

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

5-8 de junio de 2023

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

Tutoría 13 de junio a las 12 h

Remedios Melero. IATA-CSIC, Valencia, España

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

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

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

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

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

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

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

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

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

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

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

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

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)

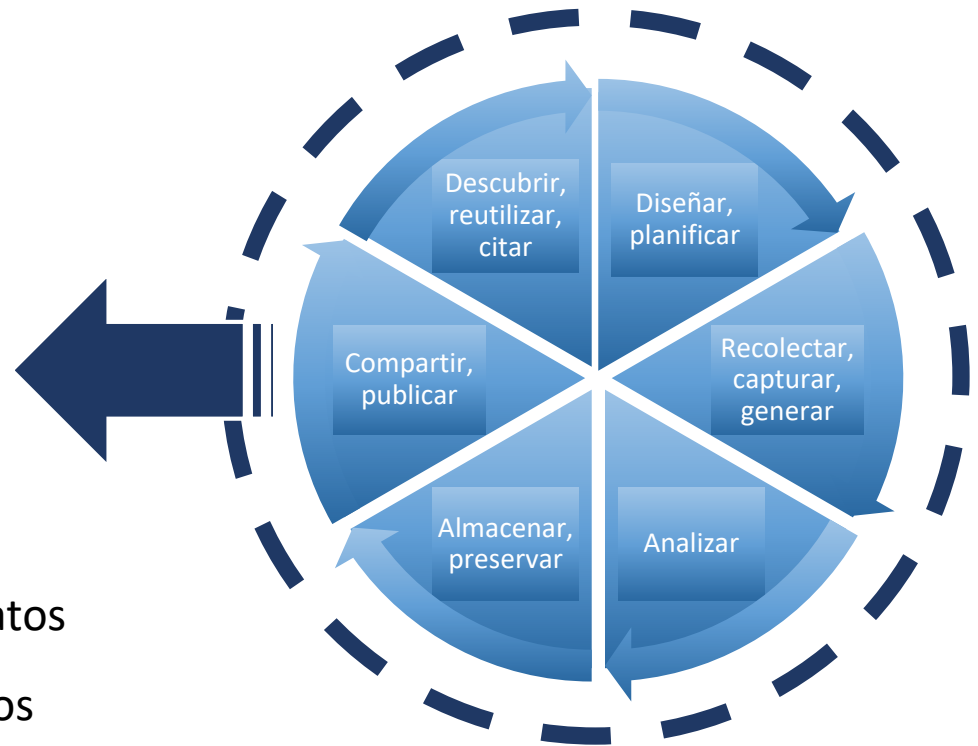
Una buena gestión de los datos ayuda a garantizar que los investigadores compartan sus datos de forma FAIR (localizables, accesibles, interoperables y reutilizables).

- 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.

¿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

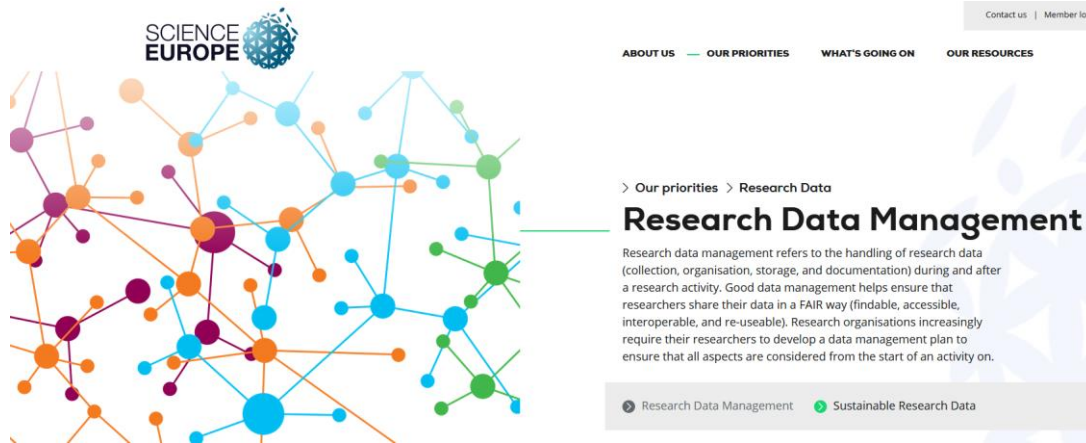
Los componentes comunes de un DMP son:

- 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 los datos y a garantizar que los datos sean FAIR (Localizables, Accesibles, Interoperables, Reutilizables)?



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.consortiomadrono.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.



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	
DMP Planner	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 web	
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	

INVESTIGAM -Portal de Ciencia Abierta



InvestigaM, el portal de ciencia abierta del [Consortio 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 los investigadores de las universidades públicas de la Comunidad de Madrid y la UNED, miembros del Consorcio 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?](#)



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.

<https://dmptool.org/>

[Learn](#) [Sign in](#) [Language](#)



Build your Data Management Plan

✔ Notice: You will receive an email with instructions to reset your password in a few minutes.

