

Open Science dalla A alla Z

8 – Data Management Plans

Università di Perugia, settembre 2021

<https://doi.org/10.5281/zenodo.5495602>

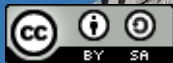
Elena Giglia

Università di Torino

elena.giglia@unito.it



@egiglia



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

In questo modulo impareremo:

1. come scrivere un Data Management Plan

2. Quali strumenti utilizzare

MESSAGGI CHIAVE

- gestire bene i dati è nell'interesse di chi fa ricerca - non c'entra Open!
- NON ci sarà una ricetta per DMP, ma strumenti utili (da imparare)



UN MODO STRUTTURATO
DI PENSARE AI DATI

REGOLE CHIARE=MENO
ERRORI DA SUBITO

UN MODO NUOVO DI PENSARE
ALLA VOSTRA RICERCA, DALLA
PROSPETTIVA DEI DATI

È UN «LIVING DOCUMENT»,
CRESCe COL PROGETTO

È LA SEDE IN CUI
GIUSTIFICATE LE SCELTE
OPEN/CLOSED

...DOVE METTERE TUTTE QUESTE INFORMAZIONI?
NEL DATA MANAGEMENT PLAN

DMP?

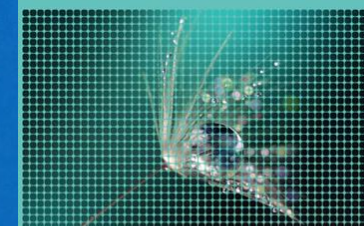
...CHIARIAMO:
IL PROBLEMA NON È
«IMPARARE» A FARE UN DMP
MA IMPARARE A GESTIRE I
DATI IN MODO FAIR E
RESPONSABILE

NEL DMP SEMPLICEMENTE
«DICHIARO» COME TRATTERÒ
I MIEI DATI

NON È UNA FORMALITÀ MA
UNA RESPONSABILITÀ (E UNO
STRUMENTO PREZIOSO)



DMP in crescita



Digital Science Report

The State of Open Data 2020

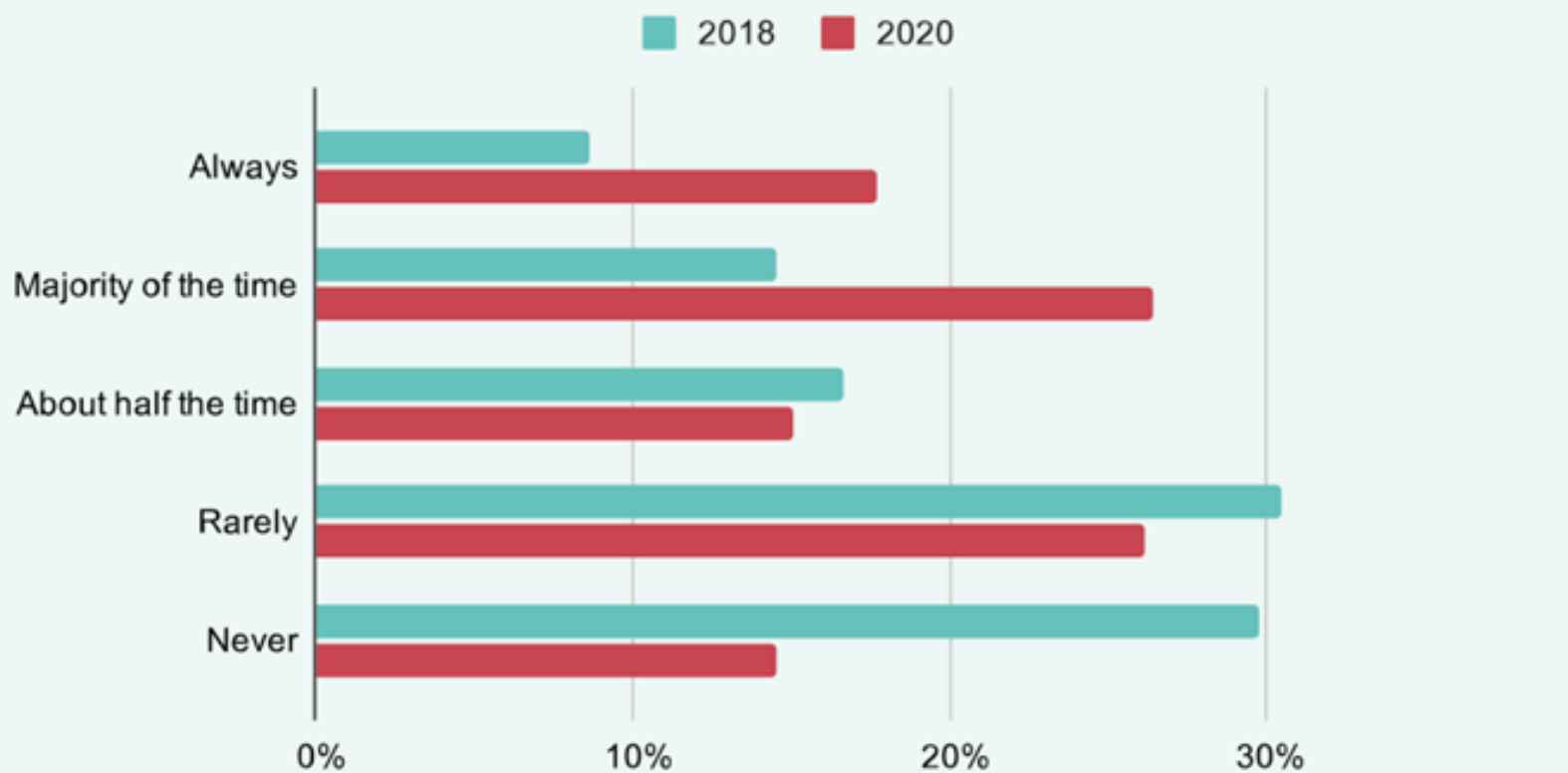
The longest-running longitudinal survey and analysis on open data

Foreword by Dr Louise McClain, CEO of RDS and Executive Director, Essential - Research Data Alliance UK

Dec.2020

December 2020

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



Data Management Plan - caveat



[DMP]

Therefore, it doesn't necessarily matter if you plan to share your data with other scholars, what matters is considering this prospect as you work out how you are going to go about your research. It will help you to understand what it is you are doing more clearly and give you the basis to share that data later on if you so wish.

PORT
postgraduate online
research training
PORT DMP

SCHOOL OF
ADVANCED STUDY
UNIVERSITY
OF LONDON

**NON IMPORTA SE ALLA FINE CONDIVIDERETE I DATI O NO.
QUI SI DOCUMENTANO IL PROCESSO DI RICERCA E LE SCELTE DI METODO**

Data management ABC – Per partire

Ask yourself this:

[DMP]

What is needed to validate the results of your research?

If you were to produce an article researching, for example, the criminal underclass in early-twentieth century New York, what data would you need to include for someone else to replicate your results? Think about it in terms of your own research.

A bibliography would be the most immediate and obvious starting point, revealing to the reader all the sources that you have used to base your research. But what of the gathering mechanisms you used? Did you create a database or undertake statistical analysis? If so you need to make the database and statistics available. This doesn't just mean providing the files in a readable format, but to provide documentation and to make sure that the data is clearly identified with explicit headings, well-structured, and easily identified.

