



Nadia Tonello
BSC Head of Data
Management
nadia.tonello@bsc.es

Como elaborar un plan de gestión de datos

Sumario

- Motivación
- Definiciones y conceptos básicos
- Plantillas
- Contenido de un DMP
- Evaluación de un DMP
- Q&A

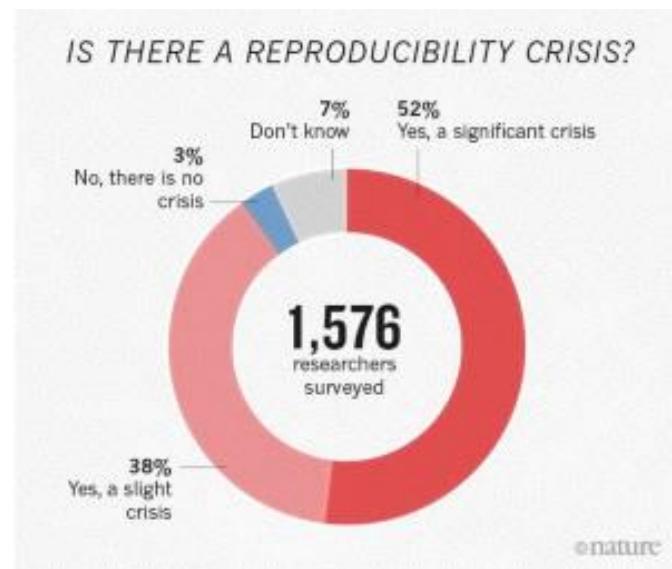
Planificación¹ de la gestión de datos de investigación

Motivación
Definiciones
Conceptos básicos

Motivación

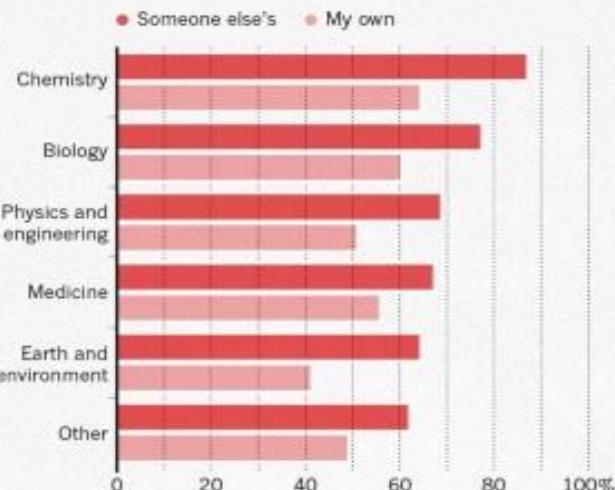
Crisis metodológica en las ciencias causada por las dificultades encontradas por los investigadores para replicar los resultados de muchos de los experimentos científicos en investigaciones posteriores, bien por investigadores independientes o bien por los investigadores originales de estos estudios.

1,500 scientists lift the lid on reproducibility
Monya Baker, *Nature* , 25 May 2016



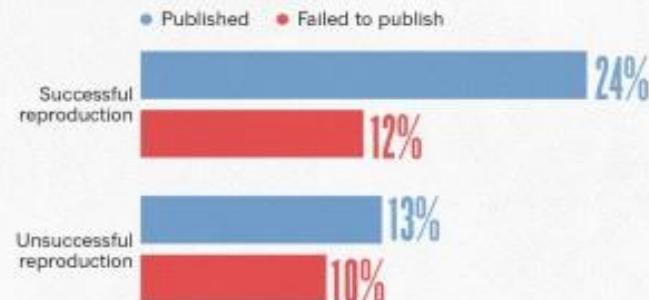
HAVE YOU FAILED TO REPRODUCE AN EXPERIMENT?

Most scientists have experienced failure to reproduce results.



HAVE YOU EVER TRIED TO PUBLISH A REPRODUCTION ATTEMPT?

Although only a small proportion of respondents tried to publish replication attempts, many had their papers accepted.



Number of respondents from each discipline:
Biology 703, Chemistry 106, Earth and environmental 95,
Medicine 203, Physics and engineering 236, Other 233

enature



**Barcelona
Supercomputing
Center**
Centro Nacional de Supercomputación

Planificación de la gestión de datos de investigación

IACS

11 de diciembre de 2024

Nadia Tonello
BSC Head of Data Management
nadia.tonello@bsc.es

Motivación

Que significa gestionar datos de investigación?

- Facilitar la coordinación entre los que participan en el Proyecto, para que los datos se gestionen de manera uniforme y coherente.
- Facilitar la comprensión de los datos del Proyecto (documentación, metadatos)
- Definir responsabilidades
- Regular los derechos de acceso a los datos

El equipo de gestión de datos



- El equipo de gestión de datos tiene que tener herramientas y conocimientos para ayudar los actores en el ciclo de vida de los datos científicos para su gestión óptimal.



Definiciones

RD - Datos de investigación

Fuentes o productos de las actividades de investigación

Ejemplos: Medidas, valores de laboratorio, ficheros audiovisuales, textos, encuestas, objetos, muestras , simulaciones....

RD Workflow

Pasos que constituyen el procesamiento de datos

- transferencia
- conversión
- transformación, ...

Responsabilidades, software y herramientas, infraestructuras, servicios.

RD Lifecycle

Pasos necesarios para mapear el proceso de un proyecto de investigación.

Colección, procesamiento, análisis, archivo y preservación, acceso and re-utilización de datos.