Welcome to the DMPTool

Create data management plans that meet institutional and funder requirements.

Get started

DMP ONLINE

Home

Public DMPs

Funder requirements

Help

Plan to make data work for you

Data Management Plans that meet institutional funder requirements.



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

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

Plan and follow your data

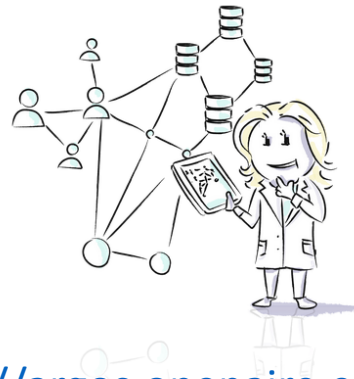
Create machine actionable DMPs.

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

< Anterior

> Siguiente

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

Acceso por EUDAT, hay que identificarse/registrarse

Choose a template

Show entries

Search:

Template	Version	Description	
Horizon 2020	1	Simplified template based on Horizon 2020 guidelines.	Use
Horizon 2020 Expert	1	A shorter template based on Horizon 2020 that assumes knowledge of data management.	Use
NINA data management plan	1	Template for the Norwegian Institute for Nature Research	Use
Science Europe	2	Template for data management plans based on the Science Europe guidelines.	Use
Sigma2 Data Management Plan	1	Data management plan for Sigma2 resource allocation.	Use

Showing 1 to 5 of 5 entries

Previous Next

! Vous souhaitez participer aux évolutions de DMP OPIDoR. Alors rejoignez-nous sur [GitHub](#) !

DMP Templates

DMP templates provided by a funder or research organisations, available on DMP OPIDoR. You can download these templates and related guidances, create a plan from these templates.

Template Name ▾	Organisation Name ▲	Organisation Type ▾	Description	Last Updated ▾	Download
ANR - Modèle de PGD (français)	Agence nationale de la recherche (ANR)	Funder	<p>Dans le cadre de sa politique science ouverte, et en lien avec le Plan national pour la science ouverte, l'Agence nationale de la recherche (ANR) demande l'élaboration d'un Plan de Gestion des Données (PGD) pour les projets financés à partir de 2019. Elle participe ainsi à l'alignement européen et international en faveur de la structuration et de l'ouverture des données de la recherche. Le principe « aussi ouvert que possible aussi fermé que nécessaire » sera au cœur de sa démarche.</p> <p>Soucieuse de développer une approche concertée, l'ANR a suivi les recommandations du Comité pour la Science Ouverte (CoSO) qu'elle a sollicité et a adopté le modèle de PGD proposé par Science Europe visant une harmonisation internationale de la gestion des données. La démarche de l'ANR est d'accompagner et de favoriser l'adaptation progressive des communautés à l'évolution des pratiques sur la question des données. Par ailleurs ce modèle s'adresse à l'ensemble des bénéficiaires de l'ANR dans le respect de leurs spécificités disciplinaires.</p> <p>Ce plan de gestion des données constitue un livrable qui devra être fourni dans les 6 mois qui suivent le démarrage scientifique du projet. Il sera mis à jour au cours de la vie du projet selon les modalités</p>	14-12-2020	

https://dmp.opidor.fr/public_templates

About the DMP Canvas Generator

This is a tool aiming to help scientists generate Data Management Plans for SNSF funded projects. The produced Word document is compliant with the [SNSF instructions for DMP creation](#) and consists of generic paragraphs corresponding to the user's inputs. The produced document structure follows the structure of the [SNSF DMP questionnaire](#). The document must be further modified before submission to reflect the specific aspects of the project. Vital-IT is not responsible for the final submitted version of the document.

Project name and description

Enter project name

Please provide a small description for your project: (optional for SNSF)

Text area for project description with a small icon in the bottom right corner.

A. Data that will be collected / generated

- | | |
|---|--|
| <input type="checkbox"/> Genomic data | <input type="checkbox"/> Proteomic data |
| <input type="checkbox"/> Transcriptomic data | <input type="checkbox"/> Lipidomic data |
| <input type="checkbox"/> Metabolomic data | <input type="checkbox"/> Microscopy images |
| <input type="checkbox"/> Molecular Biology data | <input type="checkbox"/> Flow cytometry data |
| <input type="checkbox"/> Other... | |

B. Analysis that will be generated

- | | |
|--|--|
| <input type="checkbox"/> R | <input type="checkbox"/> MATLAB |
| <input type="checkbox"/> Spreadsheet files | <input type="checkbox"/> Custom script |
| <input type="checkbox"/> Other... | |

C. Questions on data collection

1. Will there be any use of human subjects or samples?
2. Will there be any use of vertebrate subjects or samples?

- 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


Belmont Forum e-I&DM


Home DDOMP Guide Data Management Training Best Practices & Standards GitHub

E-INFRASTRUCTURES & DATA MANAGEMENT TOOLKIT

Providing training and educational resources for data discovery, management, and curation across the globe, in support of an international collaborative effort to enable open access to scientific data

 **DDOMP Researcher Guide**
Resources and tips for creating a successful DDOMP

 **Data Management Training**
Webinars, courses, certifications, and literature

 **Best Practices & Standards**
Guidelines for effective data management

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

RDM Starter Kit

Home > Resources > RDM Starter Kit

> GO FAIR Materials

This page is supposed to serve as a Starter Kit for research resources designed to help researchers get started to

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



Because good research needs good data

- About
- News
- Events
- Services
- Guidance
 - Briefing Papers
 - How-to Guides
 - Case Studies
 - Policy Analysis
 - Metadata
 - Curation Lifecycle Model
 - Data Management Plans
- Research
- Publications
- Change cookie settings

Home > [Guidance](#)

How-to Guides

Our Guides and Checklists are aimed at the those in research data management support roles.