Focusing on what is needed for validation and re-use, rather than the obvious attributes of research data, is useful. It helps you to think through the process of research from a different perspective and what it is you have actually done to come to your conclusions. It also allows you to show the process you have undertaken; revealing how valuable your approach might be and making the

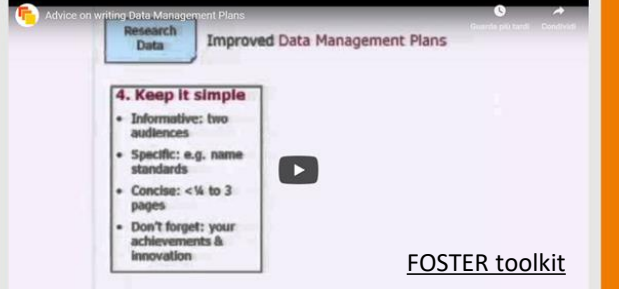
COSA SERVE A VALIDARE
LA MIA RICERCA?
TUTTO QUESTO VA
INSERITO NEL DMP.
PROSPETTIVA DIVERSA
SULLA VOSTRA RICERCA



Trucchi e suggerimenti

Top tip - keep it short and specific!

This very short extract from a presentation by Peter Dukes, Medical Research Council, is really useful advice on writing a DMP from the funding body perspective. The advice applies to all disciplines. The quality of the video isn't great, but the advice is definitely is!



SINTETICO E
SPECIFICO

NON COPIATE

OGNI DATASET È
UNICO, OGNI
INFRASTRUTTURA È
UNICA, OGNI RICERCA
HA LA SUA
IMPOSTAZIONE

ESSERE GENERICI NON SERVE
A NULLA
[we expect a huge size of data;
data will be available]

- USATE TABELLE,
ELENCHI PUNTATI
- SIATE SCHEMATICI E
NON DILUNGATEVI

CIÒ CHE DICHIARATE NEL
DMP POI VA FATTO
VERAMENTE... QUINDI

A) NON FATE GLI
SPLENDIDI

B) NON IMPEGNATEVI A
FARE COSE CHE
SAPETE IMPOSSIBILI

Es. DATI PSEUDONIMIZED, non
ANONIMIZED





10 TIPS FOR WRITING A DATA MANAGEMENT PLAN

1

START EARLY

Read the guidance and ask for advice early on in the process, as writing a DMP may take some time

2

CONSIDER REUSE

Think about reusing existing data. Describe what you will need to know about your data five years from now

3

CHECK POLICIES

Talk to your supervisor or lab members about existing data management policies and standards

4

MAKE USE OF SUPPORT

Use your in-house support services like RDM Support, the Library, IT department or legal desk

5

THINK BROAD

Also address software code, algorithms and any other valuable research assets in your DMP

6

COPY WHERE YOU CAN

Look at other (submitted) plans and copy when appropriate

7

BE UNIQUE WHERE NEEDED

Since every research project is unique, so are the data it generates. Copying from sample DMPs is not sufficient

8

BE CONCRETE

Make your answers as concrete as possible. Show that you have consulted RDM experts

9

SAY SO IF YOU DON'T KNOW

Indicate what you do not yet know and how you will resolve these questions later

10

UPDATE

DMPs add to the planning of your research methods. Therefore define, carry out and update your DMP just as you would any method

Tips&tricks / 2

A data planning process ensures that all aspects of data management are holistically explored at the start of a project. Short-term and long-term aims can be balanced, so that decisions made early in a project do not negatively impact on the ability to find and use the research data in future.

Effective management of data provides researchers with many benefits, including

- time saved through reduced duplication of effort
 - decreased risk of loss, theft or inappropriate use of data
 - good research practice ensures the integrity and quality of data
 - data can be understood and used now and in the future
-
- helps researchers find and gain access to data management – expertise and infrastructure offered at the University
 - increased researcher profile through data dissemination and re-use.

A data planning process is particularly important in the context of collaborative research projects. Researchers may identify areas of potential difficulty or conflict, and these can be resolved with colleagues and collaborators before they escalate into issues. Clarifying ownership of data, and ensuring early agreement on technical standards and frameworks across institutions, are an important part of establishing trust and ensuring that a project runs smoothly.

EVITA SFORZI DI
DUPLICAZIONE
EVITA LA PERDITA
RISOLVE I
CONFLITTI

Vantaggi di un DMP

CESSDA Guide



⊖ Benefit 3. Clarifies needed budget

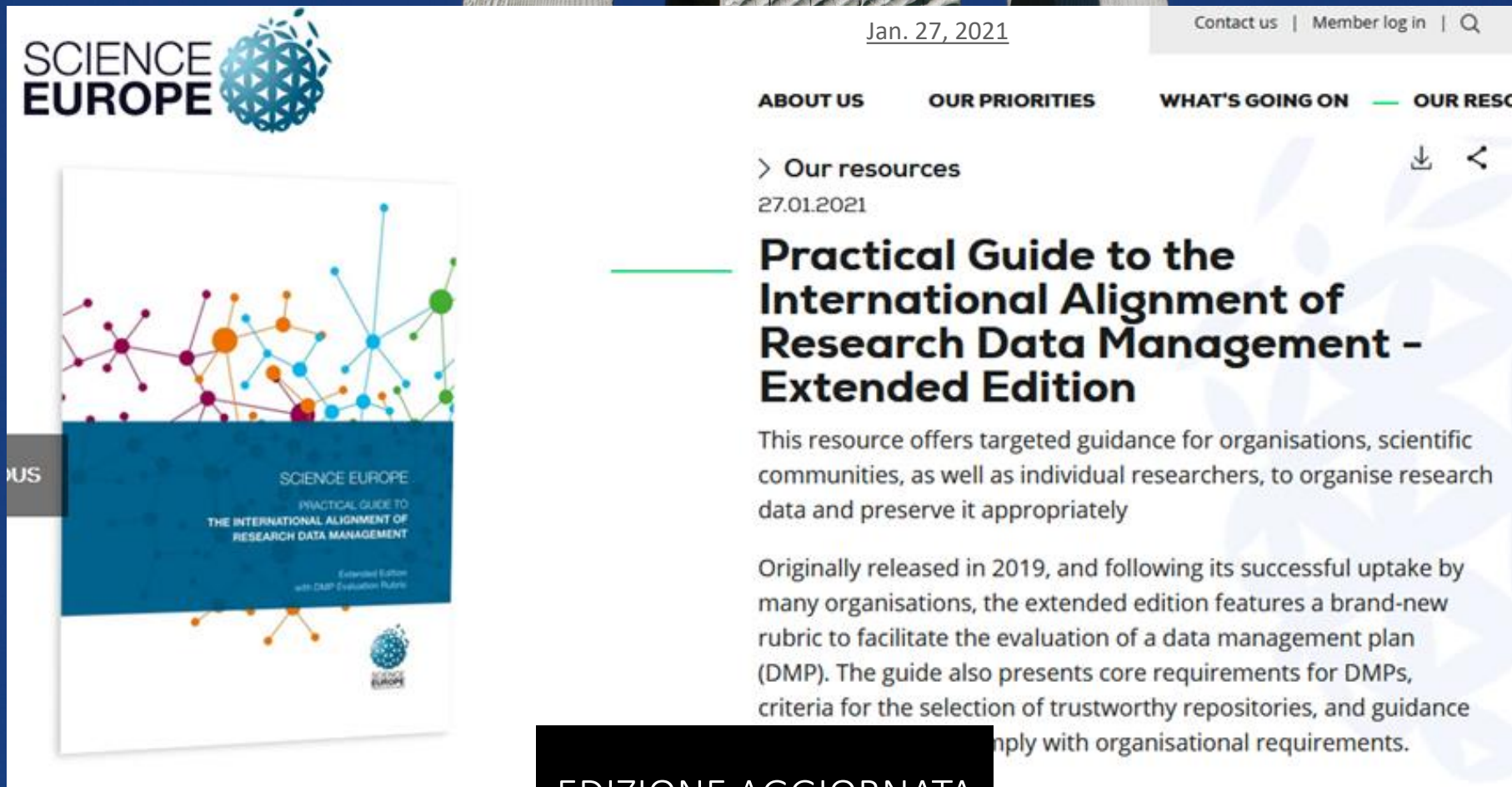
Data management is not free. You do not want to find yourself running out of funding before the end of the project because you have ignored or underestimated the cost of structured, detailed, and safe data management. Therefore, an important aspect of a DMP is its use in calculating how much money will be required for managing your research data during your research project.