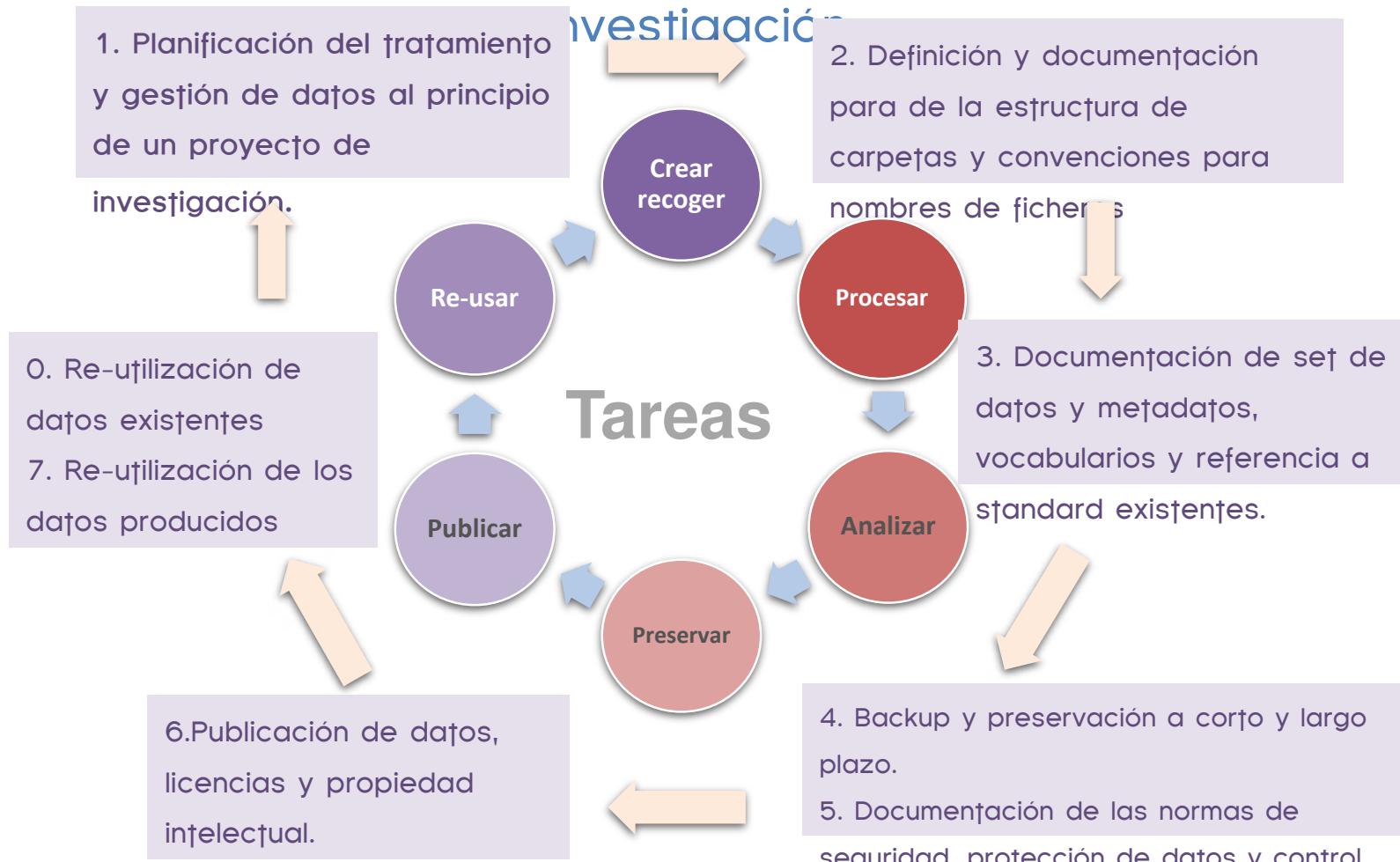


<https://rdmkit.elixir-europe.org/index.html>

Definiciones

Research Data Management

Gestión sistemática y responsable de los datos de investigación.



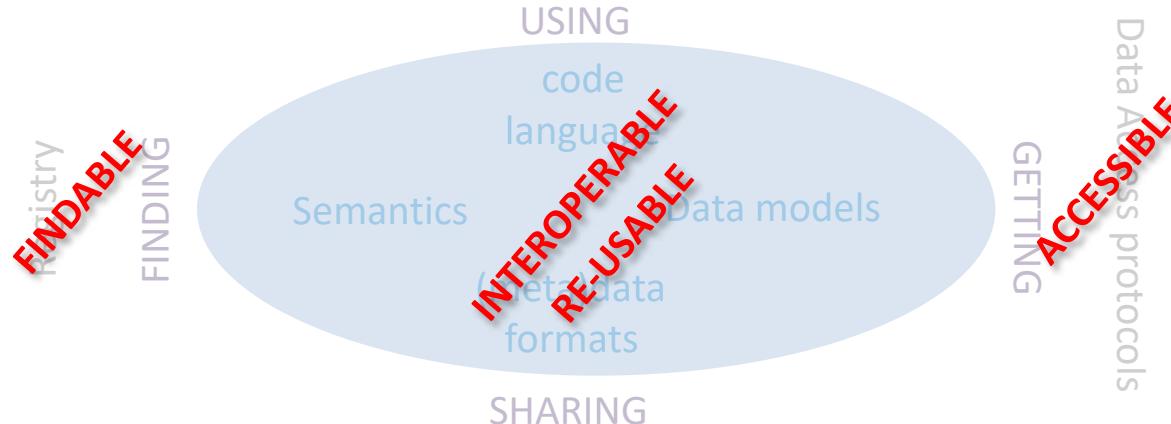
Definiciones Principios FAIR



Browser

Desktop apps

User programs



Resources providers

Adapted from:
G. Greco-IVOA



Barcelona
Supercomputing
Center
Centro Nacional de Supercomputación

Planificación de la gestión de datos de investigación

IACS

11 de diciembre de 2024

Nadia Tonello
BSC Head of Data Management
nadia.tonello@bsc.es

Definiciones Principios FAIR

THE FAIR DATA PRINCIPLES

To be Findable	
F1	(meta)data are assigned a globally unique and eternally persistent identifier
F2	data are described with rich metadata
F3	metadata clearly and explicitly include the identifier of the data they describe
F4	(meta)data are registered or indexed in a searchable resource
To be Accessible	
A1	(meta)data are retrievable by their identifier using a standardised communications protocol
A1.1	the protocol is open, free, and universally implementable
A1.2	the protocol allows for an authentication and authorisation procedure, where necessary
A2	metadata are accessible, even when the data are no longer available

