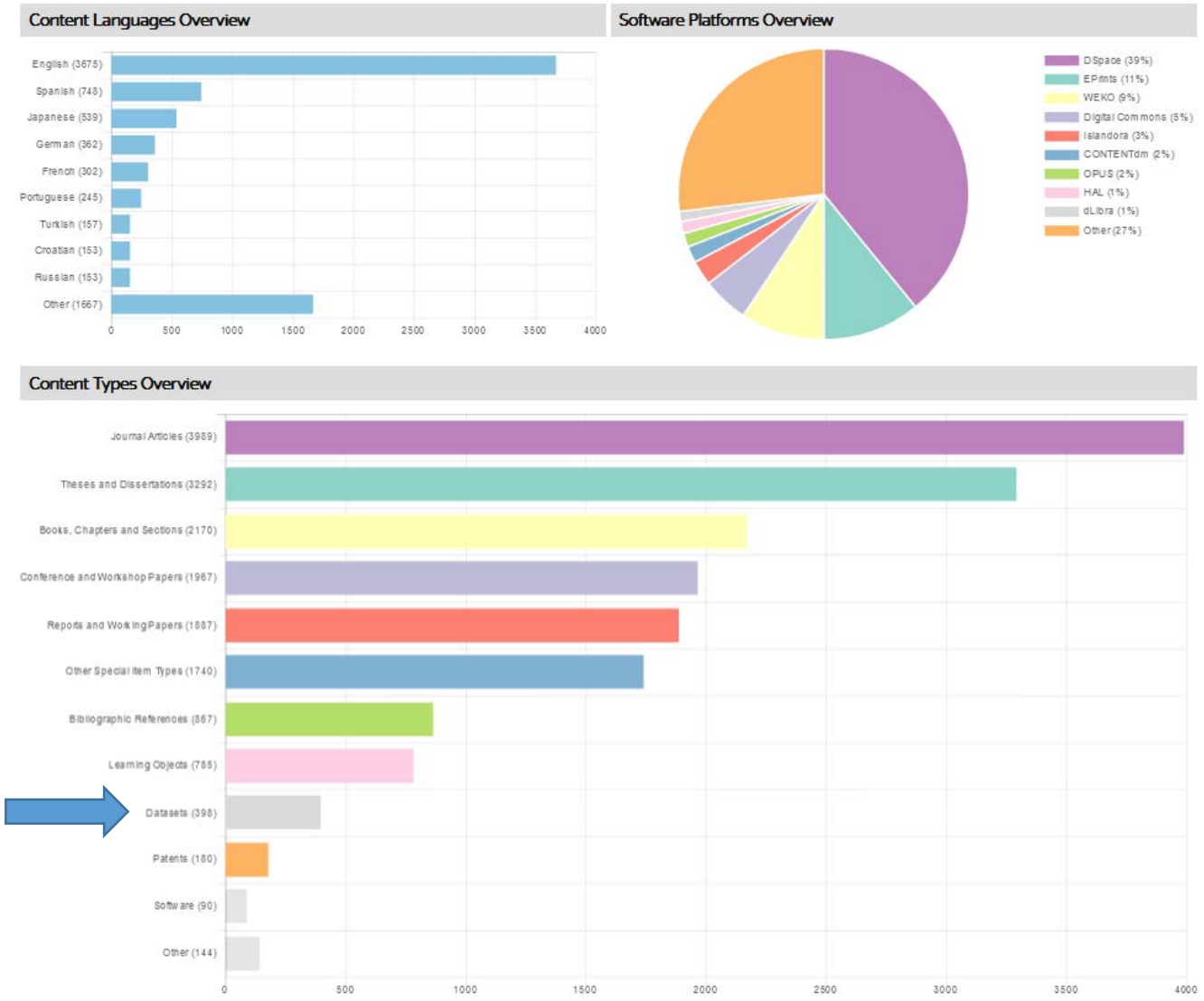
https://edatos.consorciomadrono.es/dataverse/Madrono



https://dataverse.csuc.cat/

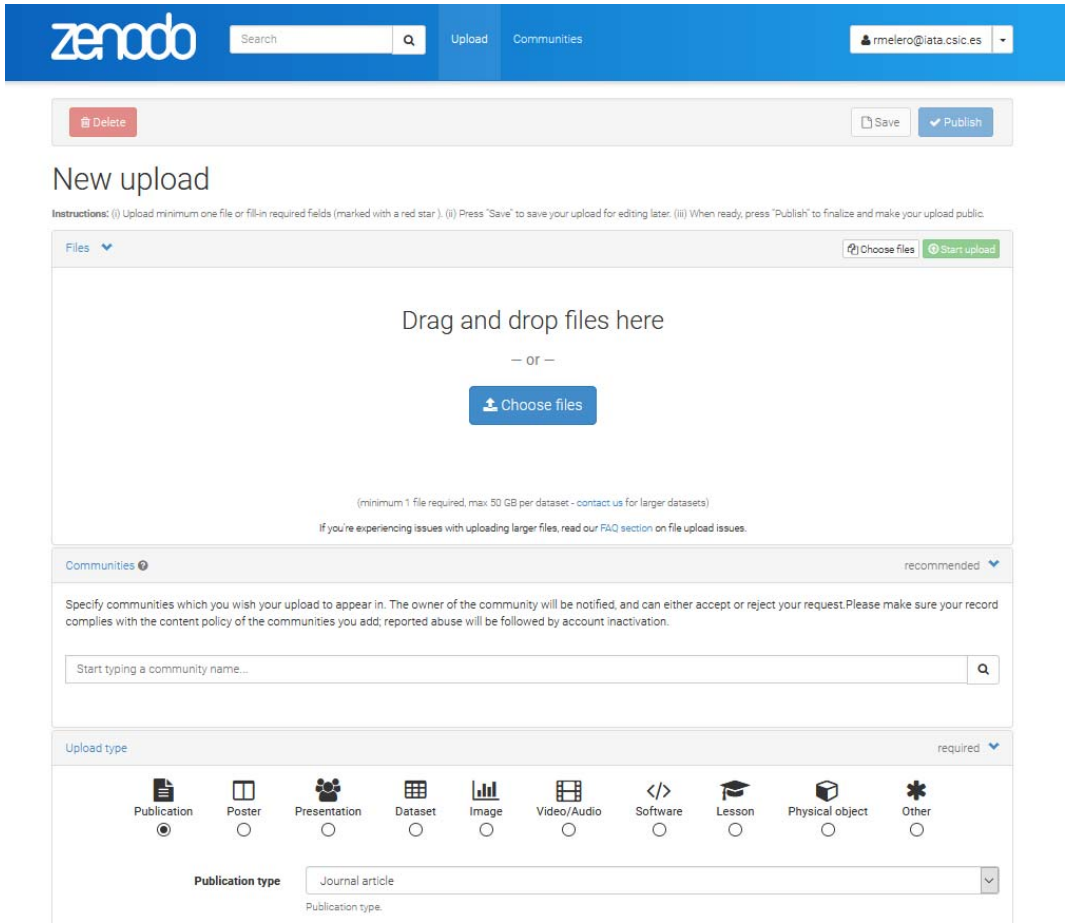


https://v2.sherpa.ac.uk/view/repository_visualisations/1.html

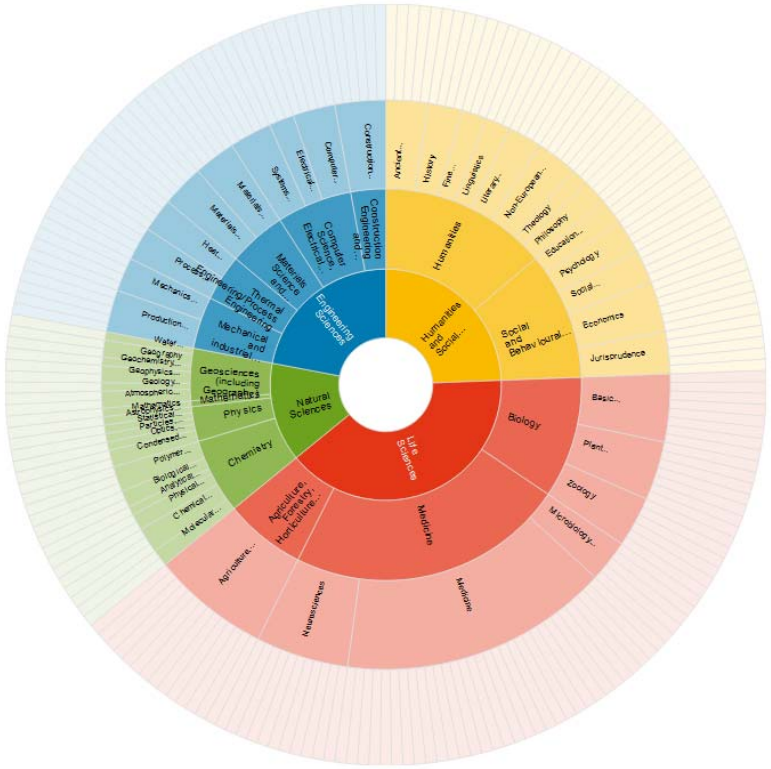


Repositorio “generalista”

https://zenodo.org/deposit/new



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

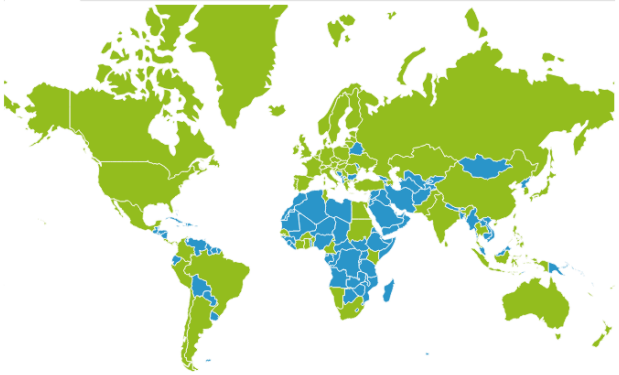
Data sharing made easier: use Repository Finder to find the right repository for your data

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

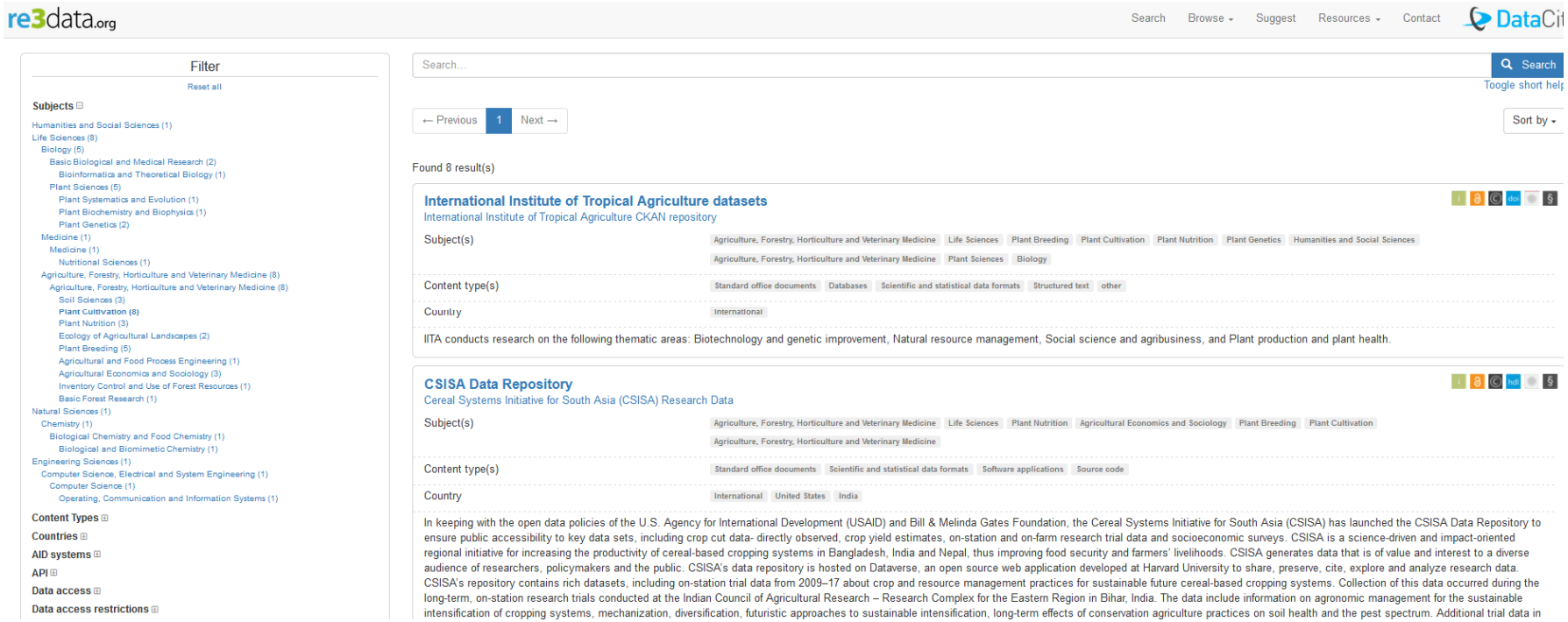
2,000 Data Repositori and Science Europe Framework for Discipline-specific Research Data Management