A DMP can be useful in the process of applying for funding. Grant applications should not only include time and resources for collecting, analysing, and publishing on data in their budget, time and resources for careful documentation as well as server space, backup solutions, and documentation software need to be included as well. A DMP is also useful once funding is granted to plan and manage your expenses. Many research funders require a DMP as part of the application and decision-making process. The arguments for making data available are several, the most popular being that the data produced by public funds should be used to the greatest extent possible and available to the public. Unless there are legal, ethical or commercial barriers, data should also be openly available so that research results can be verified, replicated and reused.

Examples of Data Management cost assessments are given by the [University of Utrecht](#) (n.d.) and the Dutch Landelijk Coördinatiepunt Research Data Management ([LCRDM](#), n.d.) inspired by the '[Data management costing tool](#)' by UK Data Service, 2013.

È FONDAMENTALE PER
STIMARE I COSTI DI GESTIONE
- STIMATE LA DATA
STEWARDSHIP (IN-KIND?)
- POSSIBILI COSTI DI STORAGE
COSTI ERANO RIMBORSABILI IN
H2020 (6.2.D.3 AMGA) E IN
HORIZON EUROPE (6.2.C.3)

Guida al DMP



The screenshot shows the Science Europe website interface. At the top left is the Science Europe logo, which consists of the text 'SCIENCE EUROPE' next to a blue geometric globe icon. To the right of the logo is the date 'Jan. 27, 2021' and navigation links for 'Contact us', 'Member log in', and a search icon. Below the logo is a large image of the book cover for 'Practical Guide to the International Alignment of Research Data Management - Extended Edition'. The cover features a colorful network diagram of nodes and lines. The text on the cover includes 'SCIENCE EUROPE', 'PRACTICAL GUIDE TO THE INTERNATIONAL ALIGNMENT OF RESEARCH DATA MANAGEMENT', and 'Extended Edition with DMP Evaluation Rubric'. The Science Europe logo is also present at the bottom right of the cover. To the right of the book cover, the website's navigation menu includes 'ABOUT US', 'OUR PRIORITIES', 'WHAT'S GOING ON', and 'OUR RESOURCES'. Below the menu, there is a breadcrumb trail '> Our resources' and a download icon. The main heading of the page is 'Practical Guide to the International Alignment of Research Data Management - Extended Edition', dated '27.01.2021'. The introductory text states: 'This resource offers targeted guidance for organisations, scientific communities, as well as individual researchers, to organise research data and preserve it appropriately'. The summary text follows: 'Originally released in 2019, and following its successful uptake by many organisations, the extended edition features a brand-new rubric to facilitate the evaluation of a data management plan (DMP). The guide also presents core requirements for DMPs, criteria for the selection of trustworthy repositories, and guidance to comply with organisational requirements.'

EDIZIONE AGGIORNATA
DELLA GUIDA 2018

DMP Core Requirements

CORE REQUIREMENTS FOR DATA MANAGEMENT PLANS



When developing solid data management plans, researchers are required to deal with the following topics and answer the following questions:

- 1. Data description and collection or re-use of existing data**
 - a. How will new data be collected or produced and/or how will existing data be re-used?
 - b. What data (for example the kinds, formats, and volumes) will be collected or produced?
- 2. Documentation and data quality**
 - a. What metadata and documentation (for example the methodology of data collection and way of organising data) will accompany data?
 - b. What data quality control measures will be used?
- 3. Storage and backup during the research process**
 - a. How will data and metadata be stored and backed up during the research process?
 - b. How will data security and protection of sensitive data be taken care of during the research?
- 4. Legal and ethical requirements, codes of conduct**
 - a. If personal data are processed, how will compliance with legislation on personal data and on data security be ensured?
 - b. How will other legal issues, such as intellectual property rights and ownership, be managed? What legislation is applicable?
 - c. How will possible ethical issues be taken into account, and codes of conduct followed?

SEZIONI MINIME IN UN DMP

- 5. Data sharing and long-term preservation**
 - a. How and when will data be shared? Are there possible restrictions to data sharing or embargo reasons?
 - b. How will data for preservation be selected, and where will data be preserved long-term (for example a data repository or archive)?
 - c. What methods or software tools will be needed to access and use the data?
 - d. How will the application of a unique and persistent identifier (such as a Digital Object Identifier (DOI)) to each data set be ensured?
- 6. Data management responsibilities and resources**
 - a. Who (for example role, position, and institution) will be responsible for data management (i.e. the data steward)?
 - b. What resources (for example financial and time) will be dedicated to data management and ensuring that data will be FAIR (Findable, Accessible, Interoperable, Re-usable)?



SCIENCE EUROPE
PRACTICAL GUIDE TO
THE INTERNATIONAL ALIGNMENT OF
RESEARCH DATA MANAGEMENT

Extended Edition
with DMP Evaluation Public

Jan. 27, 2021



CORE REQUIREMENTS
FOR DATA MANAGEMENT PLANS



DMP Core Requirements

Translating the Core Requirements into a DMP template

The following example of a data management plan template is based on the core requirements for DMPs.⁶ These core requirements should be considered as a minimum standard, leaving the flexibility to formulate additional guidelines according to the needs of specific domains or to national or local legislation.

The template presented below refers to the 15 questions covering six core requirements for good data management. Additional guidance and explanations are provided to help researchers fill out such a template and to assure that all relevant aspects of research data management are covered. The below table is an example of how the core requirements can be transformed into a DMP template. It will be up to the individual organisations and disciplines to develop templates that fit their needs.

GENERAL INFORMATION

- Administrative information**
- Provide information such as name of applicant, project number, funding programme, version of DMP.

1 DATA DESCRIPTION AND COLLECTION OR RE-USE OF EXISTING DATA

- 1 a**
- 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.

UTILI COME
SPUNTO PER LE
PRIME VOLTE,
PER CAPIRE
QUALI DOMANDE
FARSI

2 DOCUMENTATION AND DATA QUALITY

2 a

What metadata and documentation (for example the methodology of data collection and way of organising data) will accompany the data?