GoFAIR - principles definition

To be Interoperable	
I1	(meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation
I2	(meta)data use vocabularies that follow FAIR principles
I3	(meta)data include qualified references to other (meta)data
To be Re-usable	
R1	meta(data) are richly described with a plurality of accurate and relevant attributes
R1.1	(meta)data are released with a clear and accessible data usage license
R1.2	(meta)data are associated with detailed provenance
R1.3	(meta)data meet domain-relevant community standards

[CESSDA: FAIR data](#)



**Barcelona
Supercomputing
Center**
Centro Nacional de Supercomputación

Planificación de la gestión de datos de investigación

IACS

11 de diciembre de 2024

Nadia Tonello
BSC Head of Data Management
nadia.tonello@bsc.es

Definiciones

Ética e integridad	Rigor metodológico para salvaguardar contra las fuentes de sesgo, demostrando seguir estándares morales de conducta, incluyendo comportamientos como compartir datos cuanto antes, construir sobre trabajos publicados por terceros, y sometiéndose a validación externa.
EOSC - European Open Science Cloud	Asociación y programa de trabajo cuyo objetivo es crear una federación de las infraestructuras de datos de investigación existentes en Europa y realizar una red de datos y de servicios relacionados para la ciencia, según los principios FAIR.
Horizon Europe	Programa marco de investigación e innovación de la Unión Europea para el período 2021-2027. Incluye las prácticas de ciencia abierta como forma de excelencia, e introduce obligaciones específicas sobre acceso abierto a publicaciones y de gestión de datos.

Open Data

Todos los datos de investigación se tienen que publicar en abierto a menos que existan restricciones legales (GDPR, PI), o éticas de algún tipo.

Datos confidenciales	Mails internos, contratos, documentos financieros de una institución...
Datos personales	Información que directamente o indirectamente identifica o puede identificar una persona viva. Ejemplos: nombre y apellidos domicilio dirección de correo electrónico número de documento nacional de identidad, o cualquier otro numero de identificación personal datos de localización (como la función de los datos de localización de un teléfono móvil) dirección de protocolo de internet (IP), los datos en poder de un hospital o médico, que podrían ser un símbolo que identificara de forma única a una persona (datos físicos, psicológicos, genéticos, ...)

Reglamento general de protección de datos (RGPD)

General Data Protection Regulation

Protege a las personas cuando sus datos están siendo tratados

Permite que las personas controlen mejor sus datos personales

Sujeto

Persona identificable.

Sus datos se pueden usar solo por propósitos legítimos explícitos y bajo consentimiento explícito.

Tiene el derecho de retirar el consentimiento al uso de sus datos personales para los propósitos explicitados en cualquier momento.



Controlador

Persona física o jurídica que determina el propósito de los datos personales y custodia el consentimiento explícito del sujeto para el uso de sus datos para el propósito especificado.

Tiene que ser capaz de demostrar el cumplimiento de la GDPR en cada momento.

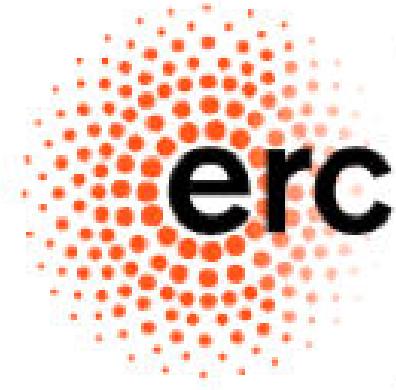


Procesador

Persona física o jurídica que trata los datos por parte del controlador.



DMP para acceder a fondos de investigación públicos



European
Research
Council





DMP para proyectos europeos

Open access to publications and research data in Horizon 2020

- ACCESO ABIERTO a publicaciones revisadas por pares **OBLIGATORIO desde el H2020**
- PUBLICACIONES DE DATOS EN ABIERTO empezó con un piloto flexible, facultativo con el H2020 .
- Perspectiva de la Ciencia Abierta
- EOSC - European Open Science Cloud: web of FAIR services and data



Horizon Europe
ACCESO ABIERTO
PUBLICACIONES DATOS
Plan de Gestión de Datos

Pilares del Open DM según el H2020

Depositar los datos en
un repositorio de datos de
investigación



Asegurar que los
datos se puedan
acceder, explotar,
reproducir, diseminar...



Reporte periódico
Tablas de publicaciones,
entregables,
procedimiento de reporte continuo
1) Inicial 2) Detallado 3) Final



DMP - Plan de gestión de datos

Documento que describe las estrategias de gestión de datos para recolectar, procesar, archivar, preservar y publicar datos de investigación.

INCENTIVOS

- Identificar potenciales problemas de gestión de datos y encontrar soluciones
- Evitar duplicación o perdida de datos, y fallos de seguridad
- Cumplir con los requerimientos
- Futura Reforma del sistema de evaluación de los trabajos de investigación

IMPACTO

- Tiempo
- Decisiones sobre procesos y datos producidos
- Descripción de los datos
- Interoperabilidad - investigación multidisciplinaria
- Difusión y reutilización de los resultados de investigación
- Automatización de los procesos
- Escalabilidad del análisis para grandes volúmenes de datos.