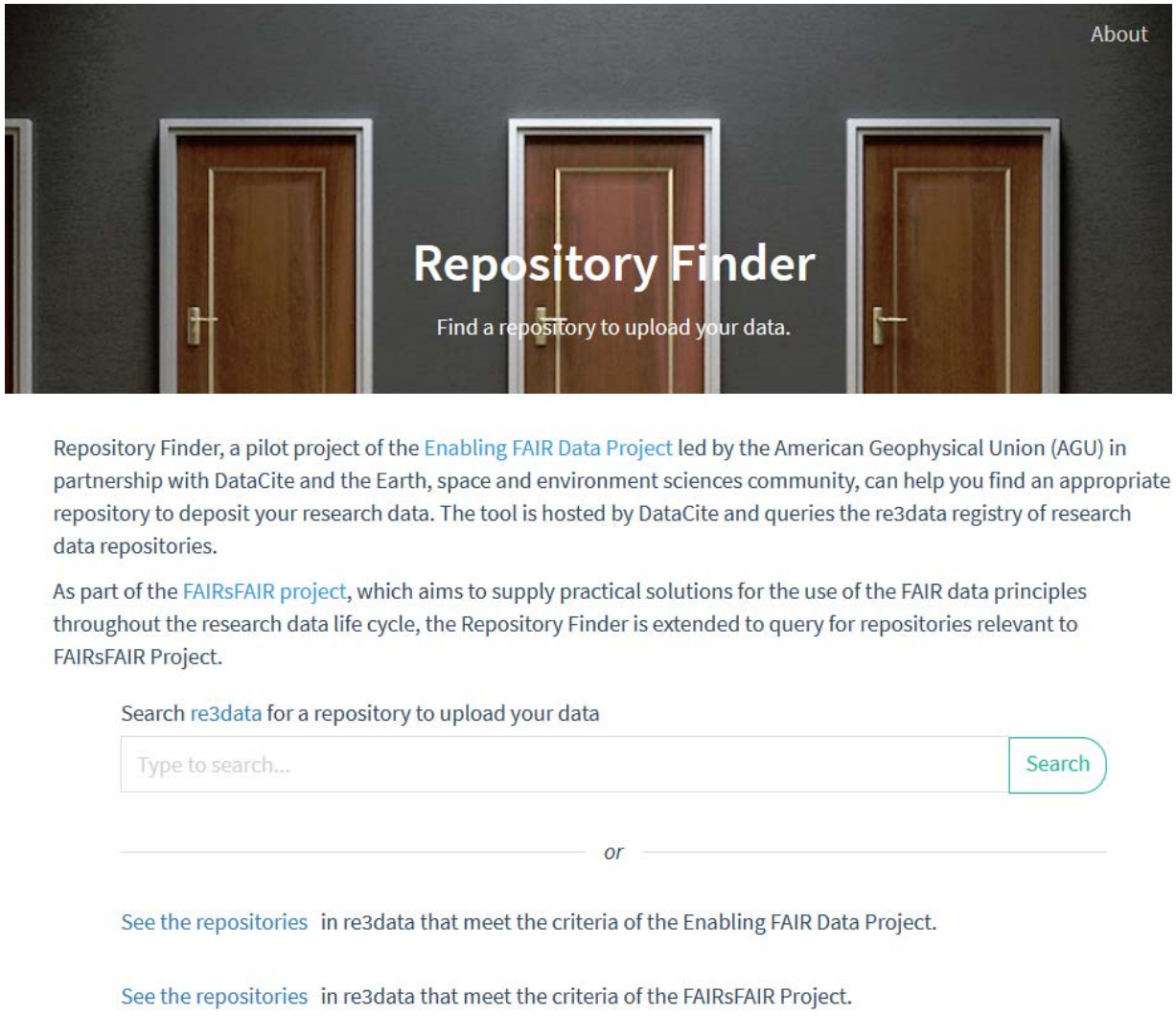
Browse by country

Graphical Text

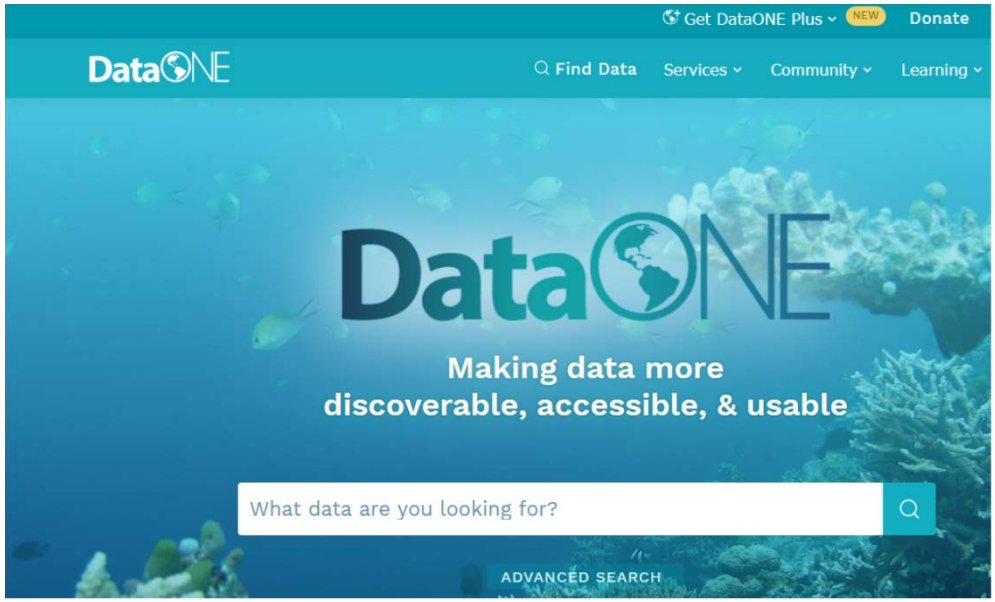


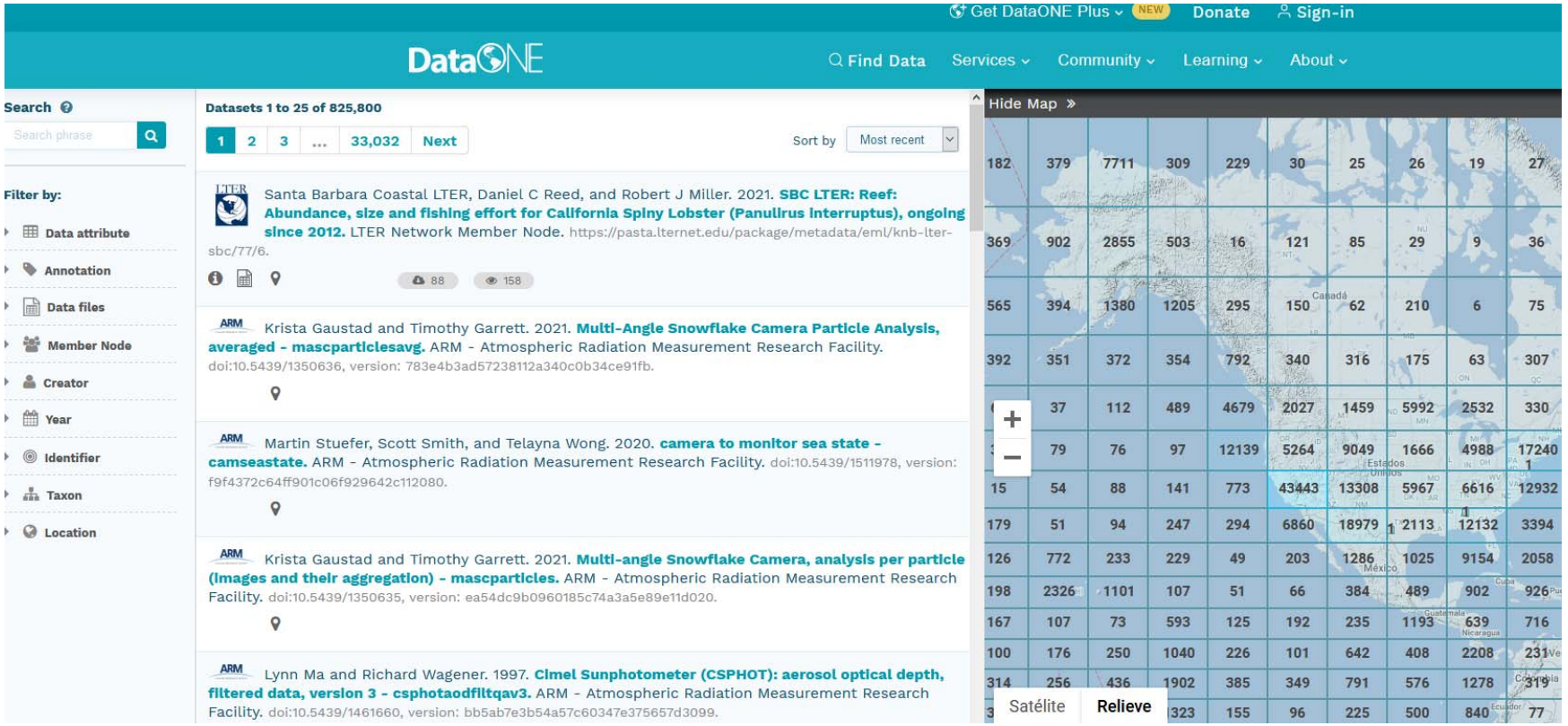
http://www.re3data.org/



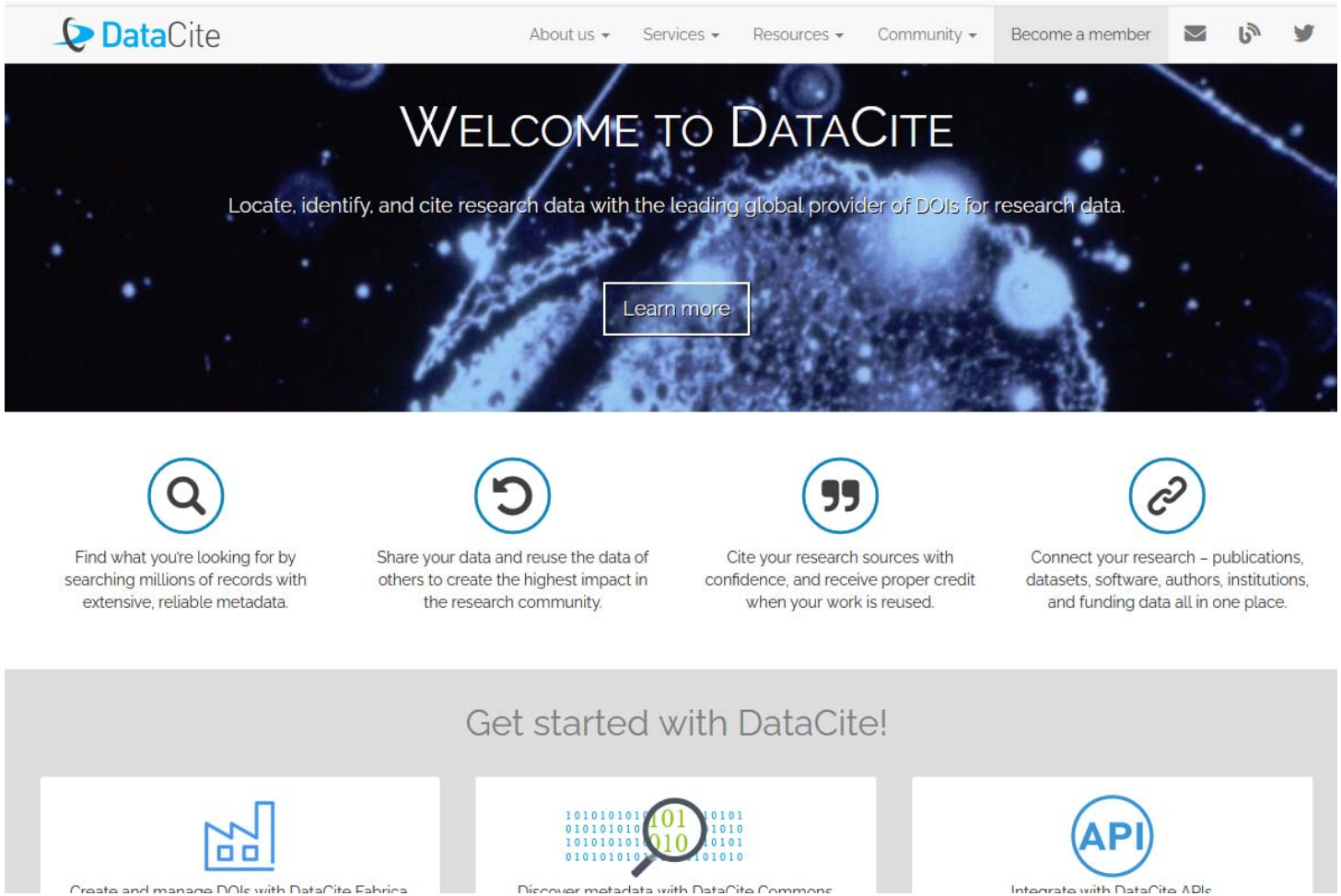


Permite localizar repositorios que siguen los principios FAIR, indexados en el directorio re3data

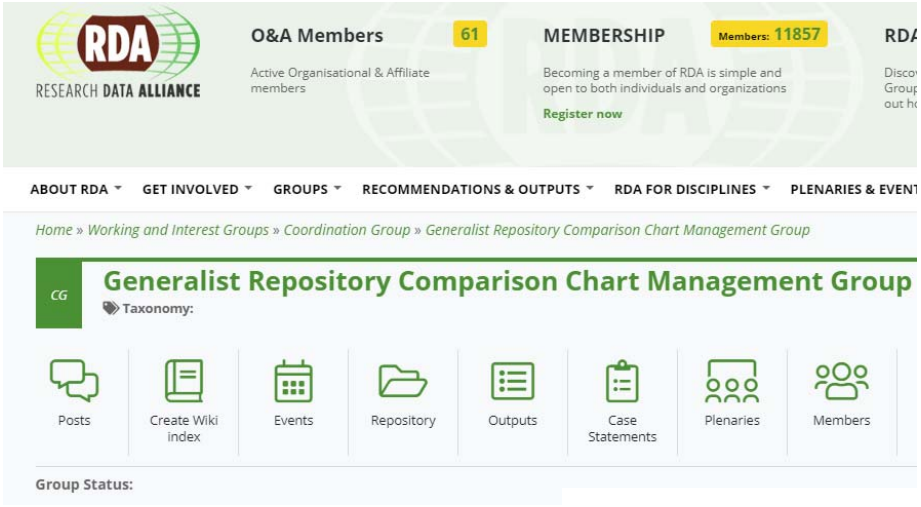

Get DataONE Plus
NEW
Donate
DataONE
Find Data
Services
Community
Learning
DataONE
Making data more
discoverable, accessible, & usable
What data are you looking for?
ADVANCED SEARCH


Get DataONE Plus
NEW
Donate
Sign-in
DataONE
Find Data
Services
Community
Learning
About
Search
Search phrase
Filter by:
Data attribute
Annotation
Data files
Member Node
Creator
Year
Identifier
Taxon
Location
Datasets 1 to 25 of 825,800
1
2
3
...
33,032
Next
Sort by
Most recent
Santa Barbara Coastal LTER, Daniel C Reed, and Robert J Miller. 2021. SBC LTER: Reef: Abundance, size and fishing effort for California Spiny Lobster (Panulirus interruptus), ongoing since 2012. LTER Network Member Node. https://pasta.lternet.edu/package/metadata/eml/knb-lter-sbc/77/6.
88
158
Krista Gaustad and Timothy Garrett. 2021. Multi-Angle Snowflake Camera Particle Analysis, averaged - mascparticlesavg. ARM - Atmospheric Radiation Measurement Research Facility. doi:10.5439/1350636, version: 783e4b3ad57238112a340c0b34ce91fb.
Martin Stuefer, Scott Smith, and Telayna Wong. 2020. camera to monitor sea state - camseastate. ARM - Atmospheric Radiation Measurement Research Facility. doi:10.5439/1511978, version: f9f4372c64ff901c06f929642c112080.
Krista Gaustad and Timothy Garrett. 2021. Multi-angle Snowflake Camera, analysis per particle (images and their aggregation) - mascparticles. ARM - Atmospheric Radiation Measurement Research Facility. doi:10.5439/1350635, version: ea54dc9b0960185c74a3a5e89e11d020.
Lynn Ma and Richard Wagener. 1997. Cimel Sunphotometer (CSPHOT): aerosol optical depth, filtered data, version 3 - csphotaodfiltqav3. ARM - Atmospheric Radiation Measurement Research Facility. doi:10.5439/1461660, version: bb5ab7e3b54a57c60347e375657d3099.
Hide Map
Satélite
Relieve

https://www.datacite.org/



[Generalist Repository Comparison Chart Management Group | RDA (rd-alliance.org)]



Generalist Repository Comparison Chart

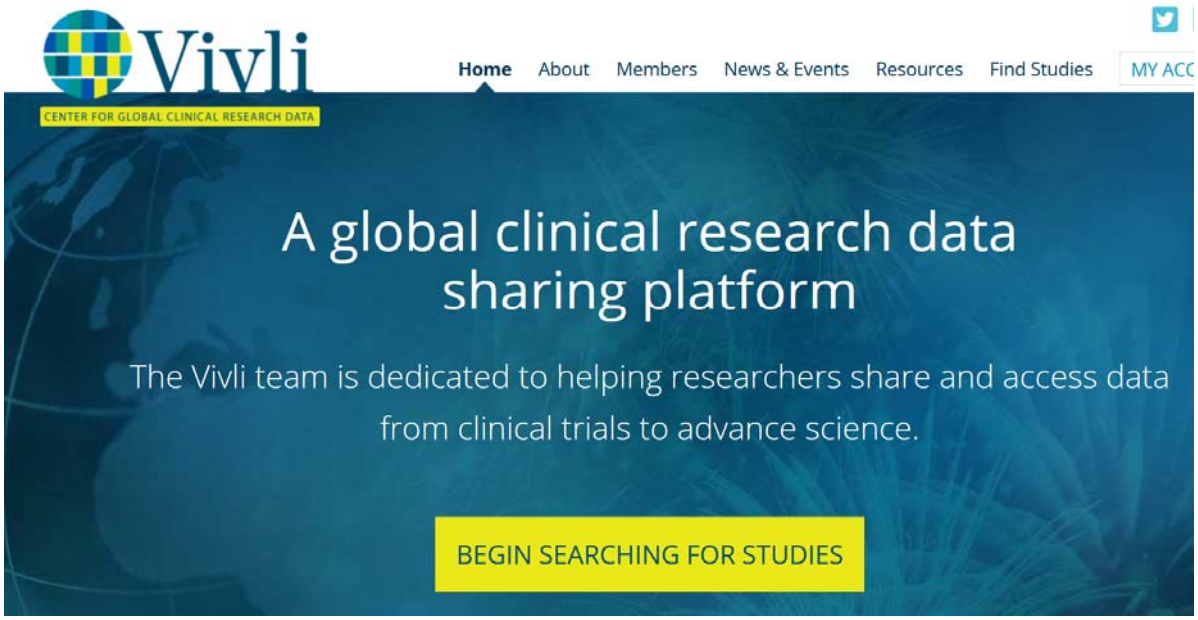
This chart is designed to assist researchers in finding a generalist repository should no domain repository be available to preserve their research data. Generalist repositories accept data regardless of data type, format, content, or disciplinary focus. For this chart, we included a repository available to all researchers specific to clinical trials (Vivli) to bring awareness to those in this field.

doi: 10.5281/zenodo.3946720

https://fairsharing.org/collection/GeneralRepositoryComparison

TOPIC	HARVARD DATAVERSE	DRYAD	FIGSHARE	MENDELEY DATA	OSF	VIVLI	ZENODO
Brief Description	Harvard Dataverse is a free data repository open to all researchers from any discipline, both inside and outside of the Harvard community, where you can share, archive, cite, access, and explore research data.	Open-source, community-led data curation, publishing, and preservation platform for CC0 publicly available research data Dryad is an independent non-profit that works directly with: • researchers to publish datasets utlizing best practices for discovery and reuse • publishers to support the integration of data availability statements and data citations into their workflows • institutions to enable scalable campus support for research data managment best practices at low cost	A free, open access, data repository where users can make all outputs of their research available in a discoverable, reusable, and citable manner. Users can upload files of any type and are able to share diverse research products including datasets, code, multimedia files, workflows, posters, presentations, and more. With discoverable metadata supporting FAIR principles, file visualizations, and integrations, researchers can make their work more impactful and move research further faster.	Mendeley Data is a free repository specialized for research data. Search more than 20+ million datasets indexed from 1000s of data repositories and collect and share datasets with the research community following the FAIR data principles.	OSF is a free and open source project management tool that supports researchers throughout their entire project lifecycle in open science best practices.	Vivli is an independent, non-profit organization that has developed a global data-sharing and analytics platform. Our focus is on sharing individual participant-level data from completed clinical trials to serve the international research community.	Powering Open Science, built on Open Source. Built by reserachers for researchers. Run from the CERN data centre, whose purpose is long term preservation for the High Energy Physics discipline, one of the largest scientific datasets in the world

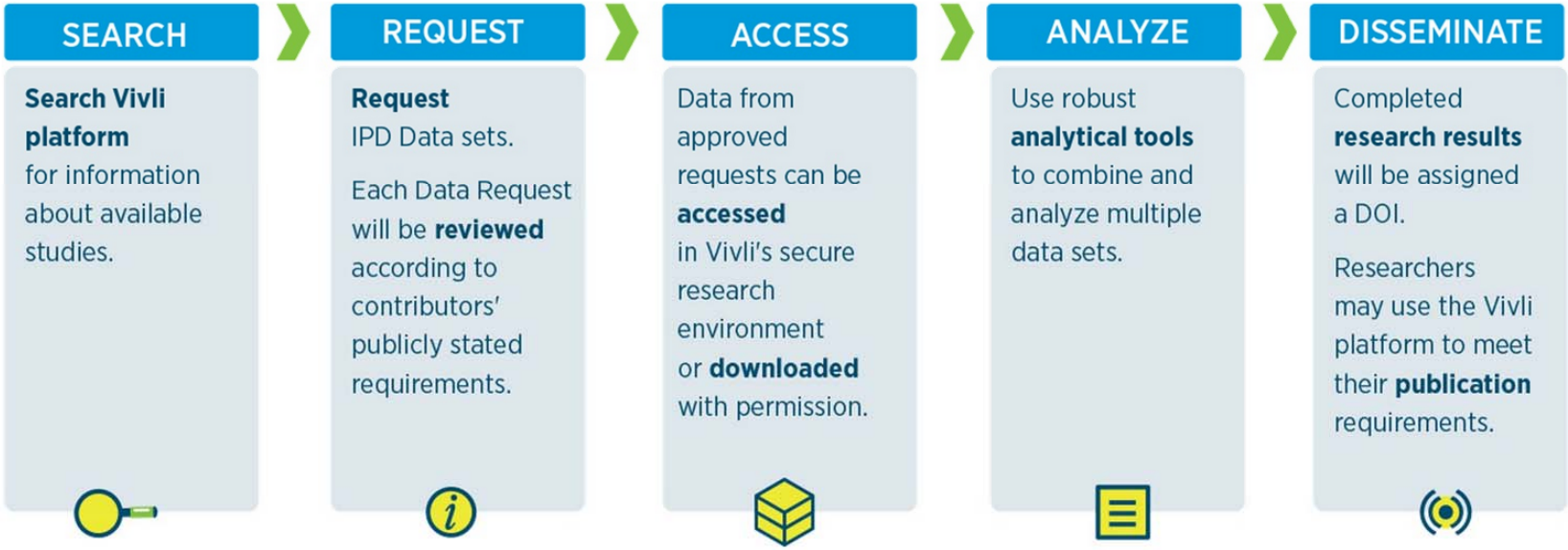
https://www.rd-alliance.org/sites/default/files/Generalist%20Repository%20Comparison%20Chart.pdf

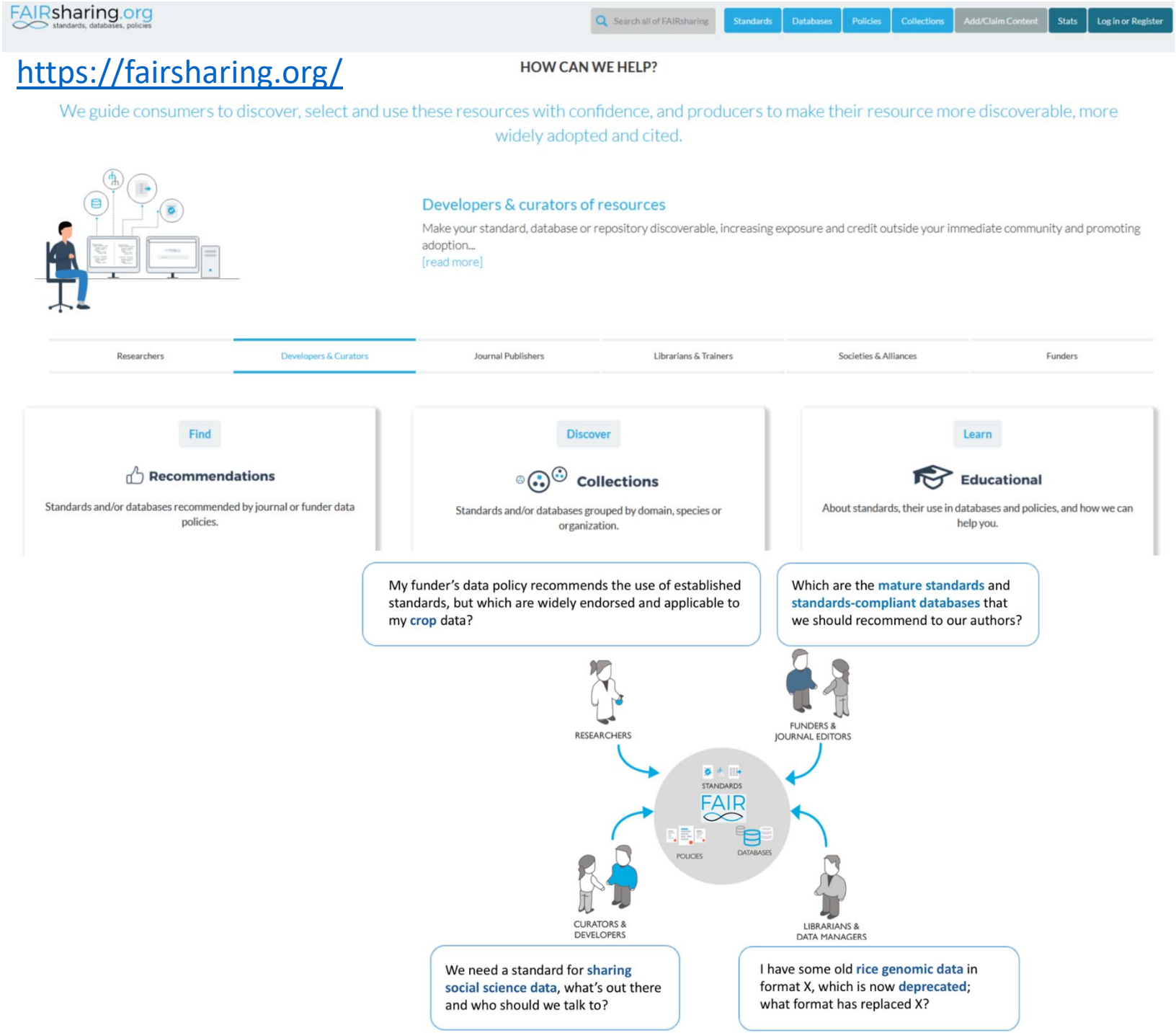


https://vivli.org/

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.