- Indicate which metadata will help others identify and discover the data.
- Indicate which metadata standards (for example DDI, TEI, EML, MARC, CMDI) will be used.
- Use community metadata standards where these are in place.
- Indicate how the data will be organised during the project, mentioning for example conventions, version control, and folder structures. Consistent, well-ordered research data will be easier to find, understand, and re-use.
- Consider what other documentation is needed to enable re-use. This may include information on the methodology used to collect the data, analytical and procedural information, definitions of variables, units of measurement, and so on.
- Consider how this information will be captured and where it will be recorded for example in a database with links to each item, a 'readme' text file, file headers, code books, or lab notebooks.

2 b

What data quality control measures will be used?

- Explain how the consistency and quality of data collection will be controlled and documented. This may include processes such as calibration, repeated samples or measurements, standardised data capture, data entry validation, peer review of data, or representation with controlled vocabularies.

Jan. 27, 2021



SCIENCE EUROPE
PRACTICAL GUIDE TO
THE INTERNATIONAL ALIGNMENT OF
RESEARCH DATA MANAGEMENT

Extended Edition
with DMP Evaluation Rubric



DMP – Rubric for evaluation



3 STORAGE AND BACKUP DURING THE RESEARCH PROCESS

Guidance for Researchers

3a

How will data and metadata be stored and backed up during the research?

- Describe where the data will be stored and backed up during research activities and how often the backup will be performed. It is recommended to store data in least at two separate locations.
- Give preference to the use of robust, managed storage with automatic backup, such as provided by IT support services of the home institution. Storing data on laptops, stand-alone hard drives, or external storage devices such as USB sticks is not recommended.

Sufficiently Addressed The DMP...

- Clearly (even if briefly) describes:
 - › The location where the data and backups will be stored during the research activities.
 - › How often backups will be performed.
 - › The use of robust, managed storage with automatic backup (for example storage provided by the home institution).

or

- Explains why institutional storage will not be used (and for what part of the data) and describes the (additional) locations, storage media, and procedures that will be used for storing and backing up data during the project.

Insufficiently Addressed The DMP...

- Provides no information or very vague reference to how data will be stored and backed up during the project.

Guidance for Researchers

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.

Sufficiently Addressed The DMP...

- 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.

Insufficiently Addressed The DMP...

- 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.

DMI

D

GEN

Guid

Admi
infor

1 D

DMP questions

CESSDA guide

Adapt your Data Management Plan

A list of Data Management Questions based on the Expert Tour Guide on Data Management



CESSDA
EUROPEAN COORDINATED
INFRASTRUCTURE

Overview

Title of the project

Date of this plan

Description of the project

- What is the nature of the project?
- What is the research question?
- What is the project time line?

Origin of Data

- What kind of data will be used during the project?
- If you are reusing existing data: What is the scope, volume and format? How are different data sources integrated?
- If you are collecting new data can you clarify why this is necessary?

Principal researchers

- Who are the main researchers involved?
- What are their contact details?

Collaborating researchers (if applicable)

- What are their contact details and their roles in the project?

Funder (if applicable)

- If funding is granted, what is the reference number of the funding granted?

Data producer

- Which organisation has the administrative responsibility for the data?

Project data contact

- Who can be contacted about the project after it has finished?

Data owner(s)

- Which organisation(s) own(s) the data?
- If several organisations are involved, which organisation owns what data?

Roles

- Who is responsible for updating the DMP and making sure that it's followed?
- Do project participants have any specific roles?
- What is the project time line?

Costs

- Are there costs you need to consider to buy specific software or hardware?
- Are there costs you need to consider for storage and backup?
- Are potential expenses for (preparing the data for) archiving covered?

UTILI COME
SPUNTO PER LE
PRIME VOLTE,
PER CAPIRE
QUALI DOMANDE
FARSI

Organising and documenting your data

Data collection

- How will the data be collected?
- Is specific software or hardware or staff required?
- Who will be responsible for the data collection?
- During which period will the data be collected?
- Where will the data be collected?

Data organisation

- How will you organise your data?
- Will the data be organised in simple files or more complex databases?
- How will the data quality during the project be ensured?
- If data consists of many different file types (e.g. videos, text, photos), is it possible to structure the data in a logical way?

Data type and size

- What type(s) of data will be collected?
- What is the scope, quantity and format of the material?
- After the project: What is the total amount of data collected (in MB/GB)?

File format

- In what format will your data be?
- Does the format change from the original to the processed/final data?
- Will your (final) data be available in an open format?

Folder structure and names

- How will you structure and name your folders?

File structure and names

- How will you structure and name your files?

Documentation

- What documentation will be created during the different phases of the project?
- How will the documentation be structured?

Metadata

- What metadata will be provided with the collected/ generated/ reused data?
- How will metadata for each object be created?
- Is there any program that can be used to document the data?
- Can metadata be added directly into the files or will the metadata be produced in another program or document?

Metadata standard (if applicable)

- What metadata standard(s) will you use?

ORGANISE &
DOCUMENT



DMP questions



Protecting your data

Ethical review (if applicable)

- Does your project require approval by a local ethics committee?

Informed consent (if applicable)

- Do you require informed consent for your project?
- If so, how will permission be obtained?
- How are consent files organised and stored?

(sensitive) Personal data /confidential information (if applicable)

- How will access to (sensitive) personal data during the project be controlled?
- How will collaborators be granted access to the data in a secure way?
- If the research project is going to have data that includes confidential information or information that requires informed consent, is there a requirement to notify a privacy officer?
- Is there any confidential information within the material that requires special treatment and/or limits the access to it during/after the project?
- How will the material be protected during/after the project?
- How will permissions and restrictions be enforced?

Intellectual property rights (IPR)/Copyrights

- Are there IPR or copyright issues to consider?
- Will permission be needed to collect/reuse the data?
- Will these rights be transferred to another organisation for data distribution and archiving?

Agreements (if applicable)

- What are the agreements with other stakeholders?

Restrictions (if applicable)

- Any other restrictions that need to be considered?

Processing your data

PROCESS



Versioning

- What is your strategy concerning versioning your data files (and scripts) during the project?
- Will you create and/or follow a convention for versioning your data?
- Who will be responsible for securing that a "Masterfile" will be maintained, documented and versioned according to the project guidelines?
- How can different versions of a data file be distinguished?

Interoperability

- Will you make use of established software and hardware? If not, how does the software and hardware you use relate to other research?

If applicable:

- Will you make use of established terminologies/ontologies (i.e. structured controlled vocabularies) in the project? If not, how do your terminologies relate to established ones?
- Which coding is used (if any)? Will you build on established coding schemes? If not, how does your coding relate to other research?

STORE



Storing your data

Storage

- How and where will the data be stored during the project?
- For how long will the data be stored?

Backup

- How, where and at what intervals will the data be backed-up?
- How will data be recovered in the case of a data loss incident?

Security

- How will sensitive data be protected? (if applicable)
- How will data access be managed?

UTILI COME
SPUNTO PER LE
PRIME VOLTE,
PER CAPIRE
QUALI DOMANDE
FARSI

Basic Information.

- State the purpose of the data collection/generation.
- Explain the relation to the objectives of the project
- Consider what data will be collected or created as part of the study (RAW data).
- Consider what data will be produced by processing the RAW data (Secondary, processed data).
- Specify if existing data is being re-used (if any)
- Specify the origin of the data
- Specify the types and formats you plan to use for the data generated/collected (raw, processed, published).
- Consider what data will be published as the result of your study (Published data).