Como se planifica la gestión de datos de un proyecto

Plantillas
Contenido

Por donde empezar?

Documento que describe las estrategias de gestión de datos para recolectar, procesar, archivar, preservar y publicar datos de investigación.

La buena gestión de los datos es necesario para

- compartir y reutilizar los datos científicos
- garantizar la calidad de los resultados
- Cumplir con los requerido por las agencias financieradoras

La gestión de datos y su planificación:

- Cubre todas las fases de un proyecto de investigación
- Se tiene que actualizar periódicamente
- Se extiende más allá de la duración de un proyecto de investigación

Hay diferentes plantillas y herramientas para escribir un DMP, que se adaptan a diferentes necesidades.



Plantillas para DMPs

FAIR Data Management Horizon 2020 DMP

1. Data summary

- Description,
- origin
- total size
- formats
- usefulness...

2. FAIR data:

- 2.1 Making data findable (metadata)
- 2.2 Making data openly accessible
- 2.3 Making data interoperable
- 2.4 Increase data re-use (licenses)

3. Allocation of resources

4. Data security

5. Ethical aspects

6. Other relevant aspects

Project Details Plan overview Write Plan Share Download



DCC Template

This plan is based on the "DCC Template" template provided by Digital Curation Centre.

The default DCC template

Template version 1, published on 08 March 2018

Instructions

The DCC default template

Write plan

Data Collection

- What data will you collect or create?
- How will the data be collected or created?

Documentation and Metadata

- What documentation and metadata will accompany the data?

Ethics and Legal Compliance

- How will you manage any ethical issues?
- How will you manage copyright and Intellectual Property Rights (IPR) issues?

Storage and Backup

- How will the data be stored and backed up during the research?
- How will you manage access and security?

Selection and Preservation

- Which data are of long-term value and should be retained, shared, and/or preserved?
- What is the long-term preservation plan for the dataset?

Data Sharing

- How will you share the data?
- Are any restrictions on data sharing required?

Responsibilities and Resources

- Who will be responsible for data management?
- What resources will you require to deliver your plan?

Plantillas para DMPs

[Discipline specific data management](#) by Science Europe

Proofs of Concept from different Communities

- Humanities (general): DARIAH
- Humanities – Archaeology: PARTHENOS/ARIADNE
- Linguistics – Language Data: CLARIN
- Social Sciences – Survey data: CESSDA
- Social/Behavioural Sciences – Psychology
- Social Sciences – Family of Studies on Longitudinal Ageing
- Life Sciences – Bio-informatics: ELIXIR and Force11/RDA FAIRSharing
- Plant Sciences: ERA-CAPS
- Climate Research: ICOS

DMP tool	dmptool.org/
DMP online	https://dmponline.dcc.ac.uk/
Easy DMP documentation	http://www.sigma2.no/easydmp-user-
DSWizard	https://ds-wizard.org/
eINA DMP	https://dmp.csuc.cat/
ARGOS	https://argos.openaire.eu



Summary Report

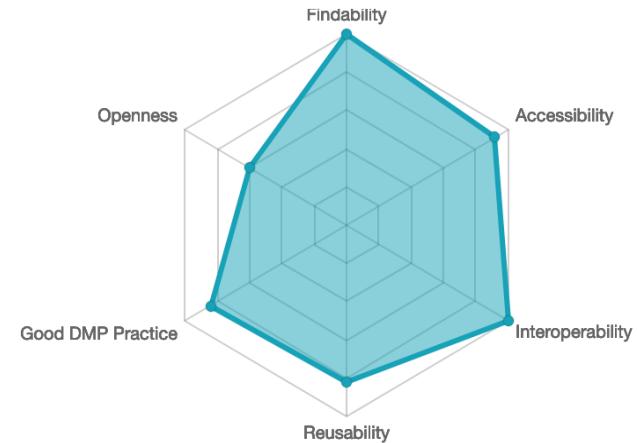
Answered (current phase): 130/163



Answered: 167/218



Metric	Measure	
Findability	1.00	
Accessibility	0.92	
Interoperability	1.00	
Reusability	0.82	
Good DMP Practice	0.84	
Openness	0.60	



Contenido de un DMP

FAIR Data Management at a glance: issues to cover in your Horizon 2020 DMP

1. Data summary

2. FAIR data:

- 2.1 Making data findable, including provisions for metadata
- 2.2 Making data openly accessible
- 2.3 Making data interoperable
- 2.4 Increase data re-use (through clarifying licenses)

3. Allocation of resources

4. Data security

5. Ethical aspects

6. Other relevant aspects

1. Data Summary

Issues to be addressed

- State the purpose of the data collection/generation
- Explain the relation to the objectives of the project
- Specify the types and formats of data generated/collected
- Specify if existing data is being re-used (if any)
- Specify the origin of the data
- State the expected size of the data (if known)
- Outline the data utility: to whom will it be useful

- A quien está dirigido el DMP?
- Que datasets podrían entrar en este DMP?
- Que preguntas se pueden contestar ANTES de que empiece el proyecto?
- Quien tiene la información para escribir esta parte del DMP?
Financiadores, colaboradores , técnicos, ...