FAIRsharing.org
standards, databases, policies
Search all of FAIRsharing
Standards
Databases
Policies
Collections
Add/Claim Content
Stats
Log in or Register
https://fairsharing.org/
HOW CAN WE HELP?
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.
Developers & curators of resources
Make your standard, database or repository discoverable, increasing exposure and credit outside your immediate community and promoting adoption...
[read more]
Researchers
Developers & Curators
Journal Publishers
Librarians & Trainers
Societies & Alliances
Funders
Find
Recommendations
Standards and/or databases recommended by journal or funder data policies.
Discover
Collections
Standards and/or databases grouped by domain, species or organization.
Learn
Educational
About standards, their use in databases and policies, and how we can help you.
My funder's data policy recommends the use of established standards, but which are widely endorsed and applicable to my crop data?
Which are the mature standards and standards-compliant databases that we should recommend to our authors?
RESEARCHERS
FUNDERS & JOURNAL EDITORS
STANDARDS
FAIR
POLICIES
DATABASES
CURATORS & DEVELOPERS
LIBRARIANS & DATA MANAGERS
We need a standard for sharing social science data, what's out there and who should we talk to?
I have some old rice genomic data in format X, which is now deprecated; what format has replaced X?

Gestión de datos de investigación

La gestión de datos de investigación se refiere a las operaciones intrínsecas al manejo de los datos de investigación durante y después de una actividad de investigación:

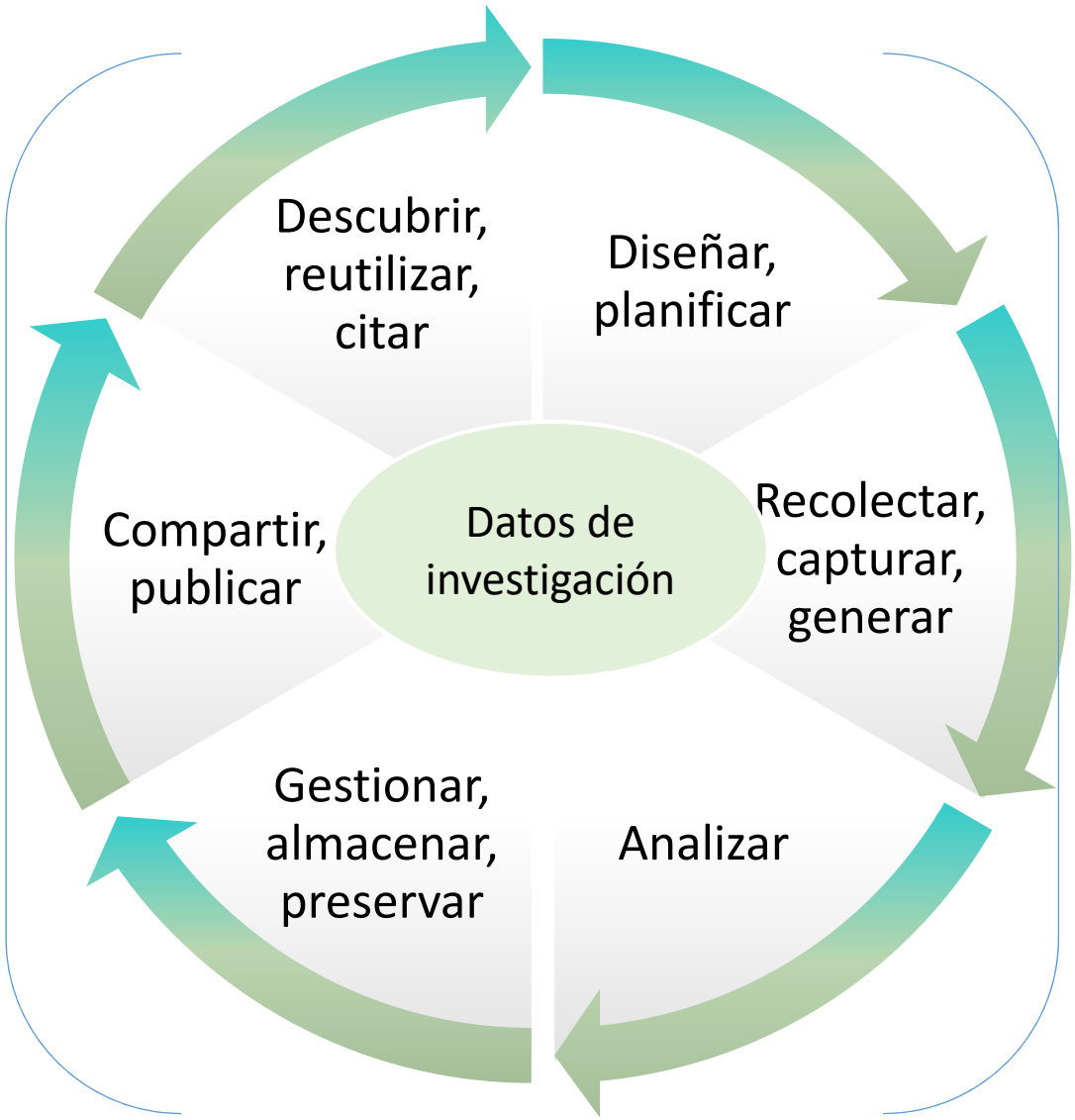
- Recopilación (qué)
- Organización (cómo)
- Almacenamiento (dónde)
- Documentación (cómo)
- Preservación (cómo)
- Puesta en Circulación (cómo, dónde)

Una buena gestión de los datos ayuda a garantizar que los investigadores compartan sus datos de forma FAIR (localizables, accesibles, interoperables y reutilizables).

- Un plan de gestión de datos (PGD) sirve para planificar, organizar y documentar cómo se van a obtener o se han obtenido los datos en el marco de un proyecto de investigación.
- El plan de gestión de datos es un documento vivo que va modificándose o actualizándose en función del desarrollo de una investigación.
- Es recomendable hacer un plan de gestión de datos al inicio del proyecto e ir creando versiones nuevas en función de los posibles cambios que se deriven durante su ejecución.

Ciclo de vida de los datos

Fase II:
"reutilización y preservación"



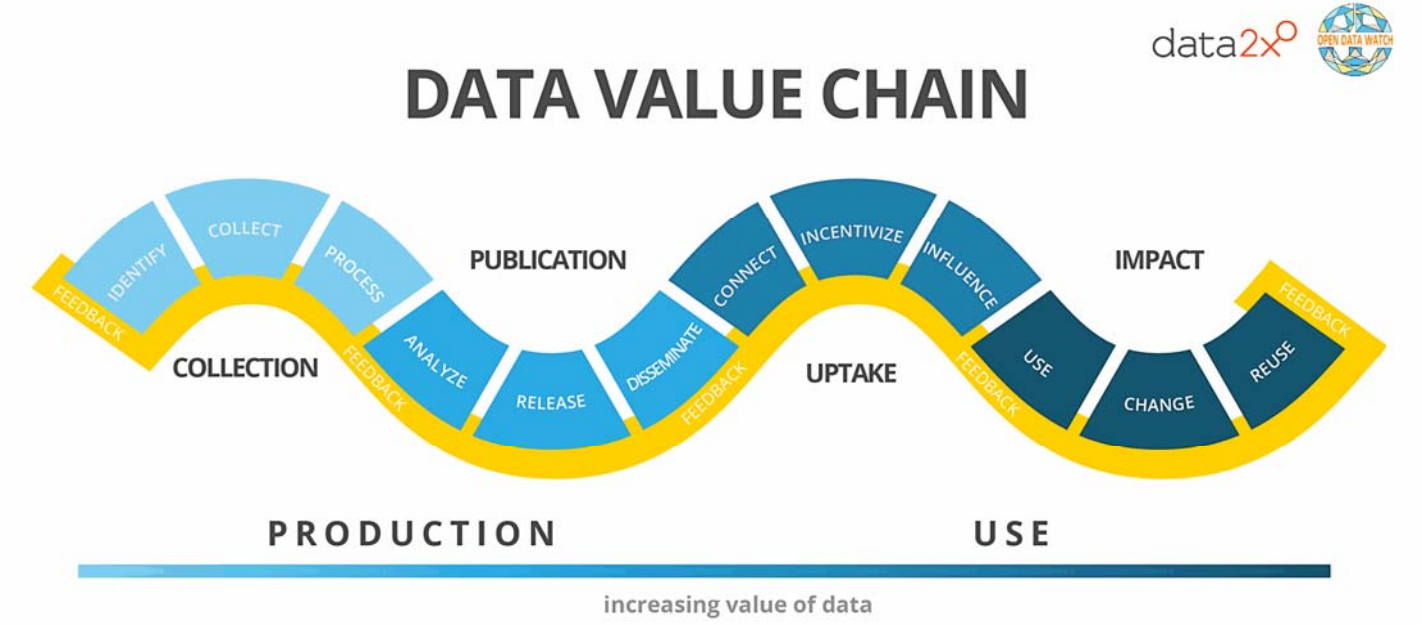
Fase I:
"generación y análisis"

La cadena de valor de los datos describe el proceso de creación y uso de los mismos, desde la identificación de una necesidad de datos hasta su uso final y su posible reutilización.

Consta de cuatro etapas principales:

- Recopilación: identificar, recopilar, procesar
- Publicación: analizar, publicar, difundir
- Absorción/asimilación : conectar, incentivar, influir
- Impacto: utilizar, cambiar y reutilizar

Con una retroalimentación constante entre los productores y las partes interesadas.

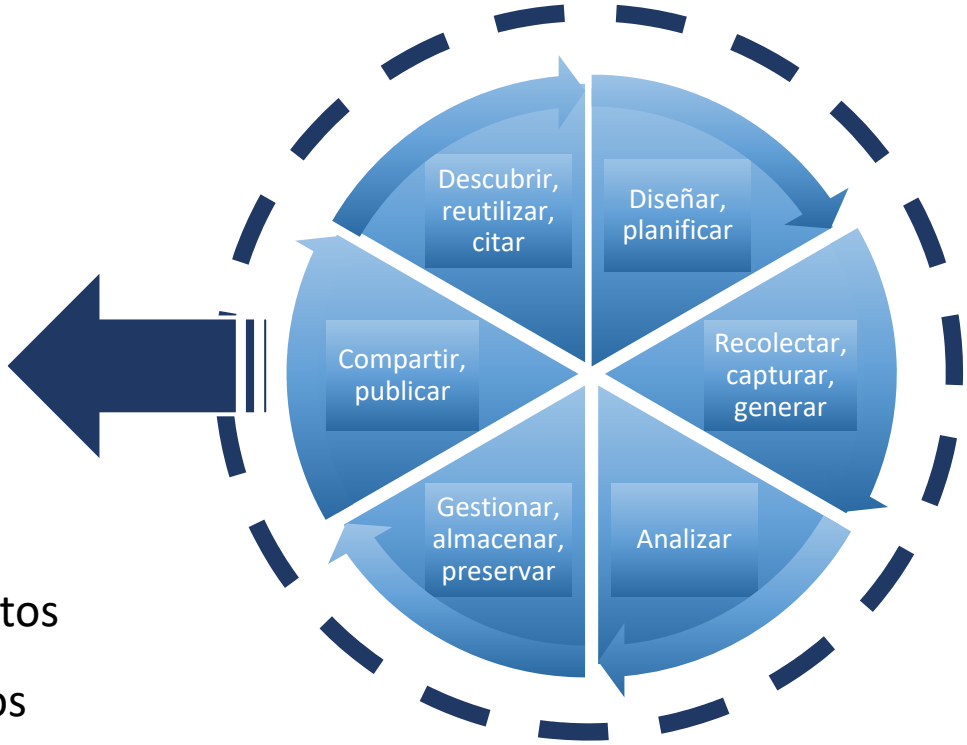


https://opendatawatch.com/wp-content/uploads/2018/03/Data_Value_Chain_OnePager-WR-180312.pdf

¿Qué debe contemplar un plan de gestión de datos?

(mínimos)

- Contexto
- Descripción de los datos que se van a tomar o crear
- La metodología y estándares para la recolección de datos
- Aspectos éticos y relacionados con la propiedad intelectual, si corresponde
- Vías para compartir y acceder a los datos
- Estrategia para la preservación de datos



Las instituciones o agencias financiadoras pueden tener especificaciones propias

Que, porqué, cómo de PGD

https://youtu.be/gYDb-GP1CA4

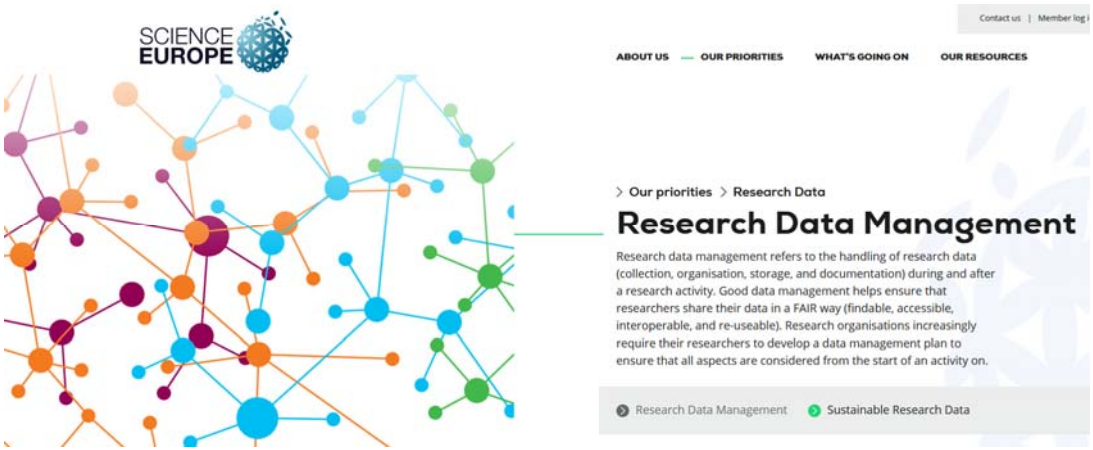
Los componentes comunes de un DMP son:

- Información general sobre el proyecto.
- Descripción de los conjuntos de datos que se utilizarán y generarán.
- Uso de metadatos, ontologías y la forma en que se proporcionará la documentación de los datos.
- Soluciones de almacenamiento, seguridad de los datos y estrategia de conservación durante y después del proyecto.
- Uso compartido de los datos.
- Costes y recursos necesarios para la gestión de los datos.
- Cuestiones éticas y jurídicas, como la privacidad, la propiedad intelectual y las licencias.

Ventajas de elaborar un plan de gestión de datos

- Suele ser un requisito de las organizaciones de investigación y los financiadores.
- Ayuda a planificar y presupuestar los recursos y equipos necesarios.
- Define las funciones y responsabilidades en la gestión de datos entre el equipo del proyecto.
- Ayuda a identificar los riesgos en el manejo de los datos y a aplicar soluciones en una fase temprana.
- Facilita el intercambio, la reutilización y la conservación de los datos.

https://scienceeurope.org/our-resources/practical-guide-to-the-international-alignment-of-research-data-management/



Practical Guide to the International Alignment of Research Data Management

Aspectos básicos a responder en un plan de gestión de datos……

1. Descripción de los datos y recogida o reutilización de los datos existentes.

a. ¿Cómo se recogerán o generarán los nuevos datos y/o cómo se reutilizarán los datos existentes?

b. ¿Qué datos (por ejemplo, el tipo, los formatos y los volúmenes) se recogerán o producirán?

2. Documentación y calidad de los datos

a. ¿Qué metadatos y documentación (por ejemplo, la metodología de recogida de datos y la forma de organizarlos) acompañarán a los datos?

b. ¿Qué medidas de control de calidad de los datos se utilizarán?

3. Almacenamiento y copia de seguridad durante el proceso de investigación

a. ¿Cómo se almacenarán los datos y los metadatos y se harán copias de seguridad durante el proceso de investigación?

b. ¿Cómo se cuidará la seguridad de los datos y la protección de los datos sensibles durante la investigación?

4. Requisitos legales y éticos, códigos de conducta

a. Si se tratan datos personales, ¿cómo se garantizará el cumplimiento de la legislación sobre datos personales y sobre seguridad de los datos?

b. ¿Cómo se gestionarán otras cuestiones legales, como los derechos de propiedad intelectual y la titularidad? ¿Qué legislación es aplicable?

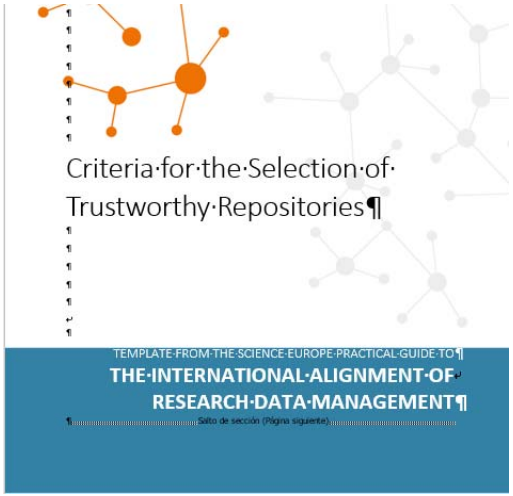
c. ¿Cómo se tendrán en cuenta las posibles cuestiones éticas y se respetarán los códigos de conducta?

5. Intercambio de datos y conservación a largo plazo

a. ¿Cómo y cuándo se compartirán los datos? ¿Existen posibles restricciones para compartir los datos o razones de embargo?
b. ¿Cómo se seleccionarán los datos para su conservación y dónde se conservarán a largo plazo (por ejemplo, en un depósito o archivo de datos)?
c. ¿Qué métodos o herramientas informáticas se necesitarán para acceder a los datos y utilizarlos?
d. ¿Cómo se garantizará la aplicación de un identificador único y persistente (p.e. un DOI) a cada conjunto de datos?

6. Responsabilidades y recursos para la gestión de los datos

a. ¿Quién (por ejemplo, función, cargo e institución) será responsable de la gestión de los datos (es decir, el administrador de los datos)?
b. ¿Qué recursos (por ejemplo, financieros y de tiempo) se dedicarán a la gestión de datos y a garantizar que los datos sean FAIR (Localizables, Accesibles, Interoperables, Reutilizables)?