Volume and Life Cycle of the Data.

If you are using FAIRDOM, we will look after data that will be retained and potentially exchanged by your projects. It will help with local storage for temporarily-held local data prior to processing.

For RAW data, please consider the following:

- How much RAW data you think will be produced (Estimates, per month, year, full project duration)?
- Will all of the RAW data be kept for the duration of the study or will the RAW data be deleted once it is processed?
- For large scale RAW data (images, sequence) have you planned the local storage capacity necessary for processing?
- Do you require help to organise a suitable local management system for RAW data?
- Do you have policies that govern the management and usage of RAW data?
- How long will RAW data be kept? Will there be a long-term archive?

For Secondary and Published data, please consider the following:

- What data processing is foreseen in the project?
- How much processed data will be produced, and stored (can you make estimates per month, year, full project)?
- How much of this data will be published? (Estimates per month, year, full project)?
- Does your institution, or the project funders, have policies governing the access and usage of processed data?

Additional for personally sensitive data (e.g medical data)

- When looking at the data flow through the project, define what data is:
 - aggregated (typically safe to share, if names cannot be recovered)
 - anonymized (name cannot be recovered from the data)
 - pseudonymized (name can be recovered by some)
 - non-anonymized (name linked to data)
- Determine which organisational boundaries have to be traversed by which data.
- Make sure with your "local" data protection officer and ethics commission that the data can be shared with your partners along the flow described with the anonymisation levels as described. Why local? Some laws change across surprising boundaries. E.g. in Germany Universities and other public organisations are subject to another data protection law than enterprises. Why seek advice? In some cases you may be required to be able to recover the name-data-relation, e.g. to enable study participants to "leave" a study.

Data Management Checklist

<https://fair-dom.org/knowledgehub/data-management-checklist/>

UTILI COME
SPUNTO PER LE
PRIME VOLTE,
PER CAPIRE
QUALI DOMANDE
FARSI

DMP in Horizon Europe

IN HORIZON EUROPE
- NELLA PROPOSTA 1 PAGINA
SULLA GESTIONE DEI DATI
DMP COME DELIVERABLE M6




V.1 Feb. 2021
Horizon Europe (HORIZON)
Euratom Research and Training Programme
(EURATOM)
General Model Grant Agreement
EIC Accelerator Contract
(HE MGA — Multi & Mono)



V.2 April 2021



Horizon Europe Programme
Standard Application Form (RIA, IA)
Application form (Part A)
Project proposal – Technical description (Part B)
Version 2.0
22 April 2021

⚠️ *Proposals selected for funding under Horizon Europe will need to develop a detailed data management plan (DMP) for making their data/research outputs findable, accessible, interoperable and reusable (FAIR) as a deliverable by month 6 and revised towards the end of a project's lifetime.*

⚠️ *For guidance on open science practices and research data management, please refer to the relevant section of the [HE Programme Guide](#) on the Funding & Tenders Portal.*

Open science: research data management

The beneficiaries must manage the digital research data generated in the action ('data') responsibly, in line with the FAIR principles and by taking all of the following actions:

- establish a data management plan ('DMP') (and regularly update it)
- as soon as possible and within the deadlines set out in the DMP, deposit the data in a trusted repository; if required in the call conditions, this repository must be federated in the EOSC in compliance with EOSC requirements

DMP in Horizon Europe



HEU DMP

Horizon Europe

Data Management Plan Template

Version 1.0
05 May 2021

Data Summary

- Will you re-use any existing data and what will you re-use it for?
- What types and formats of data will the project generate or re-use?
- What is the purpose of the data generation or re-use and its relation to the objectives of the project?
- What is the expected size of the data that you intend to generate or re-use?
- What is the origin/provenance of the data, either generated or re-used?
- To whom might your data be useful ('data utility'), outside your project?

FAIR data

- 2.1. Making data findable, including provisions for metadata: Will data be identified by a persistent identifier?
- 2.1. Making data findable, including provisions for metadata: Will rich metadata be provided to allow discovery? What metadata will be created? What disciplinary or general standards will be followed? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how.
- 2.1. Making data findable, including provisions for metadata: Will search keywords be provided in the metadata to optimize the possibility for discovery and then potential re-use?
- 2.1. Making data findable, including provisions for metadata: Will metadata be offered in such a way that it can be harvested and indexed?
- 2.2. Making data accessible - Repository: Will the data be deposited in a trusted repository?
- 2.2. Making data accessible - Repository: Have you explored appropriate arrangements with the identified repository where your data will be deposited?

- 2.2. Making data accessible - Repository: Does the repository ensure that the data is assigned an identifier? Will the repository resolve the identifier to a digital object?
- 2.2. Making data accessible - Data:
Will all data be made openly available? If certain datasets cannot be shared (or need to be shared under restricted access conditions), explain why, clearly separating legal and contractual reasons from intentional restrictions. Note that in multi-beneficiary projects it is also possible for specific beneficiaries to keep their data closed if opening their data goes against their legitimate interests or other constraints as per the Grant Agreement.
- 2.2. Making data accessible - Data:
If an embargo is applied to give time to publish or seek protection of the intellectual property (e.g. patents), specify why and how long this will apply, bearing in mind that research data should be made available as soon as possible.
- 2.2. Making data accessible - Data:
Will the data be accessible through a free and standardized access protocol?
- 2.2. Making data accessible - Data:
If there are restrictions on use, how will access be provided to the data, both during and after the end of the project?
- 2.2. Making data accessible - Data:
How will the identity of the person accessing the data be ascertained?
- 2.2. Making data accessible - Data:
Is there a need for a data access committee (e.g. to evaluate/approve access requests to personal/sensitive data)?
- 2.2. Making data accessible - Metadata:
Will metadata be made openly available and licenced under a public domain dedication CC0, as per the Grant Agreement? If not, please clarify why. Will metadata contain information to enable the user to access the data?

TROVATE GIÀ IL
TEMPLATE IN
DMPONLINE

DMP online

PERSONALIZZABILE

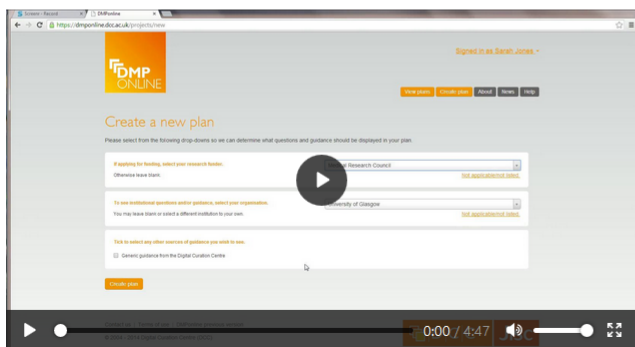


Home About Future plans Help Change language

Welcome.

DMPonline helps you to create, review, and share data management plans that meet institutional and funder requirements. It has been jointly developed by the Digital Curation Centre (DCC) and the University of California Curation Center (UC3).

Screencast on how to use DMPonline



Sign in

Veteran tapes

Project Details Plan overview Write Plan Share Download

expand all | collapse all

13/13 answered

Data Collection (2 / 2)

What data will you collect or create?