Contenido de un DMP

FAIR Data Management at a glance: issues to cover in your Horizon 2020 DMP

1. Data summary

2. FAIR data:

- 2.1 Making data findable, including provisions for metadata
- 2.2 Making data openly accessible
- 2.3 Making data interoperable
- 2.4 Increase data re-use (through clarifying licenses)

3. Allocation of resources

4. Data security

5. Ethical aspects

6. Other relevant aspects

Plataforma de gestión de datos de investigación

IACS

11 de diciembre de 2024

Principios FAIR en la práctica



Australian Research Data Commons (CC-BY)
<https://ardc.edu.au/resource-hub/making-data-fair/>



Australian Research Data Commons

Findability

Findable

To aid automatic discovery of relevant datasets, (meta)data should be easy to find by both humans and machines and be assigned a persistent identifier.

- Outline the discoverability of data (**metadata provision**)
- Outline the identifiability of data and refer to standard identification mechanism. Do you make use of persistent and unique identifiers such as **Digital Object Identifiers?**
- Outline **naming conventions** used
- Outline the approach towards search **keyword**
- Outline the approach for clear **versioning**
- Specify standards for metadata creation (if any). If there are no standards in your discipline describe what type of metadata will be created and how

Metadata

- Describen los datos y sus características
- Permiten la búsqueda automática
- Contienen la información relevante, con palabras clave significativas

[DCC standards metadata](#)



**Barcelona
Supercomputing
Center**
Centro Nacional de Supercomputación

Planificación de la gestión de datos de investigación

IACS

11 de diciembre de 2024

Nadia Tonello
BSC Head of Data Management
nadia.tonello@bsc.es

Findability Identificadores

- Persistentes: dirección permanente a los datos
 - Fiables: Identificación y localización de los datos
 - Combinables: para darle más valor.
-
- | | |
|-----------------|-------------|
| • Resources | doi.org |
| • People | ORCID.org |
| • Organisations | ROR.org |
| • Projects | RAID.org.au |

McMurry et al Identifiers for the 21st century: How to design, provision, and reuse persistent identifiers to maximize utility and impact of life science data, PloS Biology, (2017),
<https://doi.org/10.1371/journal.pbio.2001414>

Accessibility Repositories

Accesible

Limitations on the use of data, and protocols for querying or copying data are made explicit for both humans and machines.

- Specify which data will be made **openly available**? If some data is **kept closed** provide rationale for doing so
- Specify how the data will be made available
- Specify what methods or **software tools** are needed to access the data? Is documentation about the software needed to access the data included? Is it possible to include the relevant software (e.g. in open source code)?
- Specify where the data and associated metadata, documentation and code are deposited
- Specify how access will be provided in case there are any **restrictions**

- Donde se almacenan datos y metadatos, para que sean accesibles?
- Elección del repositorio
- Método de acceso

- Por cada objeto digital
- Index y buscador automático
 - Accesible con protocolos standard
 - Licencias asociadas

Accessibility Repositorios

Certificaciones

ISO 16363 - Transferencia de información y datos especiales

ISO 14721 - OAIS Conservación de datos a largo plazo

ISO 21001 - ENS Esquema Nacional de Seguridad

Genéricos

DataVERSE, Mendeley, Figshare, OSF, Zenodo

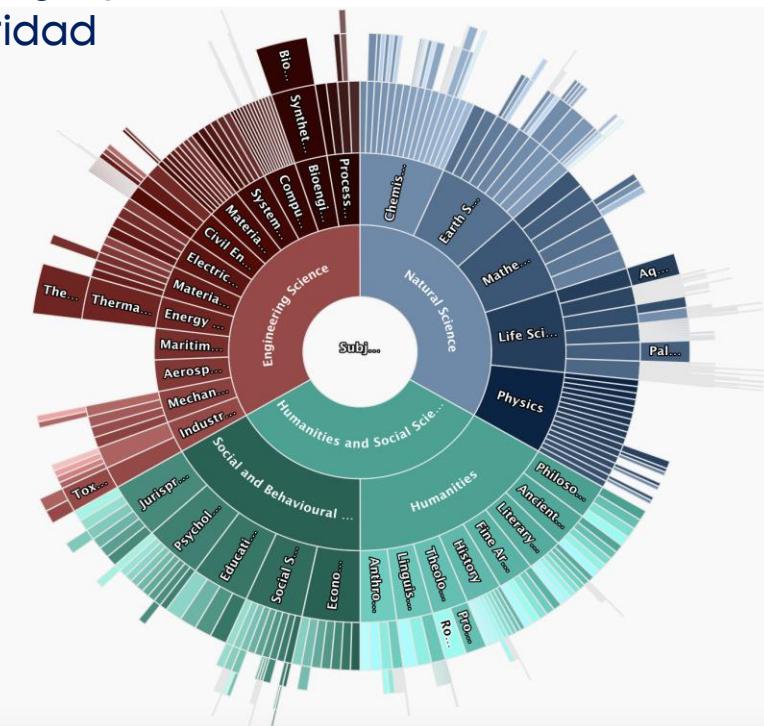
Catálogos: FAIRsharing, R3data, RRID

Estudios y comparaciones disponibles

(FairSharing)

Disciplinarios

<https://fairsharing.org/browse/subject>



Accessibility Repositorios

Artículos

Green open access (self-archiving)

El autor publica en un repositorio online el artículo que ha sido aceptado para su publicación.

Gold open access (open access)

El editor es responsable de la exposición del artículo publicado de forma abierta.

Registry of research data repositories:
re3data.org

Datos

- Generic repository: Zenodo and similar
- Life Sciences: Europe PubMed Central.
- Physical Sciences and Engineering

Requisitos:

- Identificadores persistentes, Digital Objects Identifiers
- Metadata accesibles, también después de que los datos ya no estén disponibles.
- Licencias: si los datos no son abiertos o no se pueden usar libremente, la limitación tiene que ser clara y justificada.

Interoperability

Vocabularios controlados

I
nteroperable

(Meta)data should use standardised terms (controlled vocabularies), have references to other (meta)data and be machine actionable.