Provision of Persistent and Unique Identifiers (PID)
Allow data discovery and identification
Enable searching, citing, and retrieval of data
Provide support for data versioning

Metadata
Enable finding of data
Enable referencing to related relevant information, such as other data and publications
Provide information that is publicly available and maintained, even for non-published, protected, retracted, or deleted data
Use metadata standards that are broadly accepted (by the scientific community)
Ensure that metadata are machine-retrievable

Data Access and Usage Licences
Enable access to data under well-specified conditions
Ensure data authenticity and integrity
Enable retrieval of data
Provide information about licensing and permissions (in ideally machine-readable form)
Ensure confidentiality and respect rights of data subjects and creators

Preservation
Ensure persistence of metadata and data
Be transparent about mission, scope, preservation policies, and plans (including governance, financial sustainability, retention period, and continuity plan)


10 pasos para elaborar un Plan de Gestión de Datos
Un **Plan de Gestión de Datos** (PGD) o Data Management Plan (DMP) es un **documento formal, que debe presentarse al inicio de la investigación, en el que se describe qué vas a hacer con tus datos durante y después de finalizar tu investigación** y que puede modificarse si se producen cambios en el proceso de la misma.
¿Por qué?
Es una **buena práctica,** es un **elemento clave de Open Science** y es **obligatorio** en los nuevos proyectos H2020.
Herramientas gratuitas para elaborar un PGD
PGDonline (Consorcio Madroño) http://dmp.consorciomadrono.es/
DMPonline (Digital Curation Centre, UK) https://dmponline.dcc.ac.uk/
01
Revisa los **requerimientos** de la entidad financiadora (H2020).
02
Identifica los datos: tipología, procedencia, volumen, formatos y ficheros.
03
Define cómo se organizarán y gestionarán los datos: nombre de los ficheros, control de versiones, software necesario...
04
Explica cómo se documentarán los datos: identifica la información a procesar, consulta si hay estándares o esquemas de metadatos, identifica herramientas que permitan gestionarlos.
05
Describe los procesos que aseguran una **buena calidad de los datos.**
06
Prepara una estrategia de almacenamiento (durante el proceso) y de preservación de datos (repositorio).
07
Define las políticas de datos del proyecto: cuestiones sobre propiedad intelectual y cómo se tratarán los datos sensibles y personales.
08
Describe cómo se difundirán los datos: dónde, cuáles, cuándo se van a difundir. Si publicarás los datos en un repositorio, como información suplementaria del artículo o como un "data paper".
09
Asigna roles y responsabilidades para las personas y organizaciones participantes en el proyecto.
10
Prepara un presupuesto realista: la gestión de datos cuesta tiempo y dinero en términos de software, hardware, servicios y personal.
crue Universidades Españolas Red de Bibliotecas REBIUN
Adaptado de Michener, WK. Ten simple rules for creating a good data managent plan. PLOS Comp Biol. 2015; 11(10)
El personal de tu Biblioteca te puede asesorar
BY

Personal data and the Open Research Data

How can OpenAIRE help?

Briefing paper for researchers, research administrators and project coordinators

The EC Open Research Data Pilot

Open data is data that is free to use, reuse and redistribute. The EC Open Research Data Pilot (ORDP) enables open access to and reuse of research data generated by Horizon 2020 projects. The Pilot applies primarily to the data (and metadata) needed to validate results in scientific publications, as well as other data specified in the Data Management Plan (DMP).

Projects participating in the Pilot are required to deposit their research data in a research data repository and take measures to enable third parties to access, mine, exploit, reproduce and disseminate this data.

However, the concept of the free use of research data within the Pilot may conflict with data protection rules if such data contain personal data.

What is personal data?

"Personal data" means any information relating to a natural person who is either identified or could be identified by that data (e.g., by reference to an identifier such as a name, an identification number, location data, online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that person).

Data protection rules always apply wherever personal data is being processed. Processing here includes practically any operation in connection with personal data – including collection, recording, organisation, structuring, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction.

OpenAIRE

Dealing with Personal Data

How to balance open access and data protection?

So what's the problem?

Research data – especially in fields like medicine, biotechnology and the social sciences – often contain personal data. This means that many datasets, in their raw form, cannot be made openly available as required by the ORDP due to conflicts with rules on protection of personal data.

Hence, incompatibility with data protection regulations is one of the major reasons for opting out of the ORDP. However, opting out of the Pilot is not the only way to prevent possible infringements.

Firstly, even if one particular dataset is unsuitable for sharing, the same project might produce other datasets which are. Secondly, even if datasets contain personal data, they might still be able to be shared either through (1) anonymisation, or in limited cases, (2) targeted sharing.

Anonymisation of personal data?

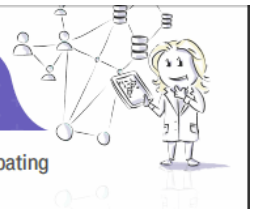
The best way to fulfil the requirements of the **Open Research Data Pilot** and **data protection rules** at the same time is to anonymise personal (research) data before making them openly available.

Anonymised data are no longer personal data, consequently data protection rules are no longer applicable.

Effective **anonymisation** prevents third parties from re-identifying individuals in anonymised datasets, i.e., associating a record to a natural person by using other sources of information. Moreover, anonymisation provides further privacy guarantees that prevent third parties from inferring that a person is associated with a certain property, e.g., a particular health condition, with high probability, or even to infer the participation of a person in a published dataset.

When possible, data anonymisation is the best solution to avert data protection risks.

Open Research Data in Horizon 2020 How can OpenAIRE help?



Are you a Researcher, Project Coordinator or Research Manager participating in a EC project? What, where and how to share your Data?

What is the Open Research Data?

Open data is data that is free to use, reuse and redistribute.
The Open Research Data enables open access and the reuse of research data generated by Horizon 2020 projects.
Open Access to data underlying their scientific publications.

Participating projects are required to:

- Develop (and keep up-to-date) a Data Management Plan (DMP).
- Deposit their data in a research data repository.
- Ensure that third parties can freely access, mine, exploit, reproduce and disseminate it.
- Make clear what tools are needed to use the raw data to validate research results (or provide the tools themselves).

What's in it for you?

- Be part of the new era of Open Science, integrating transparency, effectiveness and timeliness into all areas of scientific methods and processes.
- Reach more people, have greater impact.
- Avoid duplication of effort and help preserve data for future researchers.
- Simplify final reporting by keeping your DMP up-to-date.

OpenAIRE

How can OpenAIRE help? OpenAIRE provides a range of resources, FAQs, webinars and support pages.

OpenAIRE has local representatives in all EU countries: the National Open Access Desks, or NOADs. Contact them via our helpdesk system at

https://www.openaire.eu/support/helpdesk

If there is no disciplinary or institutional repository available, researchers are welcome to use the Zenodo repository provided by OpenAIRE and hosted by CERN.

www.zenodo.org

Do you have a Horizon 2020 grant? Do you manage your data in a FAIR way?

Are you still part of the Pilot?

Projects covered by the Work Programme 2017 are part of the Open Data Pilot by default. As of Work Programme 2017, the Pilot has been extended to all areas of Horizon 2020 (except "co-fund" and "prizes" instruments, ERC PoC, SME instrument Ph1 actions, ERA-NET Cofund actions that do not produce data).
If your project started before then and stems from one of following H2020 areas, you are automatically part of the Pilot as well:

- Future and Emerging Technologies.
- Research infrastructures – part e-Infrastructures.
- Leadership in enabling and industrial technologies – Information and Communication Technologies.
- Nanotechnologies, advanced.
- Societal Challenge: Food security, sustainable agriculture and forestry, marine and maritime and inland water research and the bioeconomy.
- Societal Challenge: Europe in a changing world – inclusive, innovative and reflective Societies.
- Science with and for Society.
- Cross-cutting activities - focus areas – part Smart and Sustainable Cities.

Can you opt out?

We hope you won't, but projects may opt out of the Pilot at any stage, partially or completely. See the EC Guide on OA for eligible reasons.
Note that in multi-beneficiary projects it is also possible for specific beneficiaries to keep their data closed if relevant provisions are made in the consortium agreement, which are in line with the reasons for opting out.

https://www.openaire.eu/factsheets

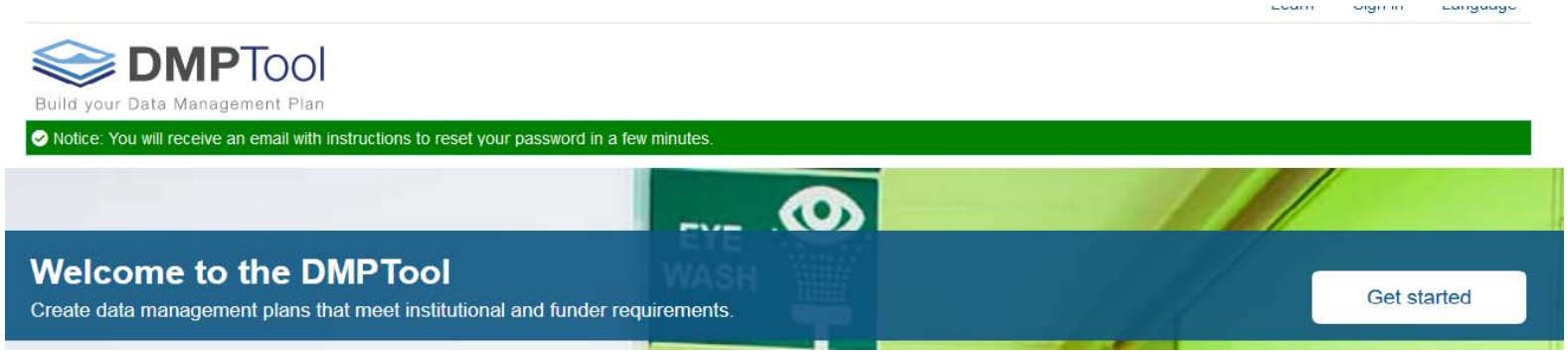
Herramientas para crear un plan de gestión de datos

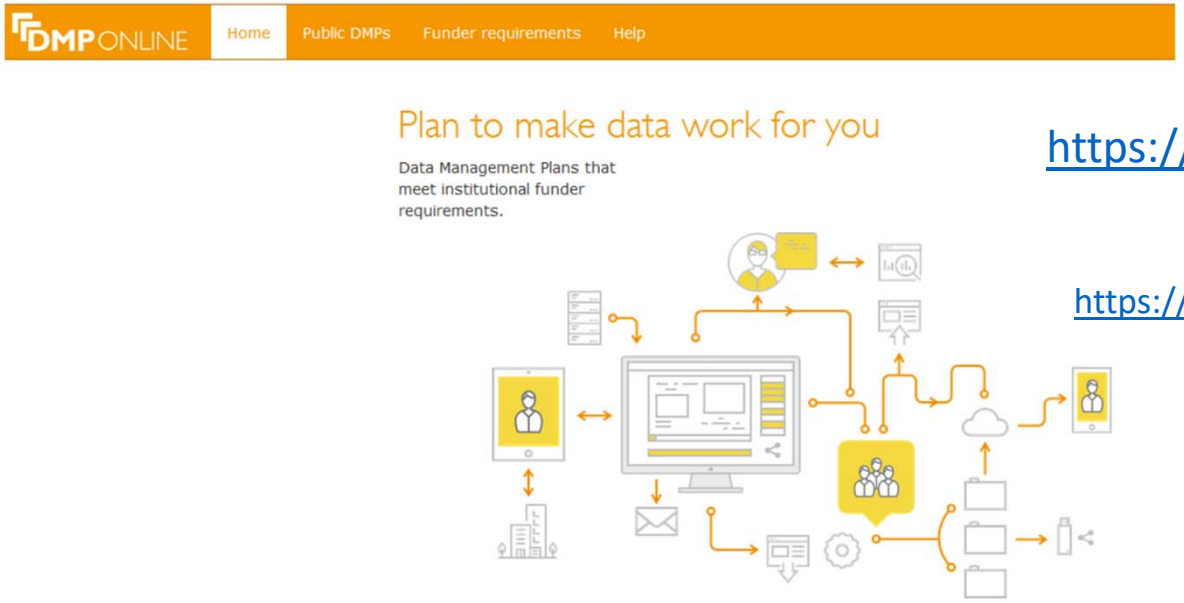
Herramientas para crear planes de gestión de datos

Tool or resource	Description	Tags	Registry
Argos	Plan and follow your data. Bring your Data Management Plans closer to where data are generated, analysed and stored.	DMP researcher data manager	
DMP Canvas Generator	Questionnaire. which generates a pre-filled a DMP	DMP researcher data manager	
DMP OPIDoR	Online questionnaire for the development of data management plans - repository of DMPs	DMP researcher data manager	
DMPlanner	Semi-automatically generated, searchable catalogue of resources that are relevant to data management plans.	DMP researcher data manager	
DMPonline	A free tool to write, share and export a data management plan. Built-in data management plan templates for many major funders.	DMP researcher data manager	
DMPonline Belgium	A free tool to write, share and export a data management plan. Instance aimed at Belgian researchers with built-in data management plan templates for the major funders.	DMP researcher data manager	
DMPTool	Build your Data Management Plan	DMP researcher data manager	
DMPTuuli Finland	Data management planning tool	DMP researcher data manager	
DS-Wizard	Data Stewardship Wizard	DMP researcher data manager IT support	
EasyDMP	DMP creation, versioning and sharing	DMP researcher data manager	
maDMP - Research Bridge	Machine-Actionable Data Management Plan | Webinar (2016) on making a good data management plan.	DMP IT support	
OTP	One Touch Pipeline (OTP) is a data management platform for running bioinformatics pipelines in a high-throughput setting, and for organising the resulting data and metadata.	human data metadata DMP data analysis	
Research Management Plan	Machine actionable DMPs.	DMP researcher data manager	

https://rdmkit.elixir-europe.org/data_management_plan.html#relevant-tools-and-resources

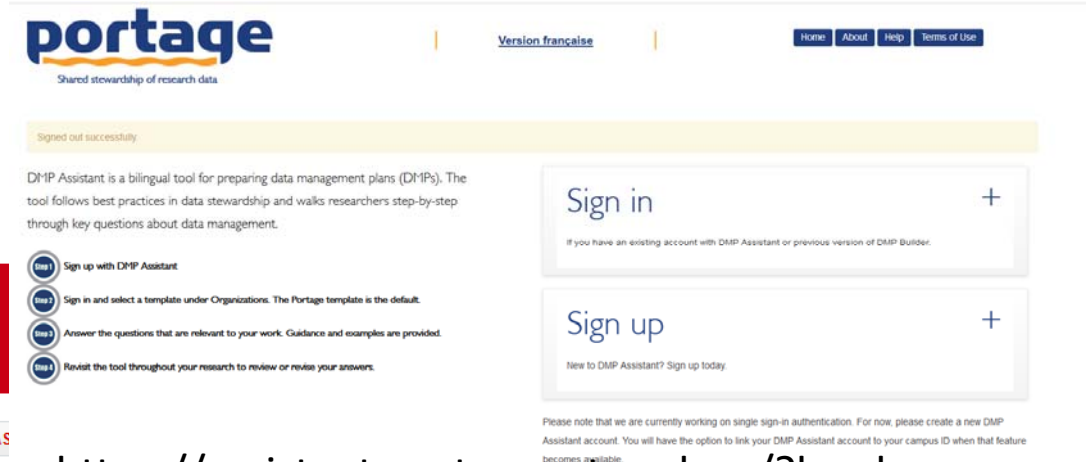
https://dmptool.org/





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

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



https://assistant.portagenetwork.ca/?locale=en

PAGODA

PGDonline

INICIO | CREAR UN PGD | PGDonline | DOCUMENTOS | FAQs | AS

PA GO DA - PlAn de GestiÓn de DAtos

Crear su Plan de Gestión de Datos

El Plan de Gestión de Datos lo solicita un agente financiador como parte de las condiciones del contrato de subvención para un proyecto científico.

El Programa Horizonte 2020 requiere que los proyectos que formen parte del Piloto de Datos de Investigación en Abierto entreguen un Plan de Gestión de Datos completo durante los 6 primeros meses del proyecto.

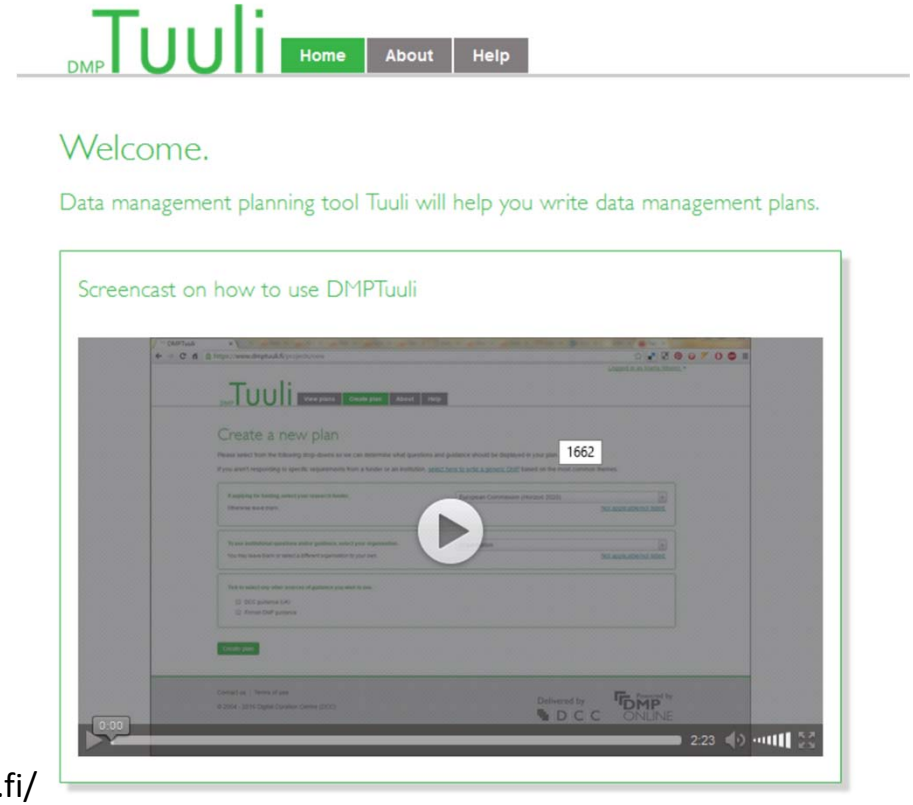
Los planes de gestión de datos son una parte integral de las solicitudes de subvenciones - no pueden ser una idea de último momento; los revisores buscarán evidencia de que la gestión de datos está incluida en su propuesta, y que forma parte integral de su proceso de investigación. En el artículo 29.3 del H2020 Model Grant Agreement: Multi-beneficiary General MGA: December 2013 se establecen las obligaciones de los participantes en el Piloto de Datos de Investigación en Abierto en lo que respecta a la gestión de los datos.

El documento Directrices sobre la Gestión de los Datos en Horizonte 2020 se dirige a los solicitantes y beneficiarios de los proyectos en el Marco del Piloto de Datos de Investigación en Abierto y su objetivo es proporcionar indicaciones sobre cómo pueden cumplir con sus responsabilidades con respecto a la calidad de los datos de investigación, su intercambio y su seguridad.

PGDonline-madroño

http://pgd.consorciomadrono.es/

https://dmp.csuc.cat/



https://www.dmptuuli.fi/

https://pgd.consorciomadrono.es/

Inicio | Acerca de | Recursos electrónicos | Servicios | InvestigaM – Ciencia Ab

> InvestigaM – Ciencia Abierta

INVESTIGAM -Portal de Ciencia Abierta

InvestigaM, el portal de ciencia abierta del Consorcio Madroño, tiene por finalidad la difusión de los resultados de la actividad investigadora que se lleva a cabo en sus instituciones miembro y asociadas, lo que incluye tanto publicaciones como datos de investigación. Se inscribe en el compromiso del Consorcio Madroño y sus universidades miembro con el Acceso Abierto y la Ciencia Abierta (OpenScience).

e-cienciaDatos

e-cienciaDatos es un repositorio de datos multidisciplinar que alberga los conjuntos de datos científicos de l investigadores de las universidades públicas de la Comunidad de Madrid y la UNED, miembros del Consorc Madroño, con el fin de dar visibilidad a dichos datos, garantizar su preservación y facilitar su acceso y reutilización.

Medición del Acceso Abierto en el Consorcio Madroño (2015-2019)

Estimación de costes de publicación por pago de APC

Declaración Ciencia Abierta

Aumento de descargas en repositorios durante el confinamiento (2020)

Publicaciones científicas

Buscador de tesis

¿Cómo depositar publicaciones?

ESPAÑOL | ENGLISH

Madroño Consorcio

Reunimos recursos, ofrecemos servicios

Inicio | Acerca de | Recursos electrónicos | Servicios | InvestigaM – Ciencia Abierta

> InvestigaM – Ciencia Abierta > PaGoDa

INVESTIGAM: PAGODA

Un Plan de Gestión de Datos es un documento que describe el tratamiento que van a recibir los datos de investigación recopilados o generados en el curso de un proyecto de investigación. El Consorcio Madroño ofrece la herramienta PGDonline, adaptación y traducción al español de la herramienta de gestión DMPonline desarrollada por el Digital Curation Centre del Reino Unido, así como también las directrices para la Gestión de Datos en Horizonte 2020.