B *I* [List icons] [Link icon] [Table icon]

The "Veteran tape " project will collect and generate different types of datasets:

Type of data	volume	Format	Storage format
Video recordings	600 x 1Gb	.mkv	.mkv
Transcriptions	600 x 1500Kb	MS Word	.txt
Structured interview text	1 x 500Kb	MS word	.txt

For the video recordings the selected format is .mkv; the same .mkv format will be used for the long-term preservation .

Transcriptions will be written in MS Word and then stored as .txt files.

We checked the format compatibility against EASY File format <https://dans.knaw.nl/en/deposit/information-about-depositing-data/before-depositing/file-formats>

As the total volume of data is greater than 50Gb, DANS requires a fee for the storage. We are currently in touch with EASY to determine the costs of archiving.

Save

Guidance

Comments (1)

DCC

DCC guidance

Guidance

Questions to consider:

- What type, format and volume of data?
- Do your chosen formats and software enable sharing and long-term access to the data?
- Are there any existing data that you can reuse?

Guidance:

Give a brief description of the data, including any existing data or third-party sources that will be used, in each case noting its content, type and coverage. Outline and justify your choice of format and consider the implications of data format and data volumes in terms of storage, backup and access.

GRATUITO
BASTA REGISTRARSI; POI SI
ACCEDE E SI TROVANO TUTTI I
PROPRI DMP NEL DESKTOP

DMP online e Data Wizard video

OA@unito.it

In UniTO **Come** Cos'è utile Perché è importante Editori italiani Eventi Corsi e formazione Video

Come scrivere un Data Management Plan

Il Data Management Plan (DMP) è un documento strutturato, vivo, che cresce con il progetto. Serve a dichiarare come si producono i dati, come li si conserverà e come li si condividerà (se possibile).

Pensatelo come le "Istruzioni per l'uso" dei vostri dati.

Deve essere

- **sintetico**: evitate sproloqui, non è una dissertazione. Frasi chiare che diano informazioni precise
- **schematico**: utilizzate il più possibile tabelle e punti elenco
- **preciso**: evitate frasi (viste davvero) tipo "we expect a huge size of data" o "data will be available". Servono solo a far perdere tempo a chi lo scrive e a chi lo legge. Quantificate: we expect max 50 GB; data will be available in Zenodo upon publication of the paper
- **specifico**: non copiate da modelli. Ogni ricerca è a sé, ogni ente ha le sue procedure
- **coerente**: scrivete solo ciò di cui siete certi, se non sapete, ditelo.

GUIDE AL DMP OVVERO: LE DOMANDE CHE DOVETE FARVI

- Science Europe **Practical Guide** to the International Alignment of Research Data Management (2021)
- **DMP tips and tricks** (2021)
- **CESSDA Adapt your DMP** (2019)

STRUMENTI ONLINE PER SCRIVERE UN DMP

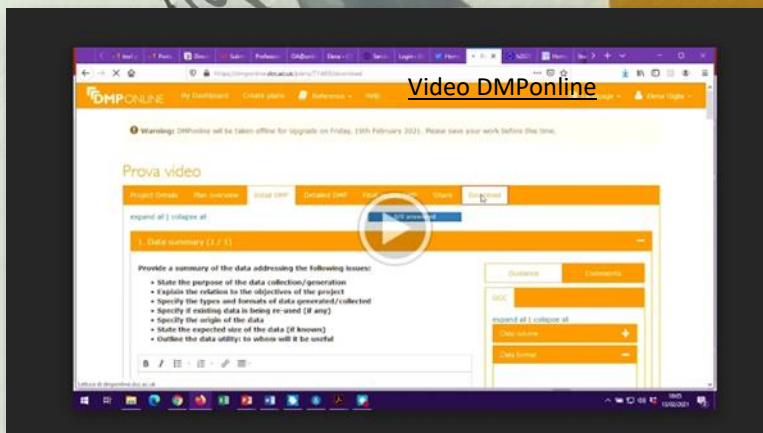
- **DMP online** con modello per Horizon2020 - **Video tutorial** su come utilizzarlo
- **Data wizard** con modello Horizon2020 e Science Europe - **Video tutorial** su come utilizzarlo
- **DMPtool** per funder USA
- **ARGOS OpenAIRE**

COME VALUTARE UN DMP (ma serve anche a capire come scriverlo bene)

In UniTO
Regolamento di Ateneo
Open Access in pratica
Open Data
Eventi
Horizon2020
Risorse

OA@UniTO

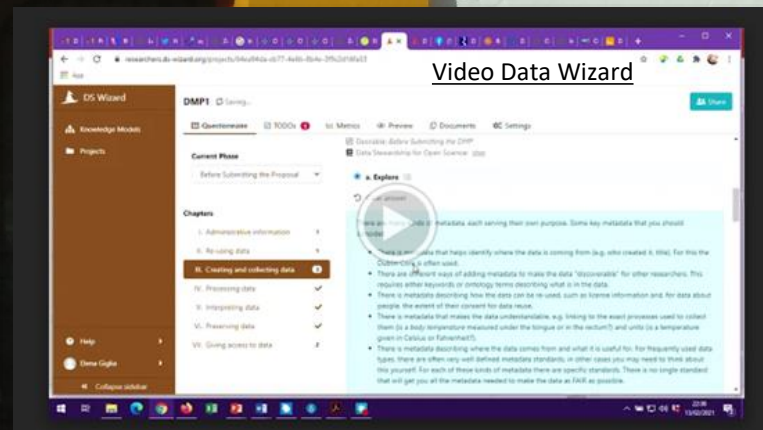
Video DMPonline



VISITA PLAZA DE TOROS



Video Data Wizard



Create a new plan

Before you get started, we need some information about your research project to set you up with the best DMP template for your needs.

* What research project are you planning?

mock project for testing, practice, or educational purposes

* Select the primary research organisation

- or - No research organisation associated with this plan or my research organisation is not listed

* Select the primary funding organisation

- European Commission (Horizon 2020)
- European Research Council (ERC)

- or - No funder associated with this plan or my funder is not listed

PREDISPOSTO PER LE 3 VERSIONI

ALLA FINE, SCARICA IL PDF

Prova

- Project Details
- Plan overview
- Initial DMP
- Detailed DMP
- Final review DM
- Share
- Download

CONDIVISIBILE PER SCRIVERLO INSIEME

Guidance

write your plan, DMPonline can show from a variety of organisations.

6 organisations to see their

Curation Centre

from additional organisations

See the full list

Save

Briefly summarise your research project to help others understand the purposes for which the data are being collected or created.

GUIDA

Principal Investigator

Name

PLEASE UPDATE Y

IDENTIFICATIVI

ORCID ID

Email

elena.giglia@unito.it

Phone

Data Contact Person

Same as Principal Investigator

Save

Set plan visibility

Public or organisational visibility is intended for finished plans. You must answer at least 50% of the questions to enable these options. Note: test plans are set to private visibility by default.

- Private: visible to me, specified collaborators and administrators at my organisation
- Organisation: anyone at my organisation can view
- Public: anyone can view

PRIVATO O PUBBLICO?

Manage collaborators

Invite specific people to read, edit, or administer your plan. Invitees will receive an email notification that they have access to this plan.