- Assess the interoperability of your data. Specify what data and metadata **vocabularies**, **standards** or methodologies you will follow to facilitate interoperability.
- Specify whether you will be using standard vocabulary for all data types present in your data set, to allow inter-disciplinary interoperability? If not, will you provide mapping to more commonly used ontologies?

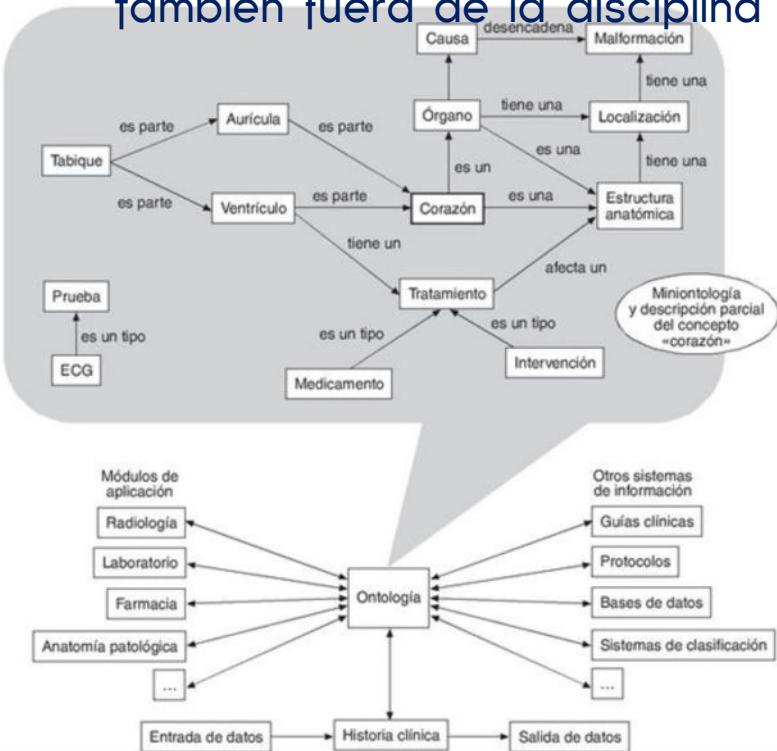
- Existe un vocabulario estandar en la disciplina?
- Hay referencias al vocabulario usado?
- Semántica (forma de implementación de una ontología) forma específica de representar información a través de relaciones entre términos



Interoperability

Vocabularios controlados

El vocabulario usado en documentación tiene que ser standard, formal, accesible, compartido y entendible también fuera de la disciplina



<https://guides.lib.unc.edu/metadata/controlled-vocab>

Health Sciences & Medicine

- [International Classification of Disease \(ICD\)](#)
- [Logical Observation Identifiers Names, and Codes \(LOINC\)](#) (for tests, observations and measurements)

- [Medical Subject Headings \(MeSH\)](#)

- [Rx Norm](#) (for drugs)

- [SNOMED-CT](#) (for clinical terms)

Sciences

- [Astronomy Thesaurus](#)

- [NASA Thesaurus](#)

- [National Agricultural Library Thesaurus](#)

- [The Getty Institute Thesaurus of Geographical Names](#)

Social Sciences

- [Ethnographic Thesaurus](#)

- [Thesaurus for Economics](#)

R Reusable

(Meta)data are sufficiently well described for both humans and computers to be able to understand them and have a clear and accessible data usage license.

Reusability Licencias

- Specify how the data will be licensed to permit the widest re-use possible.
- If an embargo is sought to give time to publish or seek patents, specify why and how long this will apply, bearing in mind that research data should be made available as soon as possible.
- If the re-use of some data is restricted, explain why.
- Specify how long is it intended that the data remains reusable.
- Describe the process for data quality assurance.

- Los datos y los productos tienen licencias tipo Creative Commons, o similares
- Datos personales (GDPR) o confidenciales necesitan un tratamiento especial

Reusability Licencias

Allowed licenses for datasets					
CC	Description	Modification of the content	Commercial Use	Free cultural works	Open definition
CC0	Free content, no restrictions (public domain)	yes	yes	yes	yes
CC-BY	Attribution	yes	yes	yes	yes
CC-BY-SA	Attribution+ShareAlike	yes	yes	yes	yes
CC-BY-NC	Attribution+NonCommercial	yes	no	no	no
CC-BY-ND	Attribution+NoDerivatives	no	yes	no	no
CC-BY-NC-SA		yes	no	no	no
CC-BY-NC-ND		no	no	no	no

Category	Permissive	Weak Copyleft	Strong Copyleft
Software	MIT, Apache 2.0	MPL 2.0, EPL	GPL v3, AGPL v3
Hardware	Solderpad, CERN OHL-P	CERN OHL-W	CERN OHL-S
LLMs	MIT, Apache 2.0	MPL 2.0, LGPL	GPL v3, AGPL v3

Las licencias GPL pueden ser virales para el desarrollo de productos derivados.

Ejemplo: [B2SHARE license selector](#)



Contenido de un DMP

FAIR Data Management at a glance: issues to cover in your Horizon 2020 DMP

1. Data summary
2. 2. FAIR data:
 3. 2.1 Making data findable, including provisions for metadata
 4. 2.2 Making data openly accessible
 5. 2.3 Making data interoperable
 6. 2.4 Increase data re-use (through clarifying licenses)
3. Allocation of resources
4. Data security
5. Ethical aspects
6. Other relevant aspects

Estimación de recursos

Los costes de la gestión de datos tienen que ser evaluados para poder incluirlos en los gastos del Proyecto.