PaGoDa además le ofrece material de apoyo y un servicio de asesoramiento en cada una de las bibliotecas miembros.

El Programa Marco de Investigación e Innovación de la Unión Europea Horizonte 2020 incluye un Piloto de Datos de Investigación en Acceso Abierto cuyo objetivo es garantizar el acceso y la reutilización de los datos generados en el ámbito de proyectos participantes.

En la Comunicación de la Comisión "Iniciativa Europea de Computación en la Nube: construir en Europa una economía competitiva de los datos y del conocimiento" (19 abril 2016) se indica que a **partir de 2017 "La Comisión convertirá la apertura de los datos de la investigación en la opción por defecto, previendo no obstante opciones de exclusión, para todos los proyectos nuevos del programa Horizonte 2020".**

El Programa Horizonte 2020 requiere que los proyectos que formen parte del Piloto de Datos de Investigación en Abierto entreguen un Plan de Gestión de Datos completo durante los 6 primeros meses del proyecto.

Crear un Plan de Gestión de Datos

argos

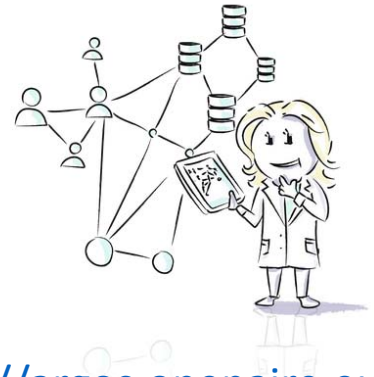
ABOUT RESOURCES CONTACT LOG IN

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



https://argos.openaire.eu/splash/

Adding dataset
(unsaved changes)
To DMP: **prueba en argos**

Descartar | Grabar | Save & Close | Save & Add New

Back to PGD

Guía paso a paso

0. Información principal (5)
1 Data Summary
2 Reusable Data
3 FAIR Data
4 Allocation of resources
5 Data Security
6 Ethical aspects
7 Other

1.1 Title of Dataset*

Title of Dataset

1.2 Descripción

A brief description of what the PGD is about it's scope and objectives.

Rellenar con la descripción

1.3 Etiquetas

Etiquetas

1.4 Plantilla*

Horizon 2020

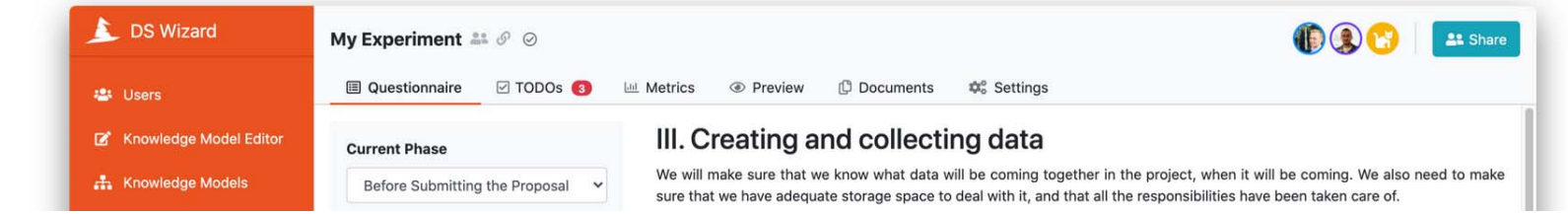
Anterior | Siguiente

DSW

About Features Resources Get Started

Create Smart Data Management Plans for FAIR Open Science

Watch video



Get Started Using the Data Stewardship Wizard

Demo

For exploring the DSW features

- Easy to sign up and use
- A shared instance with other users
- **Not for serious usage**

✓ Questionnaires & DMPs
✓ Knowledege Models Management

Researchers

For individual researchers

- Easy to sign up and use
- Ready to use Knowledge Models
- Privacy and stability

✓ Questionnaires & DMPs

Self-Managed

For organizations

- All the DSW features available
- Your own instance
- You need to host and run the instance by yourself

✓ Questionnaires & DMPs
✓ Knowledege Models Management
✓ User Management & Organization Settings

DSW Cloud

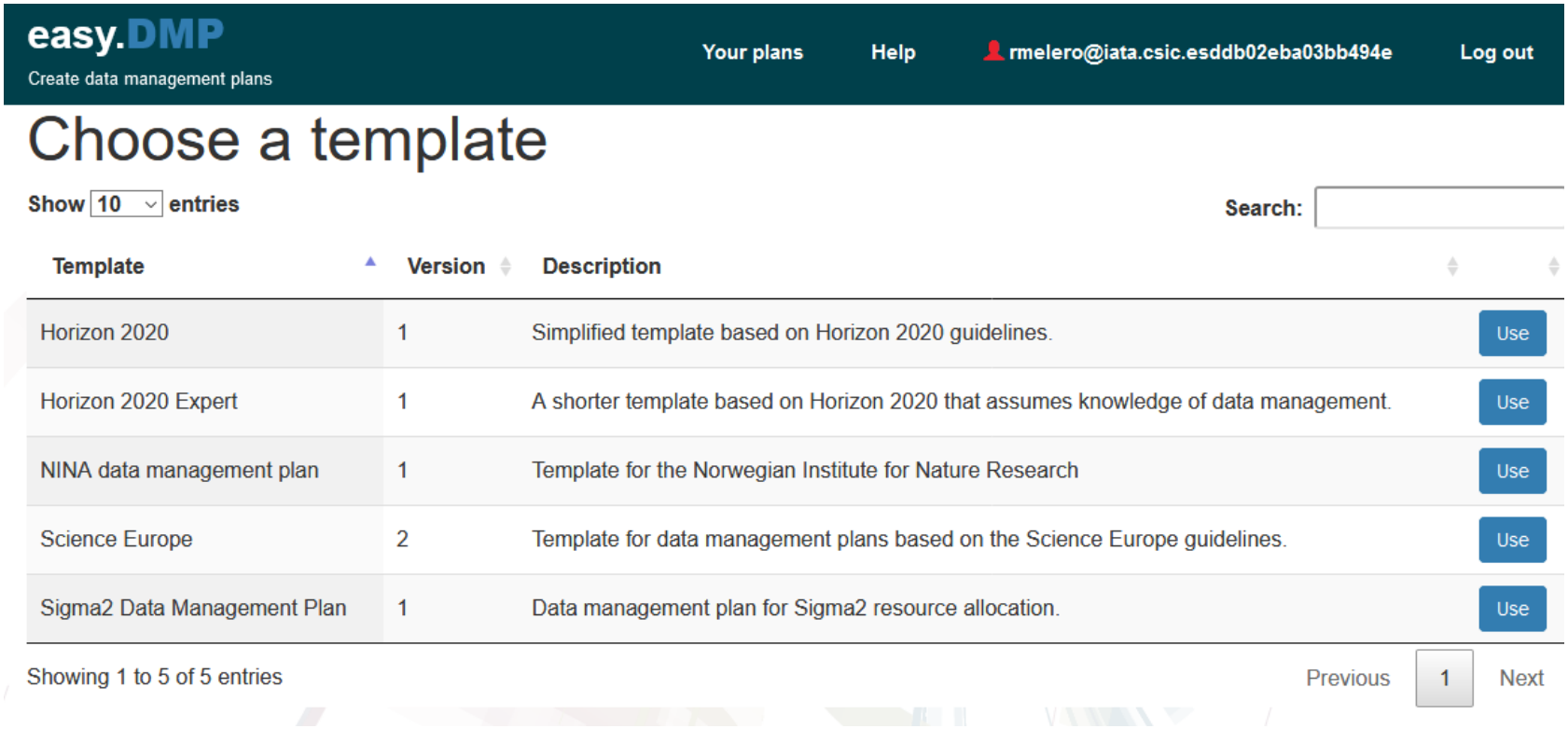
For organizations

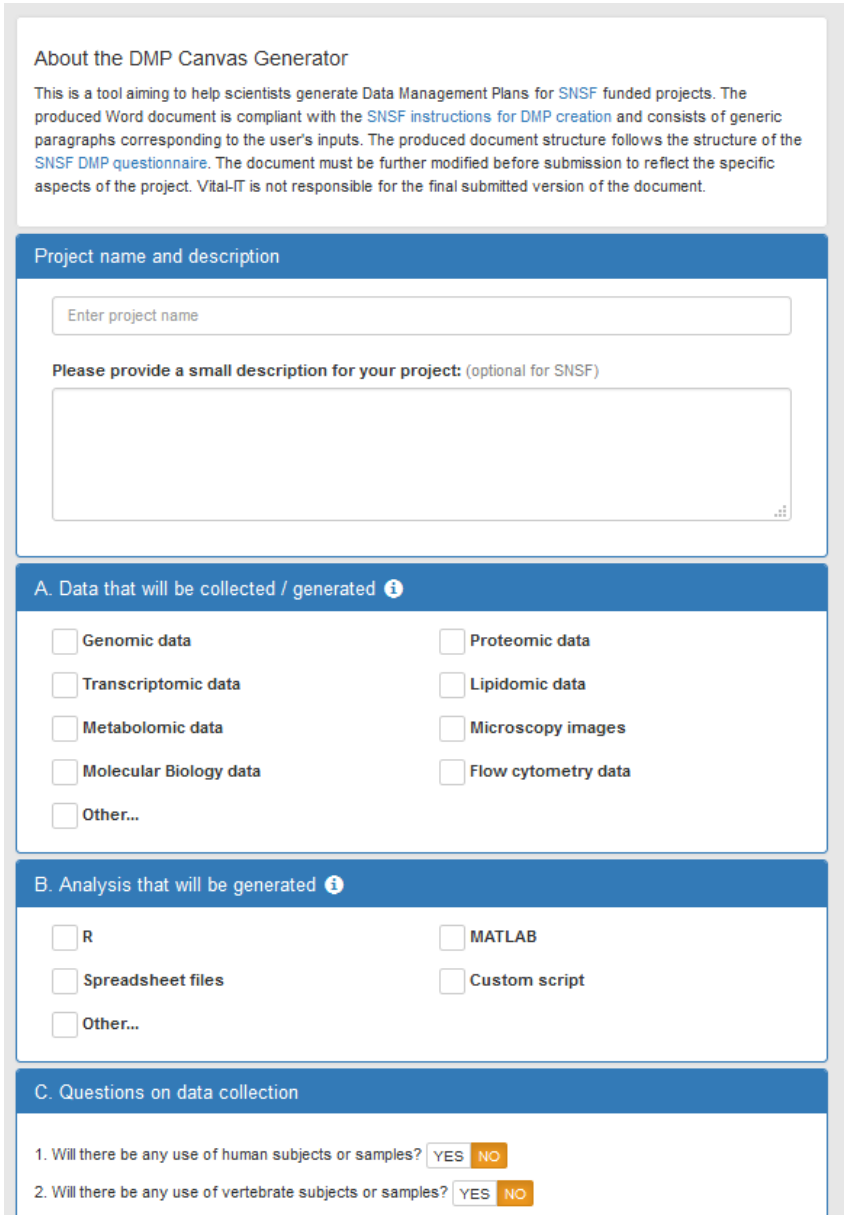
We offer managing the DS Wizarc instance for interesting project: that want to use it seriously bu don't want to run it by themselves.

✓ Questionnaires & DMPs
✓ Knowledege Models Management
✓ User Management & Organization Settings

https://ds-wizard.org/

Acceso por EUDAT, hay que identificarse/registrarse





https://dmp.vital-it.ch/#/form

https://rdmkit.elixir-europe.org/data_protection.html

RDMkit About Contribute GitHub Give feedback Search RDMkit

- Data life cycle +
- Your role +
- Your domain +
- Your problem +
- All tools and resources
- Tools assembly +

Are you working with data in the Life Sciences? Do you feel overwhelmed when you think about Research Data Management?

The ELIXIR Research Data Management Kit (RDMkit) is an online guide containing good data management practices applicable to research projects from the beginning to the end. Developed and managed by people who work every day with life science data, the RDMkit has guidelines, information, and pointers to help you with problems throughout the data's life cycle. RDMkit supports FAIR data — Findable, Accessible, Interoperable and Reusable — by-design, from the first steps of data management planning to the final steps of depositing data in public archives.

The RDMkit organises information into the six sections displayed below, which are interconnected but can be browsed independently.

Data life cycle

Start here to get an overview of research data management. Click on a section of the diagram below to get an introduction to that stage of the data management life cycle.

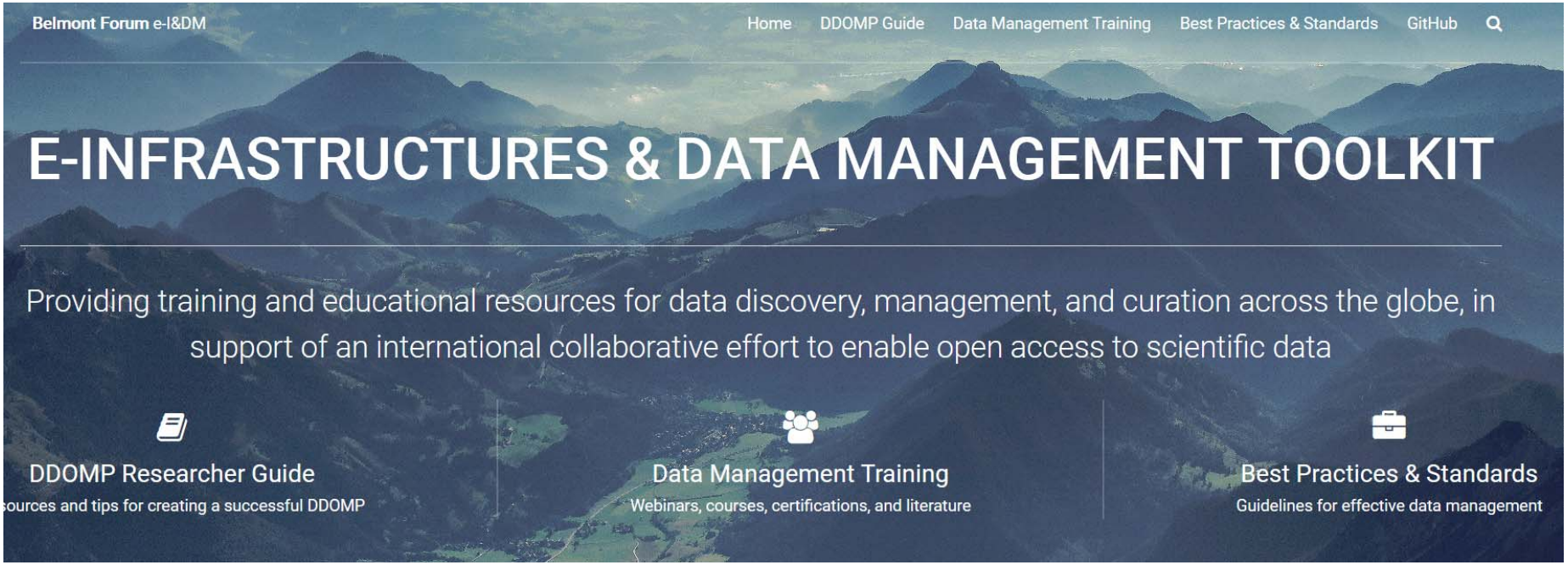
Your role

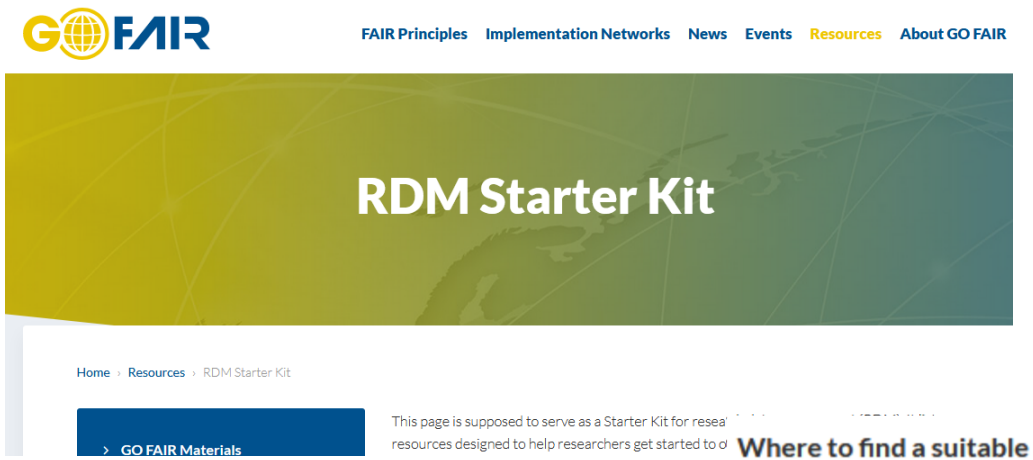
Identify your role in research data management, find data

Your domain

Learn about the data management problems that affect

https://bfe-inf.github.io/toolkit/index.html





https://www.go-fair.org/resources/rdm-starter-kit/

Where to find a suitable research data repository and RDM support

- FAIRsharing: searchable curated registry of databases, repositories, (inter-related to) data/metadata standards, and data policies by journals/publishers and funders (International)
- RatSWD: 31 data centers accredited by the German Data forum according to uniform and transparent standards (Germany)
- re3data.org: Registry of Research Data Repositories with detailed information about over 2000 research data repositories (International)

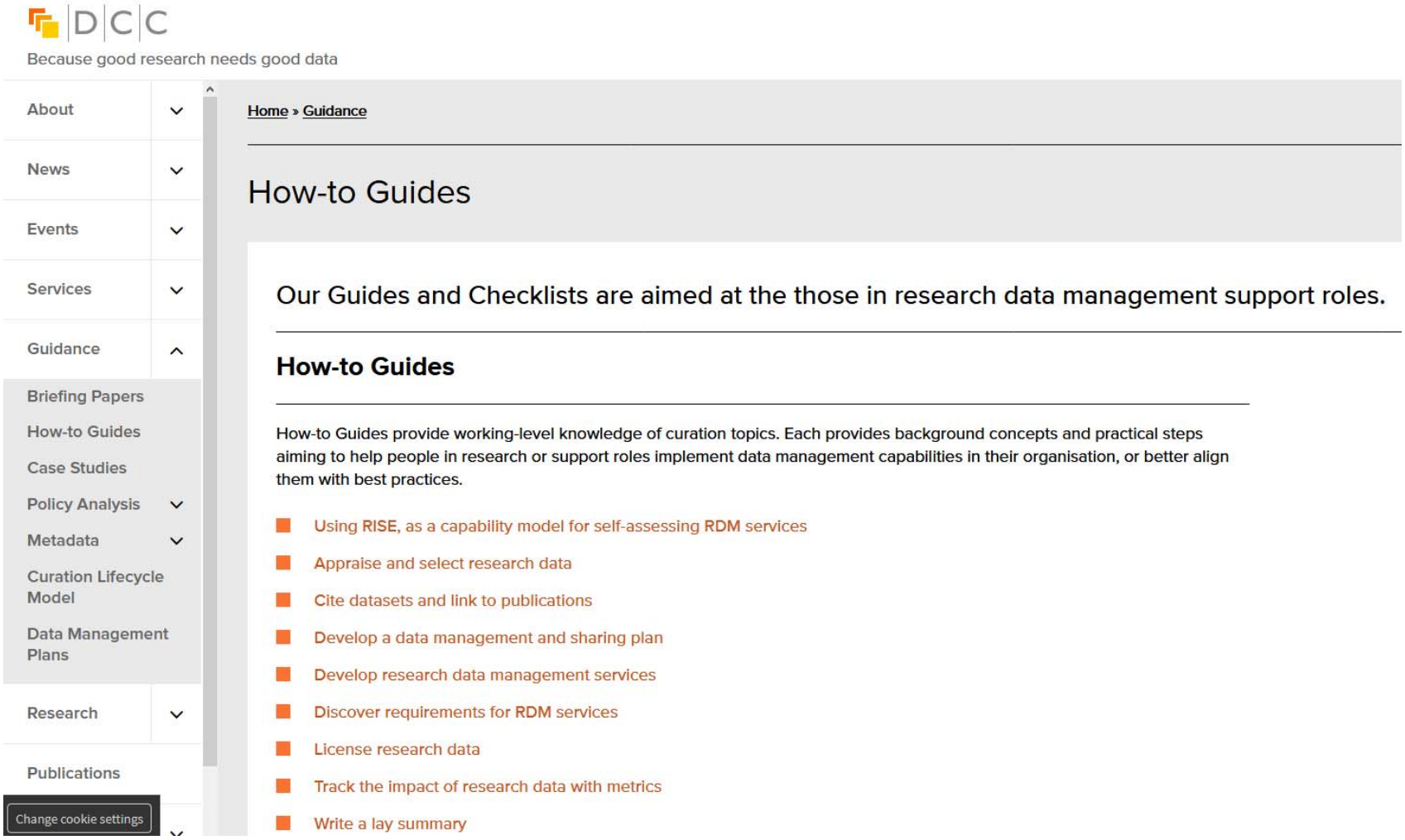
Research Data Management Plans

- Data Stewardship Wizard created by ELIXIR CZ and NL
- DMPonline of the Digital Curation Centre (DCC), UK
- DMPTool of University of California Curation Center of the California Digital Library (CDL), USA
- RDMO Research Data Management Organiser of the German Research Foundation, Germany
- Data Management Plan Catalogue of the LIBER Research Data Management Working Group
- Practical Guide on Research Data Management , developed by experts from Science Europe Member Organisations