Email address	Permissions
elena.giglia@unito.it	Owner

Invite collaborators

* Email

* Permissions

- Co-owner
- Editor
- Read only

SI POSSONO INVITARE
COLLABORATORI CON LIVELLI DI
PERMESSO DIVERSI

Template version 0, published on 04 August 2021

Instructions

Data Summary

- Will you re-use any existing data and what will you re-use it for?
- What types and formats of data will the project generate or re-use?
- What is the purpose of the data generation or re-use and its relation to the objectives of the project?
- What is the expected size of the data that you intend to generate or re-use?
- What is the origin/provenance of the data, either generated or re-used?
- To whom might your data be useful ('data utility'), outside your project?

Horizon Europe Template

Write plan

FAIR data

- 2.1. Making d
- 2.1. Making d
- 2.1. Making d

FAIR data

- 2.1. Making data findable, including provisions for metadata: Will data be identified by a persistent identifier?
- 2.1. Making data findable, including provisions for metadata: Will rich metadata be provided to allow discovery? What metadata will be created? What disciplinary or general standards will be followed? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how.
- 2.1. Making data findable, including provisions for metadata: Will search keywords be provided in the metadata to optimize the possibility for discovery and then potential re-use?
- 2.1. Making data findable, including provisions for metadata: Will metadata be offered in such a way that it can be harvested and indexed?
- 2.2. Making data accessible - Repository: Will the data be deposited in a trusted repository?
- 2.2. Making data accessible - Repository: Have you explored appropriate arrangements with the identified repository where your data will be deposited?
- 2.2. Making data accessible - Repository: Does the repository ensure that the data is assigned an identifier? Will the repository resolve the identifier to a digital object?
- 2.2. Making data accessible - Data:
Will all data be made openly available? If certain datasets cannot be shared (or need to be shared under restricted access conditions), explain why, clearly separating legal and contractual reasons from intentional restrictions. Note that in multi-beneficiary projects it is also possible for specific beneficiaries to keep their data closed if opening their data goes against their legitimate interests or other constraints as per the Grant Agreement.
- 2.2. Making data accessible - Data:
If an embargo is applied to give time to publish or seek protection of the intellectual property (e.g. patents), specify why and how long this will apply, bearing in mind that research data should be made available as soon as possible.
- 2.2. Making data accessible - Data:
Will the data be accessible through a free and standardized access protocol?

- 2.4. Increase data re-use:
How will you provide documentation needed to validate data analysis and facilitate data re-use (e.g. readme files with information on methodology, codebooks, data cleaning, analyses, variable definitions, units of measurement, etc.)?
- 2.4. Increase data re-use:
Will your data be made freely available in the public domain to permit the widest re-use possible? Will your data be licensed using standard reuse licenses, in line with the obligations set out in the Grant Agreement?
- 2.4. Increase data re-use:
Will the data produced in the project be useable by third parties, in particular after the end of the project?
- 2.4. Increase data re-use:
Will the provenance of the data be thoroughly documented using the appropriate standards?
- 2.4. Increase data re-use:
Describe all relevant data quality assurance processes.
- 2.4. Increase data re-use:
Further to the FAIR principles, DMPs should also address research outputs other than data, and should carefully consider aspects related to the allocation of resources, data security and ethical aspects.

Other research outputs

- In addition to the management of data, beneficiaries should also consider and plan for the management of other research outputs that may be generated or re-used throughout their projects. Such outputs can be either digital (e.g. software, workflows, protocols, models, etc.) or physical (e.g. new materials, antibodies, reagents, samples, etc.).
- Beneficiaries should consider which of the questions pertaining to FAIR data above, can apply to the management of other research outputs, and should strive to provide sufficient detail on how their research outputs will be managed and shared, or made available for re-use, in line with the FAIR principles.

Data Summary

- Will you re-use any existing data and what will you re-use it for?
- What types and formats of data will the project generate or re-use?
- What is the purpose of the data generation or re-use and its relation to the objectives of the project?
- What is the expected size of the data that you intend to generate or re-use?
- What is the origin/provenance of the data, either generated or re-used?
- To whom might your data be useful ('data utility'), outside your project?

...E NON «A HUGE AMOUNT OF DATA»

COSTI

Data volume

- Note what volume of data you will create in MB/GB/TB. Indicate the proportions of raw data, processed data, and other secondary outputs (e.g., reports).
- Consider the implications of data volumes in terms of storage, access and preservation. Do you need to include additional costs?
- Consider whether the scale of the data will pose challenges when sharing or transferring data between sites; if so, how will you address these challenges?

TABELLA PER DATI DI FORMATO DIVERSO

Data format

- Clearly note what format(s) your data will be in, e.g., plain text (.txt), comma-separated values (.csv), geo-referenced TIFF (.tif, .tiff).
- Explain why you have chosen certain formats. Decisions may be based on staff expertise, a preference for open formats, the standards accepted by data centres or widespread usage within a given community.
- Using standardised, interchangeable or open formats ensures the long-term usability of data; these are recommended for sharing and archiving.
- See UK Data Service guidance on [recommended formats](#) or DataONE Best Practices for [file formats](#).

SCRIVETE DIRETTAMENTE (USATE TABELLE E PUNTI ELENCO)

IN OGNI PASSAGGIO POTETE SALVARE E CONTINUARE DOPO

USATE FORMATI STANDARD (CON ELENCO)

DMP ONLINE My Dashboard Create plans Reference Help Language Elena Gigli

Perugia Prova

Project Details Contributors Plan overview Write Plan Share Download

expand all | collapse all 0/41

Data Summary (0 / 6)

Will you re-use any existing data and what will you re-use it for?

Guidance Comments

DCC

the reasons if re-use of any existing has been considered but discarded.

Save

PREVIEW DI TUTTE LE DOMANDE

FAIR data

- 2.1. Making data findable, including provisions for metadata: Will data be identified by a persistent identifier?
- 2.1. Making data findable, including provisions for metadata: Will rich metadata be provided to allow discovery? What metadata will be created? What disciplinary or general standards will be followed? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how.
- 2.1. Making data findable, including provisions for metadata: Will search keywords be provided in the metadata to help others identify and discover the data?
- 2.1. Making data findable, including provisions for metadata: Will metadata be offered in such a way that

- What metadata will be provided to help others identify and discover the data?
- Researchers are strongly encouraged to use community metadata standards where these are in place. The Research Data

Alliance offers a [Directory of Metadata Standards](#). Data

repositories may also provide guidance about appropriate metadata standards.

- Consider what other documentation is needed to enable reuse. This may include information on the methodology used to collect the data, analytical and procedural information, definitions of variables, units of measurement, any assumptions made, the format and file type of the data and software used to collect and/or process the data.
- Consider how you will capture this information and where it will be recorded, e.g., in a database with links to each item, in a 'readme' text file, in file

FAIR data (0 / 25)

2.1. Making data findable, including provisions for metadata: Will data be identified by a persistent identifier?

Guidance

DCC

Metadata & documentation

GUIDE SPECIFICHE

Save

2.1. Making data findable, including provisions for metadata: Will rich metadata be provided to allow discovery? What metadata will be created? What disciplinary or general standards will be followed? In case metadata standards do not exist in your discipline, please outline what type of metadata will be created and how.

Guidance

DCC

Metadata & documentation

B I ☰ ☷ 🔗 📄

2.2. Making data accessible - Data:

Will all data be made openly available? If certain datasets cannot be shared (or need to be shared under restricted access conditions), explain why, clearly separating legal and contractual reasons from intentional restrictions. Note that in multi-beneficiary projects it is also possible for specific beneficiaries to keep their data closed if opening their data goes against their legitimate interests or other constraints as per the Grant Agreement.