Identificar todos los procesos involucrados en la gestión de datos:

- Herramientas del centro (software)
- Infraestructuras del centro (storage)
- Costes de acceso a datos externos
- Coste de uso de herramientas externas
- Personal experto
- ...

<https://www.ukdataservice.ac.uk/media/622368/costingtool.pdf>
<https://www.uu.nl/en/research/research-data-management/guides/costs-of-data-management>

Contenido de un DMP

FAIR Data Management at a glance: issues to cover in your Horizon 2020 DMP

1. Data summary
2. 2. FAIR data:
 3. 2.1 Making data findable, including provisions for metadata
 4. 2.2 Making data openly accessible
 5. 2.3 Making data interoperable
 6. 2.4 Increase data re-use (through clarifying licenses)
3. Allocation of resources
4. Data security
5. Ethical aspects
6. Other relevant aspects

Seguridad de la gestión de datos sensibles

Impacto sobre la investigación

- Establecer el nivel de seguridad necesario para los datos
 - Análisis de riesgos: que amenazas hay, que probabilidad hay que se materialicen, que impacto pueden tener, que tiene sentido hacer para mitigar el riesgo.
- Respeto de los procesos que garantizan la seguridad
- Evaluación y aplicación de las herramientas para proteger los datos sensibles (cifrado, codificación, encriptación, anonimización, pseudonimización).

Clasificación de los datos

Caso genérico

PUBLICO

- No confidencial, de acceso público
- web, redes sociales, documentación...
- Ningún daño para personas o instituciones.

Uso interno

- Confidencialidad BAJA
- Mail interno, políticas internas, ..
- Daños: mala publicidad, desventaja sobre la concurrencia

Confidencial

- Confidencialidad MEDIA
- Contratos, datos personales, ...
- Puede causar pérdidas financieras o daños a personas

Restringido

- Confidencialidad ALTA
- Propiedad intelectual, secretos comerciales, ...
- Daños permanentes a personas o instituciones

Política de protección
Vale la pena la inversión?

METODOS

Manual
(documento s)

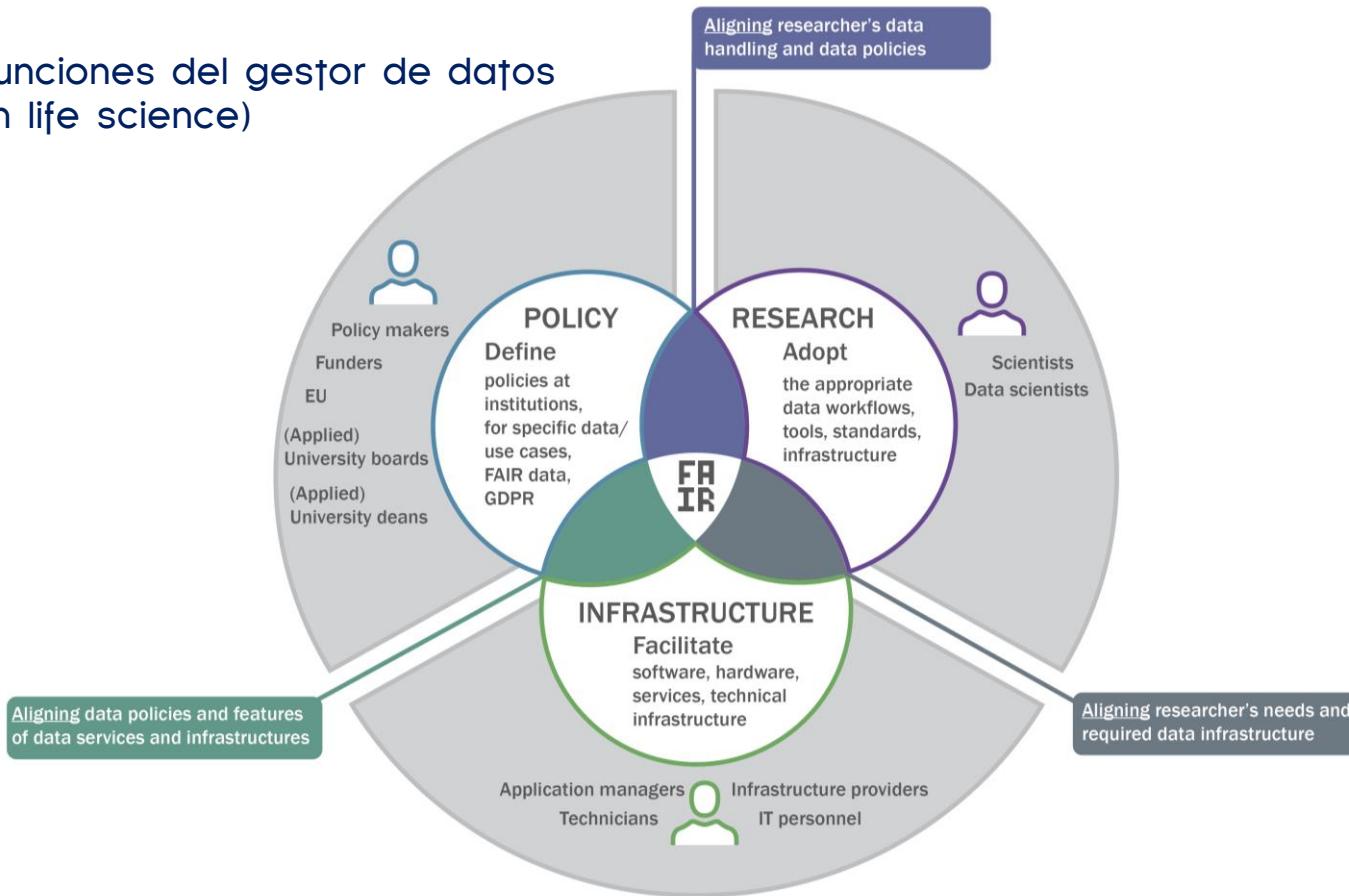
Manual con herramientas
(plantillas)

Herramientas automáticas (DLP, scan,...)