Information about Persistent Identifiers (PID)

https://www.dcc.ac.uk/guidance/how-guides/

https://data.research.cornell.edu/content/best-practices



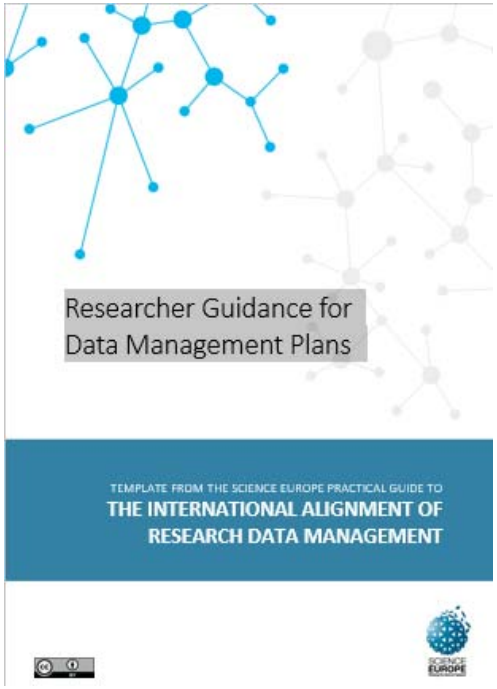
https://data.research.cornell.edu/content/best-practices

Best practices

Best practices documents created by the RDMSG to help you write a data management plan and manage your data.

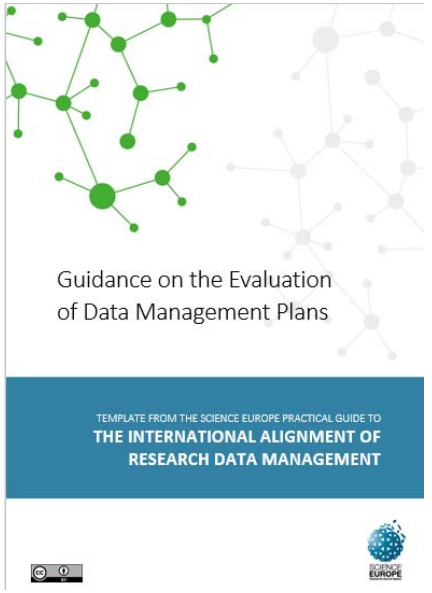
- Data citation
- Data storage and backup
- File formats
- File management
- Glossary of data management terms
- Guide to writing a Data Management Plan (DMP)
- Guide to writing "readme" style metadata
- Introduction to intellectual property rights in data management
- Metadata and describing data
- Preparing tabular data for description and archiving
- Sharing and archiving data
- Support for researchers in the Life Sciences

https://www.scienceeurope.org/our-priorities/research-data/research-data-management/



https://www.scienceeurope.org/media/411km040/se-rdm-template-3-researcher-guidance-for-data-management-plans.docx

Evaluación del RDM

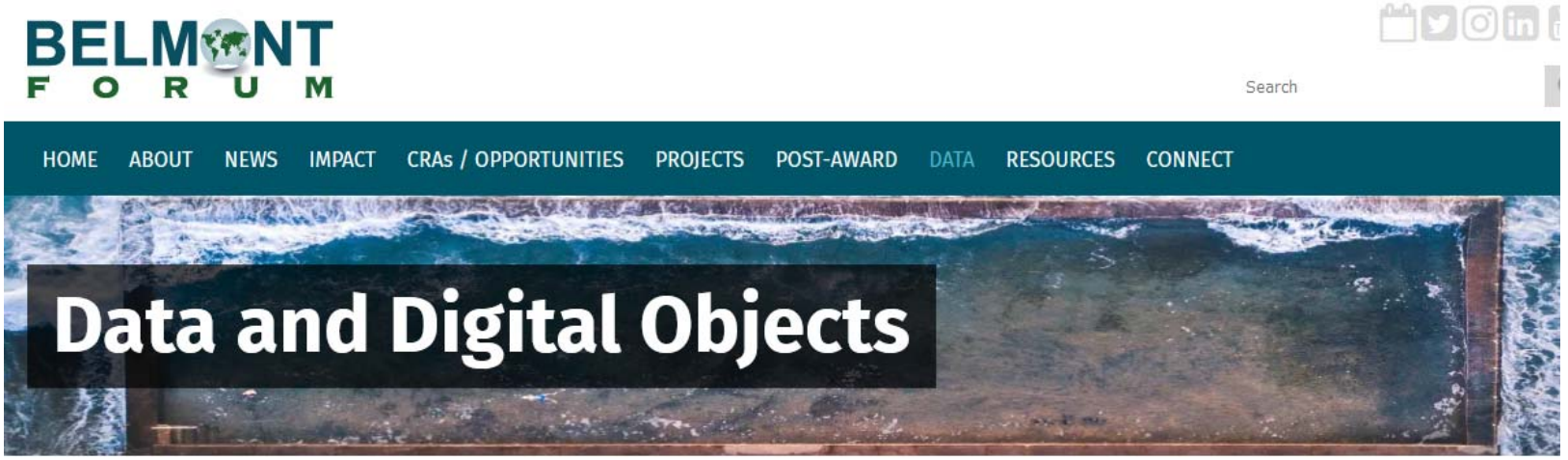


https://www.scienceeurope.org/media/22hpslfl/se-rdm-template-5-guidance-on-the-evaluation-of-data-management-plans.docx

▪ Evaluation·Rubric·for·Data·Management·Plans¶

DMP·Question¤	DMP·Guidance¤	Performance·Level¤	
Guidance·for·Researchers¤		**Sufficiently·Addressed¶** The·DMP…¤	**Insufficiently·Addressed¶** The·DMP…¤
Administrative·information¤	•→Provide·information·such·as·name·of·applicant,·project·number,·funding·programme,·version·of·DMP.¤	•→contains·the·minimal·information·required·to·identify·the·applicant·and·the·references·of·the·project.¤	•→provides·no·or·limited·information,·which·makes·it·hard·to·identify·who·is·responsible·for·the·project.¤
1···DATA·DESCRIPTION·AND·COLLECTION·OR·RE-USE·OF·EXISTING·DATA¤			
Guidance·for·Researchers¤		**Sufficiently·Addressed¶** The·DMP…¤	**Insufficiently·Addressed¶** The·DMP…¤
1a·How·will·new·data·be·collected·or·produced·and/or·how·will·existing·data·be·re-used?¤	•→Explain·which·methodologies·or·software·will·be·used·if·new·data·are·collected·or·produced.·¶ •→State·any·constraints·on·re-use·of·existing·data·if·there·are·any.·¶ •→Explain·how·data·provenance·will·be·documented.¶ •→Briefly·state·the·reasons·if·the·re-use·of·any·existing·data·sources·has·been·considered·but·discarded.¤	•→gives·clear·details·of·where·the·existing·data·come·from·and·how·new·data·will·be·collected·or·produced.·It·clearly·explains·methods·and·software·used.¶ •→explains,·if·existing·data·are·re-used,·how·these·data·will·be·accessed·and·any·constraints·on·their·re-use.¶ •→explains·clearly,·if·applicable,·why·new·data·must·be·collected,·rather·than·re-using·existing·data.¤	•→provides·little·or·no·details·on·where·the·data·come·from·and·what·data·will·be·collected·or·re-used.¶ •→does·not,·if·applicable,·provide·sufficient·rationale·for·generating·new·data.¤
1b·What·data·(for·example·the·kind,·formats,·and·volumes)·will·be·collected·or·produced?¤	•→Give·details·on·the·kind·of·data:·for·example,·numeric·(databases,·spreadsheets),·textual·(documents),·image,·audio,·video,·and/or·mixed·media.¶ •→Give·details·on·the·data·format:·the·way·in·which·the·data·is·encoded·for·storage,·often·reflected·by·the·filename·extension·(for·example·	•→clearly·describes·or·lists·what·data·types·will·be·generated·(for·example·numeric,·textual,·audio,·or·video)·and·their·associated·data·formats,·including,·if·needed,·data·conversion·strategies.·¶ •→explains·why·certain·formats·have·been·chosen·and·indicates·if·they·are·in·open·and·standard·format.·If·a·proprietary·	•→provides·no·or·little·details·on·what·data·types·will·be·generated·and·does·not·provide·a·valid·reason·for·this·omission·(for·example·a·statement·that·no·data·will·be·produced·or·generated).¶ •→only·lists/describes·the·kinds·of·data·

¶



- Data Policies and Principles
- Toolkit
- Policy Comparison Tool
- Data Management Plan Scorecard
- History

Data Policies and Principles

Recognizing the crucial role of open and effective data and information exchange to the Belmont Challenge, the Belmont Forum adopted open Data Policy and Principles based on the recommendations from the Community Strategy and Implementation Plan (CSIP) at its 2015 annual meeting of Principals in Oslo, Norway. The policy signals a commitment by funders of global environmental change research to increase access to scientific data, a step widely recognized as essential to making informed decisions in the face of rapid changes affecting the Earth's environment.

Belmont Forum Data Policy and Principles

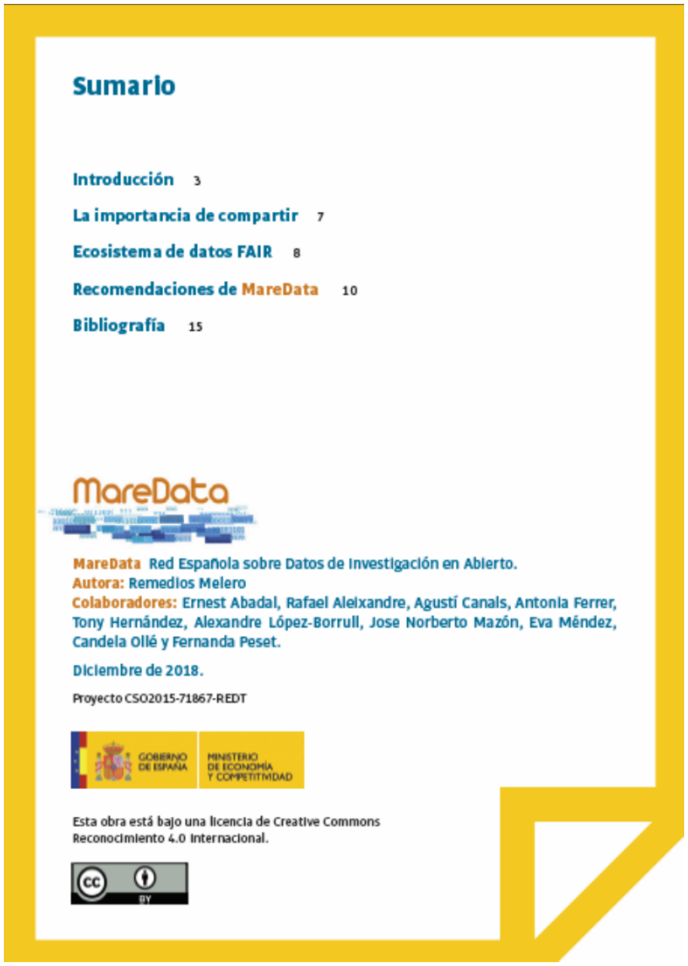
The Belmont Forum adopts this data policy and the following principles to widen access to data and promote its long-term preservation in global change research; help improve data management and exploitation; coordinate and integrate disparate organizational and technical elements; fill critical global e-infrastructure gaps; share best practices; and foster new data literacy.

Plantilla de para evaluar los DMP
https://zenodo.org/record/3530933#.YKzHpqHtbt4

Comparador de políticas
https://www.belmontforum.org/data/#policyComparison

Belmont Forum DMP Scorecard (v.20190819)

Criteria	Complete Response (score = 2)	Incomplete Response (score = 1)	No Response (score = 0)	SCORE
1. What types of datasets and other digital outputs of long-term value do you expect the project will produce or reuse?				
1.1 Plan lists the types of data and other digital outputs of long-term value. *(e.g. text, databases, images, 3D models, software, audio files, code, video files, reports, surveys, patient records, samples, and so forth)*	Datasets and other digital outputs of long-term value are identified, including data type and encoding. *"Environmental data will be delivered as NetCDF (Network Common Data Format) files. Raster files will use the raster2pgsql PostGIS module. Maps and other geographic data will use shapefiles."* *"Transcripts and coding will be provided in text files. Audio recording will be MP3 format."*	Datasets and other digital outputs of long-term value are identified, but lack detail for users beyond the project to understand. *"A combination of geo-referenced data at various spatial, temporal, and taxonomic scales (e.g., populations, regions, nations, circumpolar, biomes, habitats) will comprise our data of long-term value."* *"Long-term value data include data from anthropological field studies: transcripts of interviews and discussion workshops, associated metadata."*	No information about data types is included.	X
1.2 Plan describes how the data and other digital outputs will be collected, captured, or created. *(e.g., new observations, results from models, reuse of other data, or other)*	Clearly defines how data will be collected, captured or created, including methods, instruments, software, or infrastructure where relevant. *"The MIP compatible model runs to the ISI-MIP servers, where existing Fish-MIP runs are already stored. Scenario outputs will be archived."* *"Socio-economic data will include household food security, nutrition, and demographic data. Spatial data produced will include ground-truthed land use/land cover data ~3 km from 50 farms, land use scenario maps for 12 villages and 4 regions. All ecological/social data will be recorded on physical datasheets and entered directly into Excel or STATA."*	Missing some details regarding how some of the data will be produced; makes assumptions about reviewer knowledge of methods or practices. *"Models will produce a broad range of output simulation data."* *"Data collection includes gathering in-the-field various phytoplankton, zooplankton, fish and flooded forest biodiversity."*	No information about data collection, capture or creation.	X
1.3	Datasets and other digital outputs volume estimated.	Datasets and other digital outputs amount is vaguely estimated or	Amount of expected data is	X



Sumario

MareData

MareData Red Española sobre Datos de Investigación en Abierto.
Autora: Remedios Melero
Colaboradores: Ernest Abadal, Rafael Aleixandre, Agustí Canals, Antonia Ferrer, Tony Hernández, Alexandre López-Borrull, Jose Norberto Mazón, Eva Méndez, Candela Ollé y Fernanda Peset.

Diciembre de 2018.
Proyecto CSO2015-71867-REDT

GOBIERNO DE ESPAÑA MINISTERIO DE ECONOMÍA Y COMPETITIVIDAD

https://digital.csic.es/handle/10261/173801

1. Adoptar los principios **FAIR**	2. Seleccionar los datos de interés (limpieza de datos)	3. Aplicar los principios **FAIR** también a los metadatos y al plan de gestión de datos	4. Facilitar, en los informes de evaluación de los proyectos, los beneficios de hacer tus datos **FAIR**	5. Elaborar un plan de gestión de datos de cómo hacer tus datos **FAIR**
6. Indicar en el plan de gestión de datos si se reutilizan datos (evitar redundancia)	7. Utilizar estándares interoperables con el ecosistema de datos **FAIR**	8. Facilitar los datos de investigación de acuerdo con estándares reconocidos	9. Asignar licencias abiertas para propiciar la reutilización	10. Descripción detallada de los ensayos para facilitar su reproducibilidad
11. Identificar el repositorio de adecuado para depositar los datos	12. Verificar los términos de las políticas respecto a los datos de investigación	13. Promover la citación de los *datasets*	14. Tener en cuenta, los códigos de conducta reconocidos a nivel internacional	15. Contribuir a las buenas prácticas respecto a la gestión de los datos
16. Participar y elaborar programas de formación en gestión de datos	17. Contribuir al desarrollo de la *open science*			

Recomendaciones para la gestión de datos de investigación

MareData

Red Española sobre Datos de Investigación en Abierto

Los datos y las publicaciones

Panton Principles

Principles for Open Data in Science

Resumidamente:

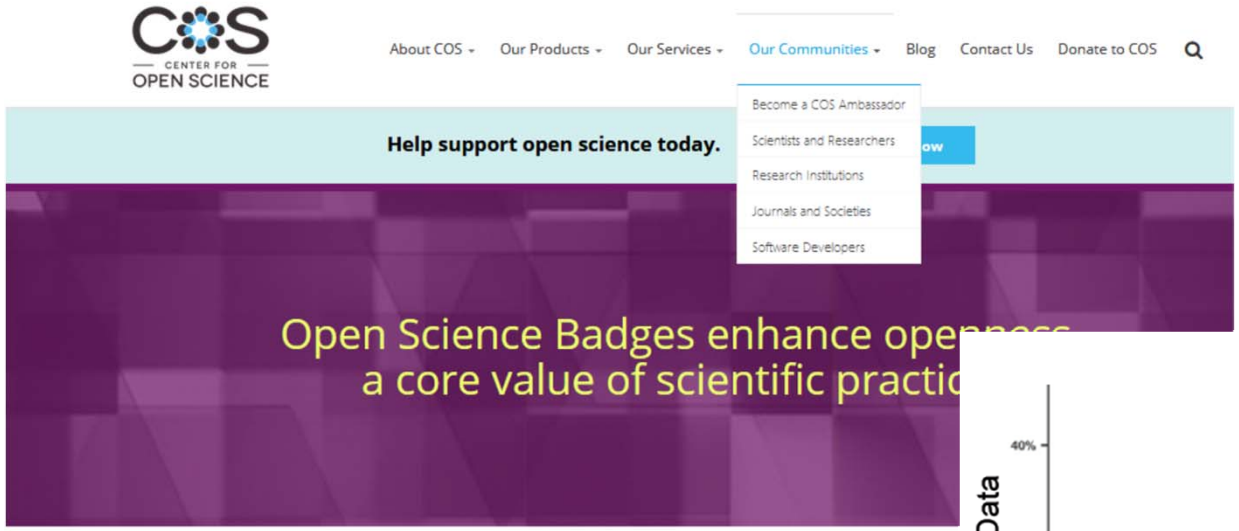
1. Términos claros del editor sobre lo que se puede hacer con los datos publicados
2. Utilizar licencias que sean adecuadas para el tratamiento de los datos
3. Evitar licencias que restrinjan limiten el uso comercial o la creación de obras trabajos derivados
4. Se recomienda encarecidamente que los datos generados con proyectos financiados con fondos públicos, sean de dominio público, mediante el uso de licencias al uso

Opciones para compartir los datos de una publicación (revista). Qué, cuando, dónde

- Adjuntar los datasets como material complementario para su evaluación
- La revista crea un repositorio propio de datos para su depósito al envío del trabajo
- La revista crea un repositorio propio de datos para su depósito después de la aceptación del trabajo
- La revista integra en su workflow el depósito de datasets (p.e., con el plugin dataverse del OJS)
- La revista recomienda el depósito en algún repositorio externo

Recomendaciones y directrices

https://cos.io/our-services/open-science-badges/

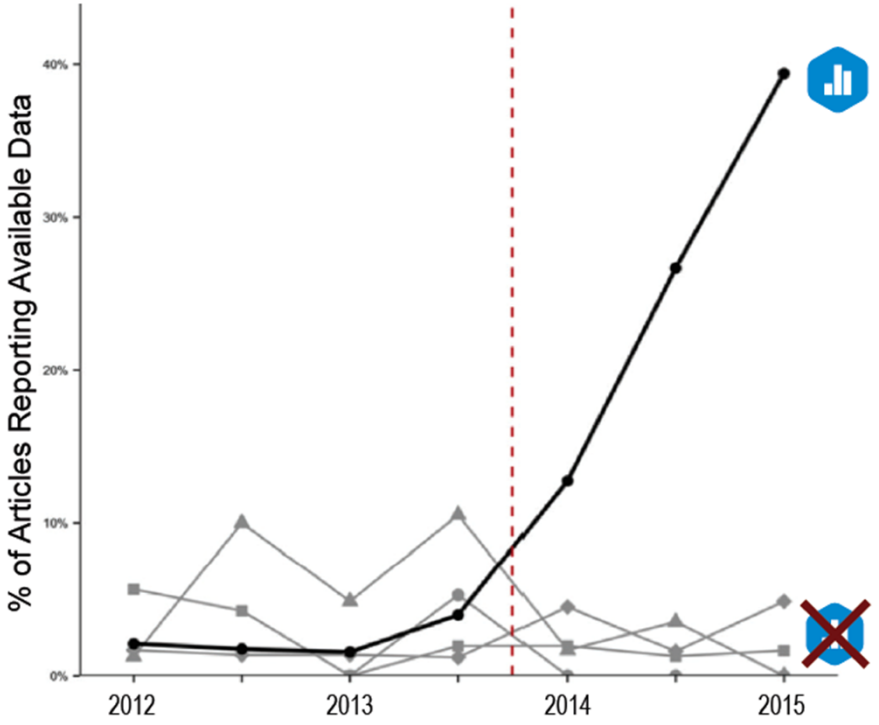


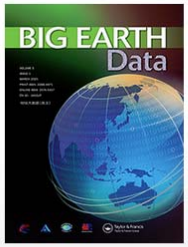
What are Open Science Badges?