B I [List Icon] [List Icon] [Link Icon] [Table Icon]

Save

AS OPEN AS POSSIBLE AS
CLOSED AS NECESSARY

RICORDATEVI CHE
AVETE FIRMATO UN
GRANT AGREEMENT...

Guidance

Comments

DCC

Data sharing

COMMENTI

- How will you share the data e.g. deposit in a data repository, use a secure data service, handle data requests directly or use another mechanism? The methods used will depend on a number of factors such as the type, size, complexity and sensitivity of the data.

Open science: research data management

The beneficiaries must manage the digital research data generated in the action ('data') responsibly, in line with the FAIR principles and by taking all of the following actions:

- establish a data management plan ('DMP') (and regularly update it)
- as soon as possible and within the deadlines set out in the DMP, deposit the data in a trusted repository; if required in the call conditions, this repository must be federated in the EOSC in compliance with EOSC requirements

- as soon as possible and within the deadlines set out in the DMP, ensure open access — via the repository — to the deposited data, under the latest available version of the Creative Commons Attribution International Public License (CC BY) or Creative Commons Public Domain Dedication (CC0) or a licence with equivalent rights, following the principle 'as open as possible as closed as necessary', unless providing open access would in particular:

- be against the beneficiary's legitimate interests, including regarding commercial exploitation, or
- be contrary to any other constraints, in particular the EU competitive interests or the beneficiary's obligations under this Agreement; if open access is not provided (to some or all data), this must be justified in the DMP

Other research outputs (0 / 2)

In addition to the management of data, beneficiaries should also consider and plan for the management of other research outputs that may be generated or re-used throughout their projects. Such outputs can be either digital (e.g. software, workflows, protocols, models, etc.) or physical (e.g. new materials, antibodies, reagents, samples, etc.).

B *I*

Save

NUOVA SEZIONE: GESTIONE DI
«ALTRI RISULTATI» (SOFTWARE,
PROTOCOLLI, MODELLI...)

Beneficiaries should consider which of the questions pertaining to FAIR data above, can apply to the management of other research outputs, and should strive to provide sufficient detail on how their research outputs will be managed and shared, or made available for re-use, in line with the FAIR principles.

B *I*

Comments

Add comments to share with collaborators

B *I*

Save

Comments

Add comments to share with collaborators

B *I*

Allocation of resources (0 / 4)

What will the costs be for making data or other research outputs FAIR in your project (e.g. direct and indirect costs related to storage, archiving, re-use, security, etc.) ?

B I ☰ ☷ 🔗 📄

How will these be covered? Note that costs related to research data/output management are eligible as part of the Horizon Europe grant (if compliant with the Grant Agreement conditions)

B I ☰ ☷ 🔗 📄

COSTI SONO
RIMBORSABILI (GRANT
6.2.C.3)

General > Article 6.2.C.3 Other goods, works and services

Costs related to research output management (HE) – Costs for research output management (e.g. *management of research data*) are eligible if the eligibility conditions are fulfilled, including open access to peer-reviewed publications (but see the additional eligibility condition referenced immediately below), research data and other outputs.

CHI RISPONDE DELLA
GESTIONE DEI DATI?

Who will be responsible for data management in your project?





B I ☰ ☷ 🔗 📄

Roles & responsibilities

- Outline the roles and responsibilities for all activities, e.g., data capture, metadata production, data quality, storage and backup, data archiving & data sharing. Individuals should be named where possible.
- For collaborative projects you should explain the coordination of data management responsibilities across partners.
- See UK Data Service guidance on [data management roles and responsibilities](#) or DataONE Best Practices: [Define roles and assign responsibilities for data management](#).

Data security (0 / 1)

What provisions are or will be in place for data security (including data recovery as secure storage/archiving and transfer of sensitive data)?

B I    

Save

SICUREZZA DEI DATI: SERVE
POLICY DI ATENEO

Storage & security

- Describe where the data will be stored and backed up during the course of research activities. This may vary if you are doing fieldwork or working across multiple sites so explain each procedure.
- Identify who will be responsible for backup and how often this will be performed. The use of robust, managed storage with automatic backup, for example, that provided by university IT teams, is preferable. Storing data on laptops, computer hard drives or external storage devices alone is very risky.
- See UK Data Service Guidance on [data storage](#) or DataONE Best Practices for [storage](#).
- Also consider data security, particularly if your data is sensitive e.g., detailed personal data, politically sensitive information or trade secrets. Note the main risks and how

Are there, or could there be, any ethics or legal issues that can have an impact on data sharing? These can also be discussed in the context of the ethics review. If relevant, include references to ethics deliverables and ethics chapter in the Description of the Action (DoA).

B I [list icon] [list icon] [link icon] [table icon]

Save

BASE LEGALE PER DATI
PERSONALI; CONSENSI
INFORMATI VANNO IN
QUESTA SEZIONE

Guidance

Comments

DCC

Ethics & privacy

+

Will informed consent for data sharing and long term preservation be included in questionnaires dealing with personal data?

B I [list icon] [list icon] [link icon] [table icon]

Save

Guidance

Comments

DCC

Refer to other national/funder/sectorial/departmental procedures for data management that you are using (if any)

B I [list icon] [list icon] [link icon] [table icon]

Save

QUI SAREBBE
URGENTE UNA
POLICY DI ATENEO O
IL NATIONAL PLAN
OPEN SCIENCE

Guidance

Comments

DCC

Related policies

- Consider whether there are any existing procedures that you can base your approach on. If your group/department has local guidelines that you work to, point to them here.
- List any other relevant funder, institutional, departmental or group policies on data management, data sharing and data security.

Perugia Prova

- Project Details
- Contributors
- Plan overview
- Write Plan
- Share
- Download**

Format

- pdf
- csv
- html
- pdf**
- text
- docx
- json

- project details coversheet
- question text and section headings
- unanswered questions

PDF formatting

Font

Face

Arial, Helvetica, Sans-Serif

Size (pt)

10

Margin (m)

Top

25

Download Plan

1. Data summary

Provide a summary of the data addressing the following issues:

- 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

Purpose: _____

Relation: _____

Data will be created as the research is original.

Reuse: videos watched on YouTube, chats coming from social platforms

Origin: interviews and observations; content analysis

Type of data	Format (ongoing)	Format (preservation)	Expected Size	Consent collected	Reuse Rights
A Textual	.doc	.txt	10 M		
A1 [Interview template]					
A2 [Consent form]					
A3 [Interview transcriptions]				x	
A4 [Code book]	.txt				
A5 [Report]					
A6 [Video transcriptions]					CC BY
B Tabular data	.xls	.csv	50 M		
B1 [Participants description]					
B2 [Thematic analysis]					
B3 [Content analysis]					
C Audio			10 G		
C1 [Recording]	.mp3	still to be decided		x	
D Video			100 G		
D1 [Recorded interviews]	.mp4			x	
D2 [Downloaded]					CC BY

A scenic view of a rocky coastline. The foreground is dominated by clear, turquoise water with visible ripples and reflections. In the background, a rugged, light-colored rock formation rises from the water's edge, topped with sparse, dark green vegetation. The sky is a clear, bright blue. A black rectangular box is centered in the middle of the image, containing the text "BUON LAVORO!" in white, uppercase letters.

BUON LAVORO!