RESPONSABILIDADES

Implementation areas for data stewardship

Funciones del gestor de datos
(in life science)



<https://doi.org/10.5281/zenodo.3460552>, All project output

[Final report](#) (Oct 3, 2019)

Funciones y responsabilidades

A. Responsable de políticas y procedimientos

Define e implementa:

- políticas institucionales
- políticas y procedimientos específicos
- FAIR data y GDPR

B. Responsable de datos y relación con los científicos

Adopta e implementa en base al proyecto:

- Workflows en acorde con las políticas
- Elije herramientas, standards e infraestructuras

POLITICAS

INVESTIGACIÓN

INFRAESTRUCTURA

Data consultant

traduce políticas institucionales y específicas en requerimientos y servicios

C. Responsable de servicios e infraestructuras

Facilita e implementa:

- soluciones software y hardware
- Evalua infraestructuras IT

Data manager

traduce necesidades de los investigadores en requerimientos y servicios

Christine Staiger, DTL

Política de seguridad

Directrices que rigen la forma de gestión y protección de la información y de los servicios.

Documentos electrónicos y físicos de los que han derivado.

- Objetivos/misión de la organización
- Marco regulatorio
- Roles/funciones de seguridad
- Estructura y composición del Comité de Seguridad
- Documentación de Seguridad
- Riesgos derivados del tratamiento de datos personales.

Políticas de RDM

Principio y requerimientos de una institución sobre la gestión de los datos de investigación (creación, colección, curación, publicación)

Guía para los investigadores sobre cómo elegir:

- Repositorios de datos
- Licencias de Software y datos, y su publicación
- Principios de rigor y calidad de la investigación

Ejemplos:

- Datos de investigación deberían ser publicados en abierto...
- Los mecanismos de gestión de datos deben ser eficientes y efectivos...
- Datos de reconocido valor a largo plazo deberían ser preservados, accesibles y reutilizables...

Otras políticas a tener en cuenta

Instituciones	Universidades e institutos de investigación
Revistas	Guías para compartir datos, citaciones y accesibilidad
Proyectos	Standards y regulación interna
Disciplinas	Formatos, índices o gestión específica
Agencias de financiación	Requisitos y sistemas para proveer información sobre la gestión de datos: DMPs

3 Como evaluar la gestión de datos de un proyecto

Resumen de
contenidos
Check list

Resumen

Un plan de gestión de datos es necesario para

- compartir y reutilizar los datos científicos
- garantizar la calidad de los resultados

El data manager da soporte a los investigadores, los técnicos y los responsables de establecer las políticas de gestión de datos para escribir y ejecutar un DMP.

Un DPM:

- Cubre todas las fases de un proyecto de investigación
- Se tiene que actualizar periódicamente
- Se extiende más allá de la duración de un proyecto de investigación

DMP cheatsheet
→

Principios FAIR

Findability	Metadata + PIDs
Accessibility	Repositories
Interoperability	Vocabularies
Reusability	Licenses

Evaluación de un DMP

El DMP puede ser:

- un entregable para la EC evaluado por un experto externo
- un documento interno detallado para los colaboradores
- documentación para investigadores ajenos al proyectos que quieren reutilizar los datos
- ...

Según el destinatario, el DMP puede ser más o menos detallado.

Pueden ser documentos diferentes.

Reference: [*Rubric for the Evaluation of Research Data Management Plans*](#) by Science Europe

Evaluación de un DMP

Check list 2/2

1. Objetivo científico para el almacenamiento de los datos.
2. Plan técnico de gestión: volumen y tipo de datos, con planificación mínima a 3 años.
3. Definición del nivel de interoperabilidad (vocabularios controlados).
4. Uso de formatos estándar, públicamente accesibles.
5. Descripción completa y accesible (documentación de apoyo al proyecto, metadatos, modelo de datos) para que los datos sean comprensibles para otros grupos.

Reference: *Rubric for the Evaluation of Research Data Management Plans* by Science Europe

Evaluación de un DMP

Check list 2/2

6. Permisos sobre el acceso a los datos (públicos, privados, embargados o restringidos, con definición de la licencia correspondiente).
7. Plataformas y servicios necesarios para poder acceder a los datos.
8. Duración del proyecto y estimación del tiempo de vida útil de los datos, más allá del proyecto (incluye el plan de finalización del proyecto)
9. Responsables de los datos.
10. Calendario de actualización de los datos y del plan.

Reference: [*Rubric for the Evaluation of Research Data Management Plans*](#) by Science Europe

Q&A

Plataformas de training

Ref: <https://marketplace.eosc-portal.eu/services/c/training-support>

Ejemplos:

- [Data Management Expert Guide \(DMEG\)](#) Social Sciences
- [E-Learning Platform of GBIF Spain](#) Biodiversity
- [EGI FitSM Training](#) IT services management standard
- [Galaxy training network](#) Computational biology
- [FAIRsharing](#) Generic, ELIXIR
- [Pan-Learning.org](#) Photon and neutrón sciences
- [PRACE Advanced Training Centres](#) Advanced computing engineering
- [TRANSITS Training](#) Computer security, GEANT
- [Open Science MOOC](#) Generic online course
- [Open Science Training](#) Practice Open and FAIR Science, OpenAire
- [RDMkit](#)
- [FAIR cookbook](#)



Gracias!

Nadia Tonello
BSC Head of Data
Management
nadia.tonello@bsc.es