#### **CORSI DI FORMAZIONE**









21 - 23 MAGGIO 2024

DMP – Data Management Plan

Elena Giglia



elena.giglia@unito.it
@egiglia



[CODATA-international] Last chance to register: CODATA/DDI Alliance webinar - Questions and Survey Instruments in DDI: Maximizing the Value of Your Metadata. 22 May 2024, 13.00 – 14.30 UTC. Free, online. D Posta in arrivo x

00:48 (7 ore fa)

Laura Molloy Annulla iscrizione

a codata-international -



Webinar: Questions and Survey Instruments in DDI - Maximising the Value of Your Metadata

22 May 2024, 13.00 - 14.30 UTC. Registration closing soon: register here!

CODATA and the DDI Alliance are delighted to announce our new collaborative webinar, "Questions and Survey Instruments in DDI: Maximizing the Value of Your Metadata", free and online tomorrow, Wed 22 May 2024, 13.00 - 14.30 UTC.

Register here: https://us02web.zoom.us/meeting/register/tZAscOggqzItE9Yzvez84dGsymesbNX\_IRdm#/registration

Questions and the questionnaires in which they are used serve as an important focus when reusing data, but also serve as a resource which can be used to maximise comparability across studies, and to make the design of questionnaires more efficient. Further, a rigorous description of survey logic and flow can provide the basis for automation of the data collection process, heightening efficiencies for longitudinal and other repeated surveys.

This webinar looks at the DDI structures for describing questions and survey instruments. Our guest speakers illustrate how this metadata can be leveraged to best effect, with examples from CLOSER, a collaborative digital infrastructure for longitudinal population studies, data and research, and two national statistics agencies: Insee (the French National Institute of Statistics and Economic Studies), and StatsCan (Statistics Canada). Becky Oldroyd, Metadata Officer at CLOSER, shares the ways that rich metadata at the variable-, question- and questionnaire-level improves the discoverability of data for longitudinal, cross-study research. Éric Sigaud describes how Insee is developing open-source tooling to support metadata-driven workflows, such as automated creation of survey instruments. Finally, Flavio Rizzolo talks about StatsCan's ongoing work of adopting DDI-Lifecycle tools to design and manage questionnaires using Colectica together with the agency's in-house data tools and systems.

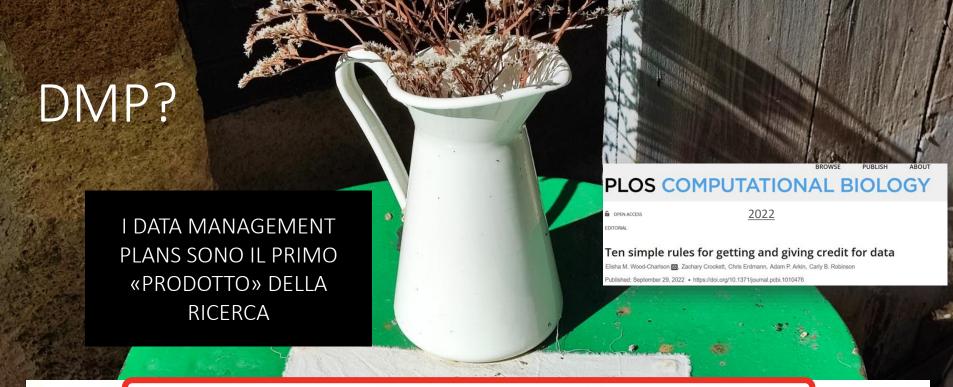
Time will be reserved for questions and discussion with the audience.

Register here: https://us02web.zoom.us/meeting/register/tZAscOgggzItE9Yzvez84dGsymesbNX\_IRdm#/registration

We look forward to seeing you on Wednesday!

Best, Laura





## Rule 3: Data management plans are your first research product

Now that you have mastered the complexity (or at least scratched the surface) of what it takes to create FAIR, comparable, and reproducible data, we need to talk about data management plans (DMPs). These are often required by funders as supplementary documents to research grants, where you outline when, where, and how data from the project will be preserved and shared. We won't go into best practices for creating a DMP, as that is well articulated by Michener [28]. However, we do want to emphasize that DMPs are no longer just supplementary pdfs. They can (and should) be created as FAIR, machine-actionable, living documents [29]. DMPs establish the initial node in your upcoming research product network (data, code, etc.). DMPs connect the people and data to the funding agency and put a stake in the ground for the

UN DOCUMENTO FORMALE SULLA GESTIONE DEI DATI CHE NE ASSICURA INTEGRITÀ 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

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

È LA SEDE IN CUI

1) GIUSTIFICATE LE SCELTE

OPEN/CLOSED

2) STIMATE I COSTI

...IL DATA MANAGEMENT PLAN

# Data management ABC -Per partire



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 COSA SERVE A VALIDARE

LA MIA RICERCA?