- Badges to acknowledge open science practices are incentives
  or to preregister.
- Badges signal to the reader that the content has been made a
  persistent location.

Badges seem silly. Do they work?

- Yes. Implementing these badges dramatically increases the ra
- A recent systematic review identified this badging program as
  program that this effective at increasing the rates of data shar
- View a list of journals and organizations that have adopted ba





Big Earth Data

An open access journal

This journal supports Open Science Badges

Publishes research on Big Data in the earth sciences, including earth observation, earth systems monitoring, atmospheric science, marine science and geology.

Enter keywords, authors, DOI, ORCID etc

This Journal

Advanced search

ARTICLE OPEN ACCESS

Announcing the journal of the medical library association's data sharing policy

Akers K, Read K, Amos L et al. See more

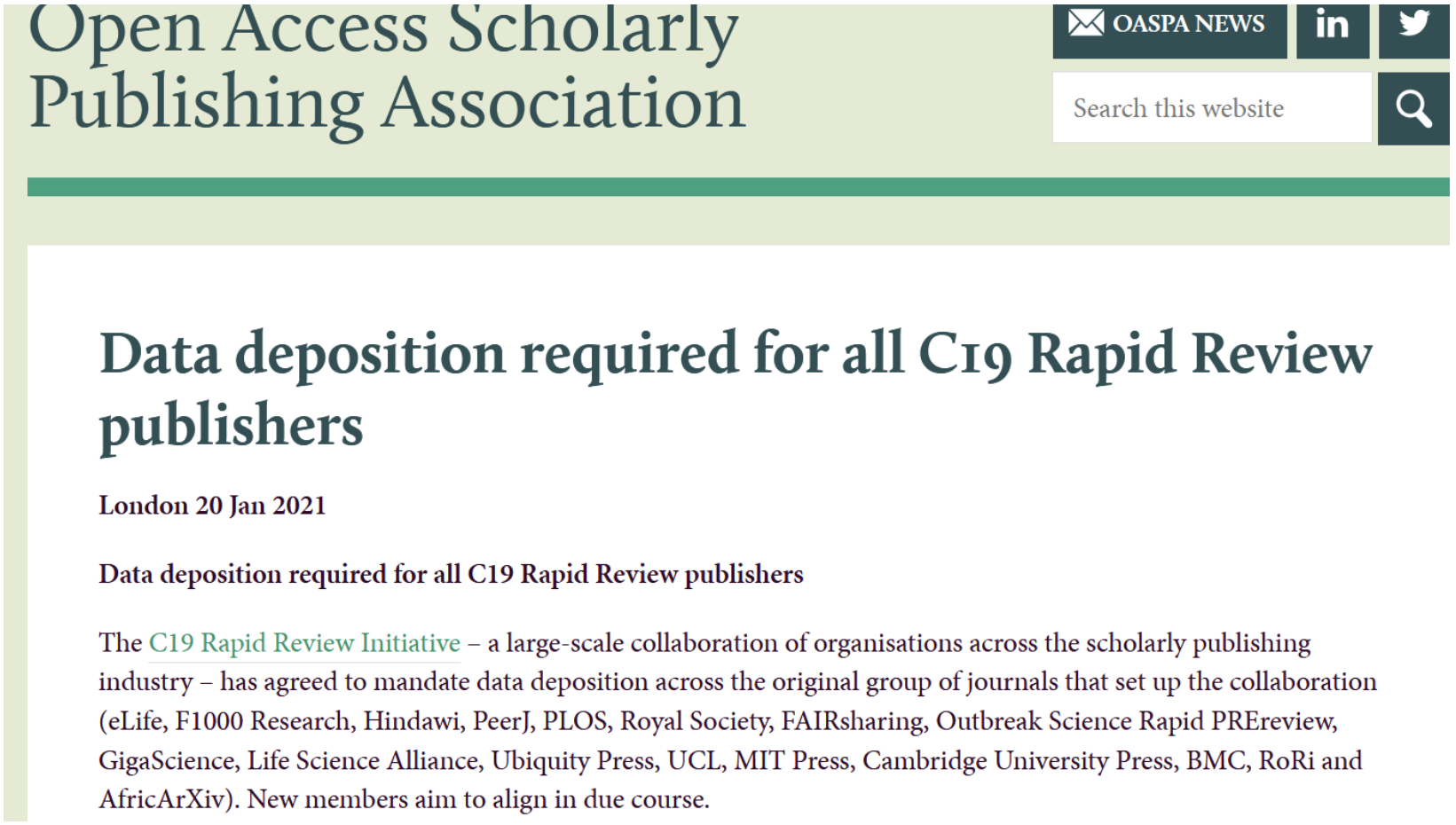
Journal of the Medical Library Association

DOI: 10.5195/jmla.2019.801

- Starting October 1, 2019, authors of Original Investigation and Case Report manuscripts are required to deposit the de-identified data associated with their manuscripts in a repository and include a "Data Availability Statement" in their manuscripts describing where and how the data can be accessed. …

- The *JMLA* defines "data" as the digital materials underlying the results described in the manuscript, including spreadsheets, text files, interview recordings or transcripts, images, videos, output from statistical software, or computer code or scripts.
- Shared data should be appropriately de-identified to prevent revealing the identity of study participants. MLA, the *JMLA*, and individual members of the *JMLA* editorial team are not liable for any harm or damage resulting from the insufficient de-identification of data associate with *JMLA* articles

- When possible, authors are encouraged to apply a license that is at least as permissive as a Creative Commons Attribution (CC BY) license to the data



Open Access Scholarly Publishing Association

OASPA NEWS

Search this website

Data deposition required for all C19 Rapid Review publishers

London 20 Jan 2021

Data deposition required for all C19 Rapid Review publishers

The C19 Rapid Review Initiative – a large-scale collaboration of organisations across the scholarly publishing industry – has agreed to mandate data deposition across the original group of journals that set up the collaboration (eLife, F1000 Research, Hindawi, PeerJ, PLOS, Royal Society, FAIRsharing, Outbreak Science Rapid PREreview, GigaScience, Life Science Alliance, Ubiquity Press, UCL, MIT Press, Cambridge University Press, BMC, RoRi and AfricArXiv). New members aim to align in due course.

COVID Rapid Review Initiative members **must have data shared in a public repository rather than just available on request**. The new common policy is to meet the TOP Data Transparency Level II that requires that "Data must be posted to a trusted repository. Exceptions must be identified at article submission". This means mandating data sharing in a public repository rather than just ensuring the authors publish a Data Availability Statement (DAS). Any DAS must now explicitly list the repositories where the data are publicly available (subject to ethical considerations).

https://copdess.org/copdess-suggested-author-instructions-and-best-practices-for-journals/

- Data Policy Statement (data accesibles en el momento de la publicación, depositados en un repositorio fiable)

- Data Citation (siguiendo los *open data principles)*

- Sample Citation and Identification:
- Crossref Funder Registry
- ORCIDs

} Identificadores únicos y persistentes

The Transparency and Openness Promotion (TOP) guidelines

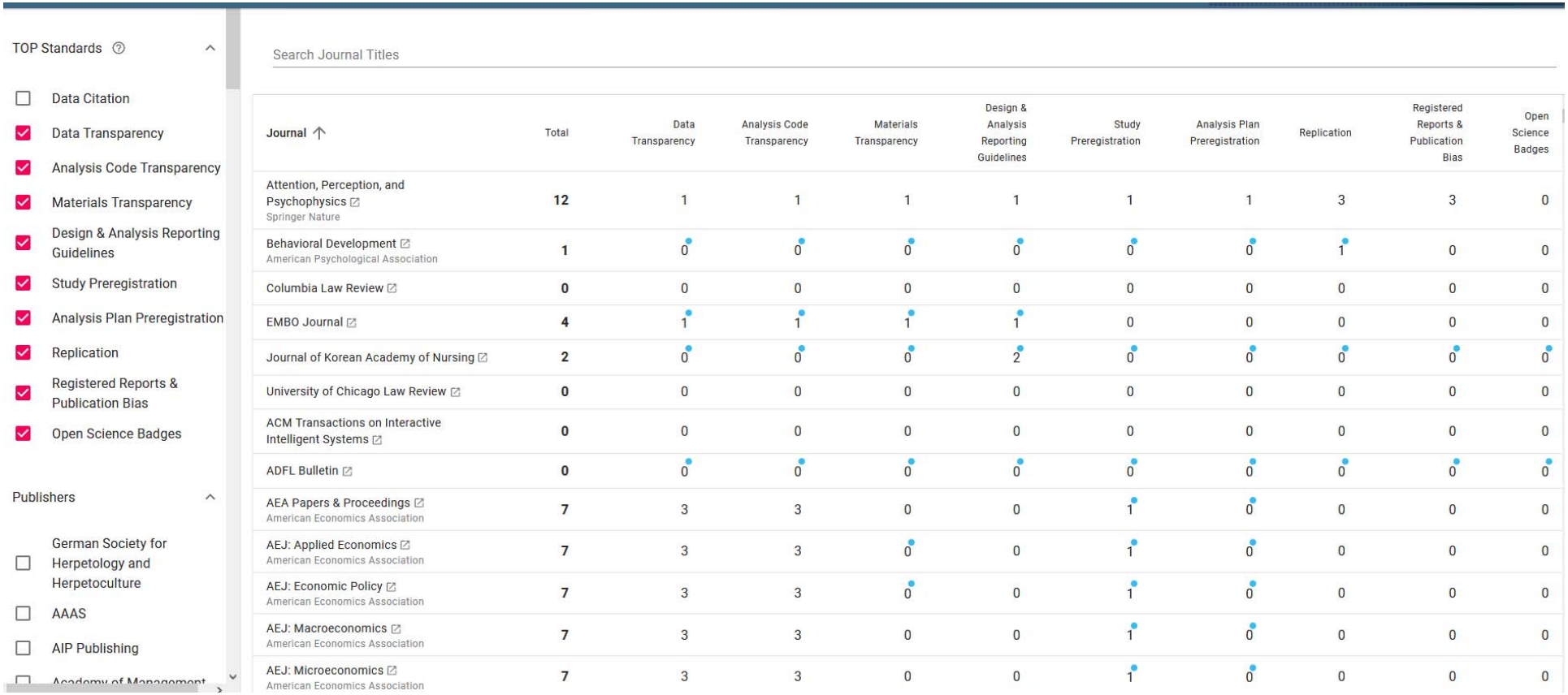
https://www.cos.io/initiatives/top-guidelines

https://osf.io/t2yu5/

Standard	Level 1	Level 2	Level 3
Data citation	Journal describes citation of data in guidelines to authors with clear rules and examples.	Article requires appropriate citation for data used consistent with the journal's author guidelines.	Article is not published until providing appropriate citation for data following journal's author guidelines.
Data transparency	Articles must state whether or not data are available.	Articles must have publicly available data, or explain why ethical/legal constraints prevent it.	Articles must have publicly available data and must be used to computationally reproduce or confirm results prior to publication.
Analytical code transparency	Articles must state whether or not code is available.	Articles must have publicly available code, or explain why ethical/legal constraints prevent it.	Articles must have publicly available code and must be used to computationally reproduce or confirm results prior to publication.
Materials transparency	Articles must state whether or not materials are available.	Articles must have publicly available materials, or explain why ethical/legal constraints prevent it.	Articles must have publicly available materials and must be used to computationally reproduce or confirm results prior to publication.
Reporting guidelines	Journal articulates design transparency standards.	Journal requires adherence to design transparency standards for review and publication.	Journal requires and enforces adherence to design transparency standards for review and publication.
Study preregistration	Articles will state if work was preregistered.	Article states whether work was preregistered and, if so, journal verifies adherence to preregistered plan.	Journal requires that confirmatory or inferential research must be preregistered.
Analysis plan preregistration	Articles will state if work was preregistered with an analysis plan.	Article states whether work was preregistered with an analysis plan and, if so, journal verifies adherence to preregistered plan.	Journal requires that confirmatory or inferential research must be preregistered with an analysis plan.
Replication	Journal encourages submission of replication studies.	Journal will review replication studies blinded to results.	Registered Reports for replications as a regular submission option.
Publication bias	Journal states that significance or novelty are not criteria for publication decisions.	Journal will review (novel) studies blinded to results.	Journal accepts Registered Reports for novel studies as a regular submission option.
Open science badges	Journal awards 1 or 2 open science badges	Journal awards all 3 open science badges	

TOP factor. Mide el grado de cumplimiento con las directrices TOP.
https://www.topfactor.org/

Se basa en tres niveles en función de su implementación y especificación: nivel 1, nivel 2 y nivel 3, el nivel 0 corresponde a “se recomienda” o “no se especifica”



TOP Standards

- [ ] Data Citation
- [x] Data Transparency
- [x] Analysis Code Transparency
- [x] Materials Transparency
- [x] Design & Analysis Reporting Guidelines
- [x] Study Preregistration
- [x] Analysis Plan Preregistration
- [x] Replication
- [x] Registered Reports & Publication Bias
- [x] Open Science Badges

Publishers

- [ ] German Society for Herpetology and Herpetoculture
- [ ] AAAS
- [ ] AIP Publishing
- [ ] Academy of Management

Search Journal Titles

Journal	Total	Data Transparency	Analysis Code Transparency	Materials Transparency	Design & Analysis Reporting Guidelines	Study Preregistration	Analysis Plan Preregistration	Replication	Registered Reports & Publication Bias	Open Science Badges
Attention, Perception, and Psychophysics (Springer Nature)	12	1	1	1	1	1	1	3	3	0
Behavioral Development (American Psychological Association)	1	0	0	0	0	0	0	1	0	0
Columbia Law Review	0	0	0	0	0	0	0	0	0	0
EMBO Journal	4	1	1	1	1	0	0	0	0	0
Journal of Korean Academy of Nursing	2	0	0	0	2	0	0	0	0	0
University of Chicago Law Review	0	0	0	0	0	0	0	0	0	0
ACM Transactions on Interactive Intelligent Systems	0	0	0	0	0	0	0	0	0	0
ADFL Bulletin	0	0	0	0	0	0	0	0	0	0
AEA Papers & Proceedings (American Economics Association)	7	3	3	0	0	1	0	0	0	0
AEJ: Applied Economics (American Economics Association)	7	3	3	0	0	1	0	0	0	0
AEJ: Economic Policy (American Economics Association)	7	3	3	0	0	1	0	0	0	0
AEJ: Macroeconomics (American Economics Association)	7	3	3	0	0	1	0	0	0	0
AEJ: Microeconomics (American Economics Association)	7	3	3	0	0	1	0	0	0	0

Propuesta de 6 tipos tipo de políticas editoriales sobre los datos de investigación, en función de 14 variables.
Se presenta lo resultados de políticas de revistas analizadas de Springer Nature, Elsevier, Wiley, PLOS

https://www.rd-alliance.org/group/data-policy-standardisation-and-implementation-ig/outcomes/developing-research-data-policy

Hrynaszkiewicz, I., Simons, N., Hussain, A., Grant, R., & Goudie, S. (2020). Developing a Research Data Policy Framework for All Journals and Publishers. Data Science Journal, 19(1), 5. http://doi.org/10.5334/dsj-2020-005

14 journal research data policy features arranged as 6 policy types (tiers)

	Policy 01	Policy 02	Policy 03	Policy 04	Policy 05	Policy 06
Definition of the research data	○	○	○	○	○	○
Exceptions to policy	○	○	○	○	●	●
Embargoes	○	○	○	●	●	●
Supplementary materials	○	○	○	●	●	●
Data repositories	○	○	○	●	●	●
Data citation	○	○	○	○	●	●
Data licensing	○	○	○	○	○	○
Researcher/ author support	●	●	●	●	●	●
Data availability statements		○	●	●	●	●
Data formats and standards				○	○	●
Mandatory data sharing (specific data types)				●	●	●
Mandatory data sharing (all papers)				○	●	●
Peer review of data				○	○	●
Data Management Plans (DMPs)				○	○	○

○ **Provide information**

The text for the policy feature will be included in the policy template but it is clear that the feature will not be enforced and checked as part of the publishing or peer review process

● **Provide information and action**

The text of the policy feature is included and makes clear where applicable that the feature will be checked and enforced in the publishing or peer-review process

Recomendaciones para la elaboración de una política editorial sobre los datos de investigación después del análisis de 201 revistas indexadas en WOS y/o Scopus en Library and Information Science Journals

Table 2. Recommended journal open data policy elements.

Category	Recommended Policy Elements
Application	• Describe precisely the types of data to which open data policies apply; • Target policies at a clearly defined body of research (e.g., quantitative data, original research, or methods-based requirements).
Exemptions	• Identify the circumstances that warrant exemptions as narrowly as possible; • Describe valid concerns related to privacy or proprietary data; • Require researchers to justify exemptions during the submission process.
Timing	• Indicate whether data should be publicly accessible prior to submission, prior to publication, or within a set time period after publication; • Indicate if datasets will be considered in the peer-review process.
Licensing and Access	• Prescribe specific open licensing terms that authors should apply to datasets; • Describe how authors should proceed if repository policies do not accommodate the preferences of the journal; • If access to data will be justifiably restricted, require authors to detail any conditions under which they will be made available.
Formatting and Repositories	• Encourage authors to use preservation-friendly formats for data, with examples provided based on data type; • Identify preferred data repositories; • If journals do not specify data repositories, identify required characteristics of eligible repositories, including preservation and persistent linking features.
Data Availability Statements	• Require data availability statements for all eligible articles; • Identify each element required in statements, such as location of and links to data, full citations, and terms of access; • Require a specific justification when data are not available, rather than providing templates.
Principles	• State the underlying principles of the open data policy; • Clearly outline the potential contributions of open data to the transparency and reproducibility of research, the potential for secondary analysis, and the ethos of research data as a public good.