How-to Guides

How-to Guides provide working-level knowledge of curation topics. Each provides background concepts and practical steps aiming to help people in research or support roles implement data management capabilities in their organisation, or better align them with best practices.

- Using RISE, as a capability model for self-assessing RDM services
- Appraise and select research data
- Cite datasets and link to publications
- Develop a data management and sharing plan
- Develop research data management services
- Discover requirements for RDM services
- License research data
- Track the impact of research data with metrics
- Write a lay summary



RESEARCH DATA MANAGEMENT SERVICE GROUP

Comprehensive Data Management Planning & Services

Best practices

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

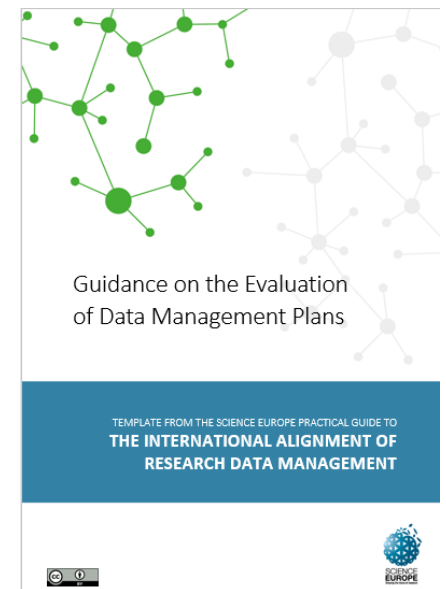
- [Data citation](#)
- [Data storage and backup](#)
- [File formats](#)
- [File management](#)
- [Glossary of data management terms](#)
- [Guide to writing a Data Management Plan \(DMP\)](#)
- [Guide to writing "readme" style metadata](#)
- [Introduction to intellectual property rights in data management](#)
- [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>

Un ejemplo de DMP

Exploring the Psychosocial Needs of Adults with Haematology Cancer under Watch and wait

https://dmponline.dcc.ac.uk/plans/94183/export.pdf?export%5Bquestion_headings%5D=true

<https://www.scienceurope.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¶	<ul style="list-style-type: none"> • Provide information such as name of applicant, project number, funding programme, version of DMP.¶ 	<ul style="list-style-type: none"> • contains the minimal information required to identify the applicant and the references of the project.¶ 	<ul style="list-style-type: none"> • 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?¶	<ul style="list-style-type: none"> • 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.¶ 	<ul style="list-style-type: none"> • 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.¶ 	<ul style="list-style-type: none"> • 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?¶	<ul style="list-style-type: none"> • 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- 	<ul style="list-style-type: none"> • 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- 	<ul style="list-style-type: none"> • 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 and Digital Objects

▶ 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#.YKzHpgHtbt4>

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?				
<p>1.1 Plan lists the <u>types</u> of data and other digital outputs of long-term value.</p> <p><i>(e.g. text, databases, images, 3D models, software, audio files, code, video files, reports, surveys, patient records, samples, and so forth)</i></p>	<p>Datasets and other digital outputs of long-term value are identified, including data type and encoding.</p> <p><i>"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."</i></p> <p><i>"Transcripts and coding will be provided in text files. Audio recording will be MP3 format."</i></p>	<p>Datasets and other digital outputs of long-term value are identified, but lack detail for users beyond the project to understand.</p> <p><i>"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."</i></p> <p><i>"Long-term value data include data from anthropological field studies: transcripts of interviews and discussion workshops, associated metadata."</i></p>	No information about data types is included.	X
<p>1.2 Plan describes how the data and other digital outputs will be <u>collected, captured, or created</u>.</p> <p><i>(e.g., new observations, results from models, reuse of other data, or other)</i></p>	<p>Clearly defines how data will be collected, captured or created, including methods, instruments, software, or infrastructure where relevant.</p> <p><i>"The MIP compatible model runs to the ISI-MIP servers, where existing Fish-MIP runs are already stored. Scenario outputs will be archived."</i></p> <p><i>"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."</i></p>	<p>Missing some details regarding how some of the data will be produced; makes assumptions about reviewer knowledge of methods or practices.</p> <p><i>"Models will produce a broad range of output simulation data."</i></p> <p><i>"Data collection includes gathering in-the-field various phytoplankton, zooplankton, fish and flooded forest biodiversity."</i></p>	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