Jackson, B. Open Data Policies among Library and Information Science Journals. Publications **2021**, 9, 25. https://doi.org/10.3390/publications9020025

Ejemplos….

Data Availability https://reviewer.elifesciences.org/author-guide/journal-policies

To maintain high standards of research reproducibility, and to promote the reuse of new findings, eLife requires all datasets associated with an article to be made freely and widely available (unless there are strong reasons to restrict access, for example in the case of human subjects data), in the most useful formats, and according to the relevant reporting standards.

Wherever possible, authors should make major datasets available using domain-specific public archives (for example, GenBank, Protein Data Bank, and ClinicalTrials.gov), or generic databases (for example, Dryad, Dataverse or the Open Science Framework) where a domain specific archive does not exist.

Authors using original data must:

- make the data available at a trusted digital repository (however, if all data required to reproduce the reported analyses appears in the article text, tables, and figures then it does not also need to be posted to a repository);
- include all variables, treatment conditions, and observations described in the manuscript;
- provide a full account of the procedures used to collect, pre-process, clean, or generate the data;
- provide research materials and description of procedures necessary to conduct an independent replication of the research.

Trusted repositories adhere to policies that make data discoverable, accessible, usable, and preserved for the long term. Trusted repositories also assign unique and persistent identifiers. Author-maintained websites are not compliant with this requirement.

Data Availability

The following policy applies to all PLOS journals, unless otherwise noted.

Introduction

PLOS journals require authors to make all data necessary to replicate their study's findings publicly available without restriction at the time of publication. When specific legal or ethical restrictions prohibit public sharing of a data set, authors must indicate how others may obtain access to the data.

When submitting a manuscript, authors must provide a Data Availability Statement describing compliance with PLOS' data policy. If the article is accepted for publication, the Data Availability Statement will be published as part of the article.

Acceptable data sharing methods are listed below, accompanied by guidance for authors as to what must be included in their Data Availability Statement and how to follow best practices in research reporting.

PLOS believes that sharing data fosters scientific progress. Data availability allows and facilitates:

- Validation, replication, reanalysis, new analysis, reinterpretation or inclusion into meta-analyses;
- Reproducibility of research;
- Efforts to ensure data are archived, increasing the value of the investment made in funding scientific research;
- Reduction of the burden on authors in preserving and finding old data, and managing data access requests;
- Citation and linking of research data and their associated articles, enhancing visibility and ensuring recognition for authors, data producers and curators.

Publication is conditional on compliance with this policy. If restrictions on access to data come to light after publication, we reserve the right to post a Correction, an Editorial Expression of Concern, contact the authors' institutions and funders, or, in extreme cases, retract the publication.

Neuropsychology (https://www.apa.org/pubs/journals/neu/index?tab=4)

Data, Materials, and Code

Authors must state whether data and study materials are available and, if so, where to access them. Recommended repositories include APA's repository on the Open Science Framework (OSF), or authors can access a full list of other recommended repositories.

In both the Author Note and at the end of the Method section, specify whether and where the data and material will be available or note the legal or ethical reasons for not doing so. For submissions with quantitative or simulation analytic methods, state whether the study analysis code is available, and, if so, where to access it (or the legal or ethical reason why it is not available).

Journal of Animal ecology https://besjournals.onlinelibrary.wiley.com/hub/journal/13652656/author-guidelines

Data Availability Statement

To enable readers to locate archived data from papers, we require that authors list the database and the respective accession numbers or DOIs for all data from the manuscript that has been made publicly available. For example, "Data available from the Dryad Digital Repository http://dx.doi.org/10.5061/dryad.41qh7 (Kiere & Drummond 2016)." When a DOI is available for the data, the full data citation should also be given in the reference list.

GigaScience. https://academic.oup.com/gigascience/pages/editorial_policies_and_reporting_standards

Supporting Data

All data on which the conclusions given in the publication are based should be available to readers, either in our *Giga*DB repository, or in community established repository where available and appropriate (for examples see below). *GigaScience* fully supports the recommendations of the National Academies regarding data sharing (see Board on Life Sciences, Sharing Publication-Related Data and Materials: Responsibilities of Authorship in the Life Sciences).

Authors using unpublished genomic data are expected to abide by the guidelines of the Fort Lauderdale and Toronto agreements. Based on broadly accepted scientific community standards, the key requirement of third parties using genomic data is to contact the owners of unpublished data (i.e. the principal investigator and sequencing center) prior to undertaking their research, to advise them about their planned analyses.

Should the manuscript be approved for publication, the accompanying datasets must be accessible by any researcher wishing to use them under a Creative Commons CC0 waiver, without restrictions, such as the need for a material transfer agreement. Exceptions to the use of a Creative Commons CC0 waiver for data are made for data that has specific legal restrictions for open sharing, but these must be submitted to a persistent, but restricted database for other researchers to submit a request for use form.

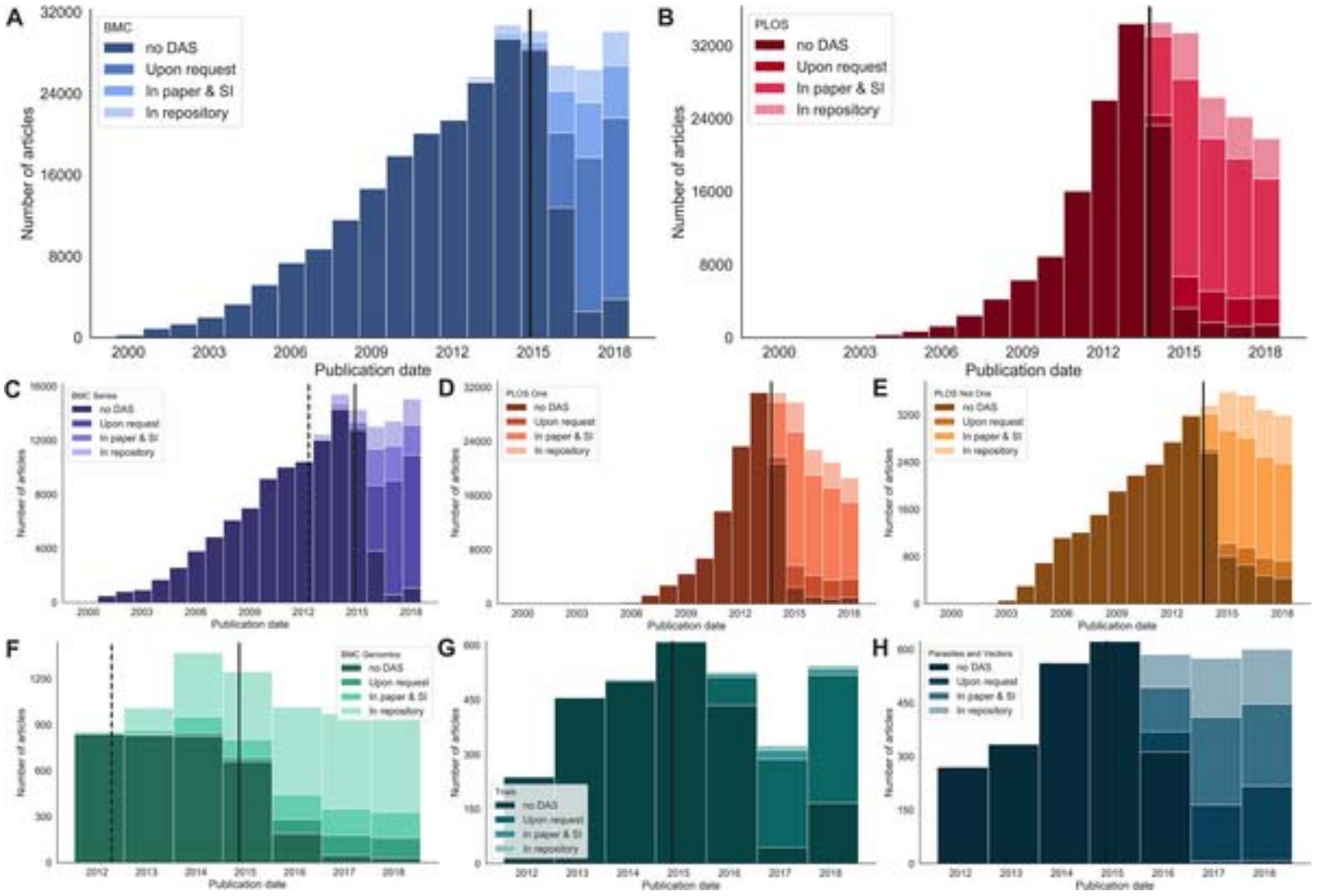
Scientific Data (https://www.nature.com/sdata/policies/data-policies)

Data Policies

Data Descriptors, *Scientific Data*'s primary article type, describe scientifically valuable datasets. These datasets must be made available to editors and referees at the time of submission, and must be shared with the scientific community as a condition of publication. Here, we provide information on the types of data that should be archived, guidance for authors on selecting a suitable repository for their data, and how to archive sensitive data. *Scientific Data*'s data policies are compatible with the standardised research data policies set out by Springer Nature, and the requirements of the Data Policy Standardisation and Implementation Interest Group of the Research Data Alliance.

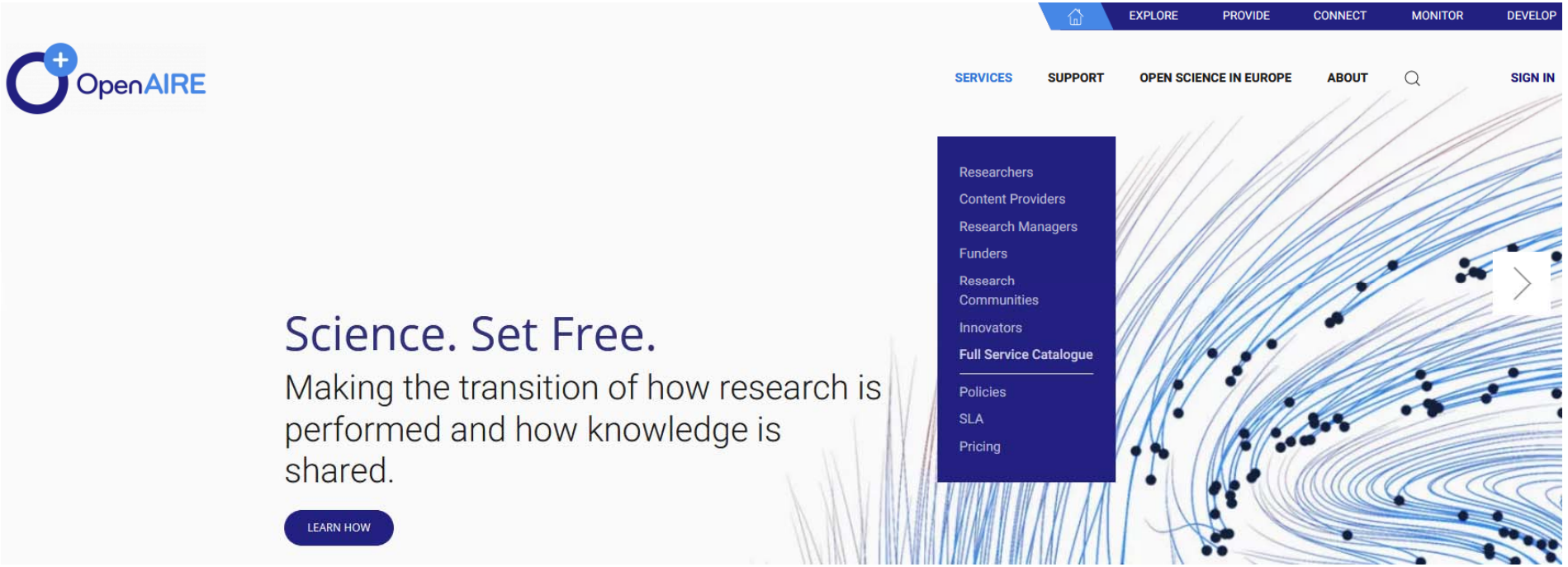
Please read on for our data deposition policies, and please contact us if you would like additional advice on how best to meet these requirements for your own data.

Efecto de la implementación de una política editorial sobre los datos de investigación subyacentes a las publicaciones. Caso de revistas de BMC y PLoS



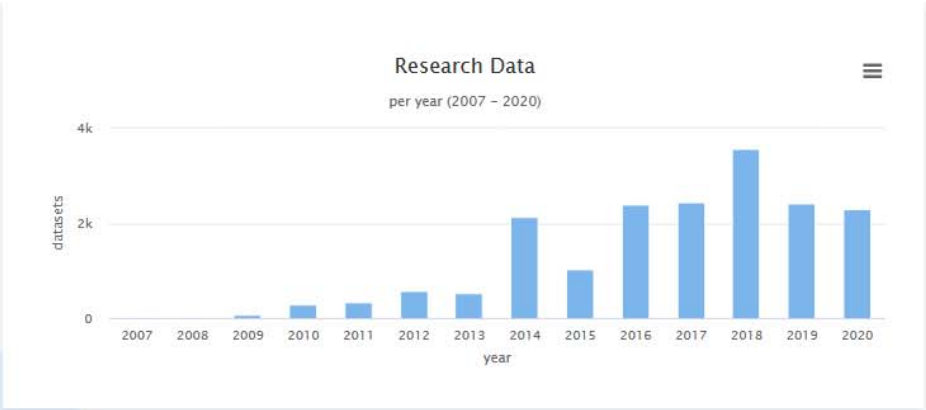
Colavizza G, Hrynaszkiewicz I, Staden I, Whitaker K, McGillivray B (2020) The citation advantage of linking publications to research data. PLOS ONE 15(4): e0230416. https://doi.org/10.1371/journal.pone.0230416
https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0230416

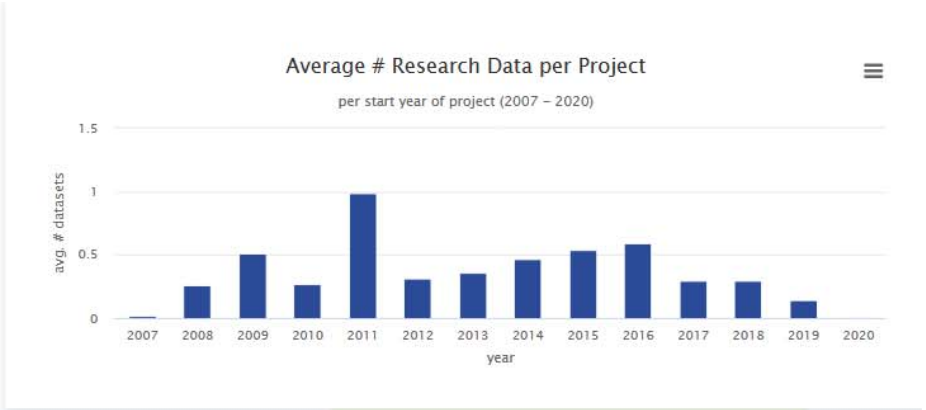
Servicios OpenAire

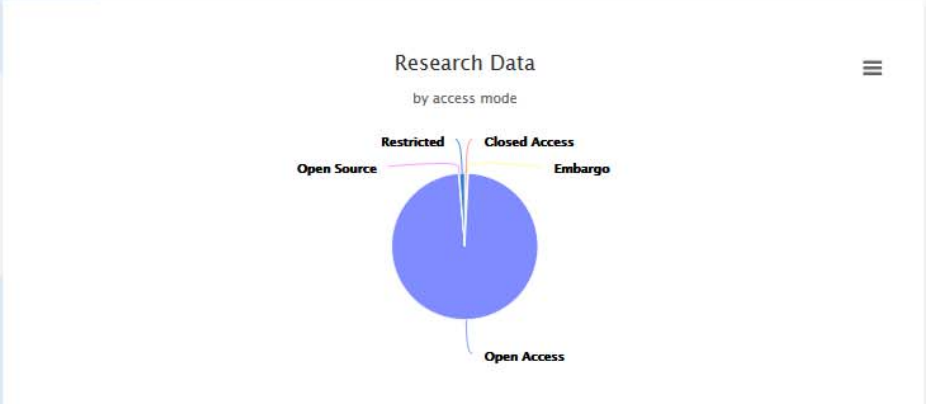


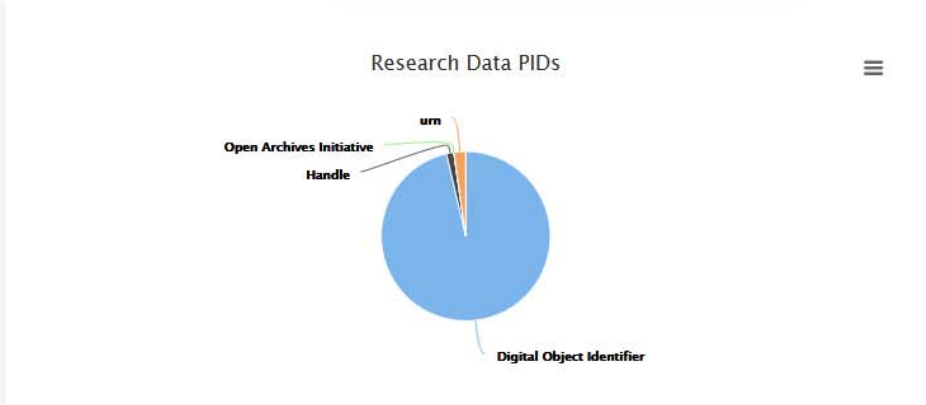
https://www.openaire.eu/

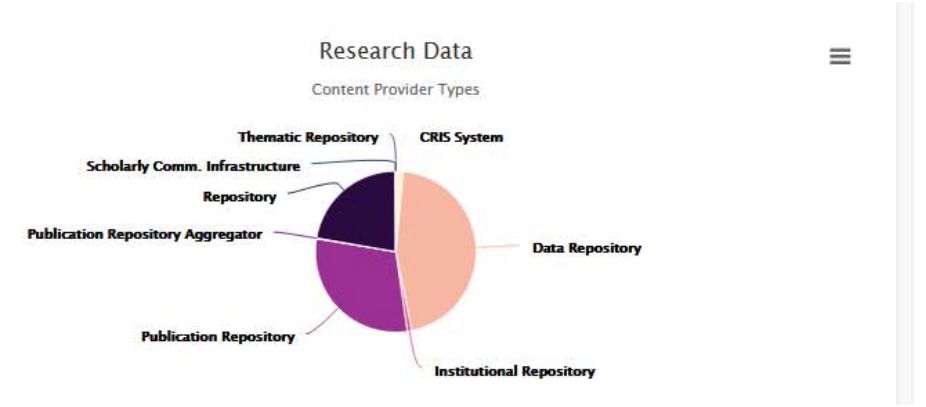
https://www.openaire.eu/openaire-nexus-public-launch-event

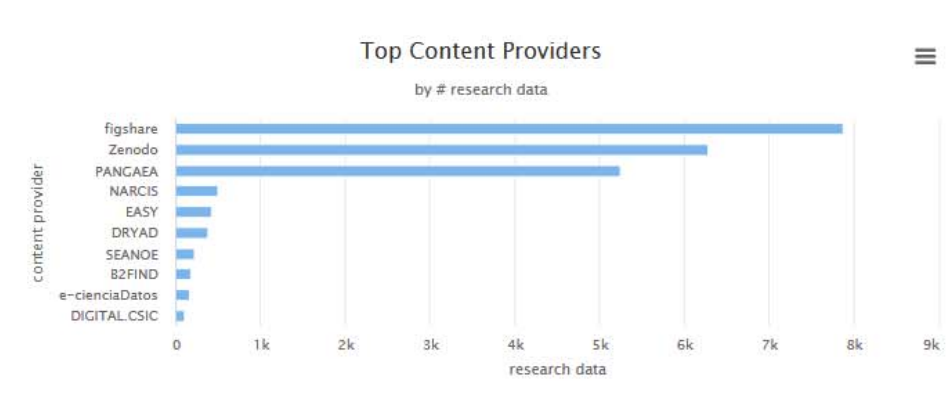












https://monitor.openaire.eu/dashboard/ec/output/research-data

¡Terminamos!