**NEL DMP** 

PROSPETTIVA DIVERSA

include for someone else to replicate your results? Think about it in t your own research.

TUTTO QUESTO VA INSERITO A bibliography would be the most immediate and obvious starting po revealing to the reader all the sources that you have used to base you research. But what of the gathering mechanisms you used? Did you database or undertake statistical analysis? If so you need to make the SULLA VOSTRA RICERCA database and statistics available. This doesn't just mean providing the mes 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





**SCHOOL OF** ADVANCED STUDY OF LONDON



# Trucchi e suggerimenti

Top tip - keep it short and specific!

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

SINTETICO E SPECIFICO

## NON COPIATE

Advice on within Data Management Plans

Research
Data

Improved Data Management Plans

4. Keep it simple

Informative: two
audiences

Specific: e.g. name
standards

Concise: < ¼ to 3
pages

Don't forget: your
achievements &
innovation

FOSTER toolkit

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

ESSERE GENERICI NON SERVE A NULLA

[we expect a huge size of data; data will be available]

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

- SE NON LO SAPETE, DITELO
- SE NO, SEMBRA CHE NON SIATE AL CORRENTE DI QUELLO SPECIFICO
   ASPETTO [STESSA DIFFERENZA FRA UNA CELLA VUOTA E CELLA CON N.A.]

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 PSEUDONIMYZED, non

**ANONIMYZED** 



# e suggerimenti / 2

- 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
- MAKE USE OF SUPPORT
  Use your in-house support services like RDM Support, the Library, IT department or legal desk
- THINK BROAD

  Also address software code, algorithms and any other valuable research assets in your DMP
- COPY WHERE YOU CAN

  Look at other (submitted) plans and copy when appropriate

- BE UNIQUE WHERE NEEDED
  Since every research project is unique, so are the data it
  - Since every research project is unique, so are the data it generates. Copying from sample DMPs is not sufficient
- BE CONCRETE

  Make your answers as concrete as possible. Show that you have consulted RDM experts
- SAY SO IF YOU DON'T KNOW
  Indicate what you do not yet know and how you will
  resolve these questions later
- UPDATE

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

# re Requirements





SCIENCE EUROPE

THE INTERNATIONAL ALIGNMENT OF RESEARCH DATA MANAGEMENT





#### **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?
  - 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. SE NON AVETE UN MODELLO RICHIESTO DAL FUNDER (ES. HORIZON FUROPE PER LA COMMISSIONE EUROPEA) UTILIZZATE QUESTO MODELLO DI **SCIENCE EUROPE** 

#### 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)?
- 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)?

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.

#### DATA DESCRIPTION AND COLLECTION OR RE-USE OF **EXISTING DATA**

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 o 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.



What metadata and documentation (for example the methodology 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 · 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.
- · 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.



of data collection and way

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

**FARSI** 

What data quality control measures will be used?



Jan. 27, 2021

SCIENCE EUROPE

THE INTERNATIONAL ALIGNMENT OF RESEARCH DATA MANAGEMENT

Extended Editio with DMP Evaluation Rubri



# DMP – Guida



#### 3 STORAGE AND BACKUP DURING THE RESEARCH PROCESS

#### Guidance for Researchers

3a

DMI

GEN

Guid

Admi

inform

1 D

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).

0

 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.

# DMP - domande

CESSDA guide

#### Adapt your Data Management Plan

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





ORGANISE DOCUMENT

#### 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?

#### 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

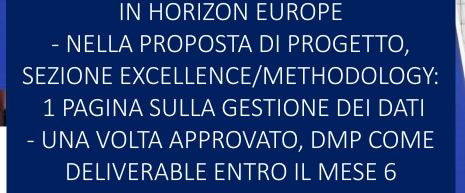
https://fair-dom.org/knowledgehub/data-management-checklist/

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

**FARSI** 



#VisitEP





Horizon Europe Programme Standard Application Form (RIA, IA)

Application form (Part A)
Project proposal – Technical description (Part B
Version 2.0



<u>V.1 Feb. 2021</u>

Horizon Europe (HORIZON)

Euratom Research and Training Programme
(EURATOM)

General Model Grant Agreement EIC Accelerator Contract

(HE MGA - Multi R. Mono

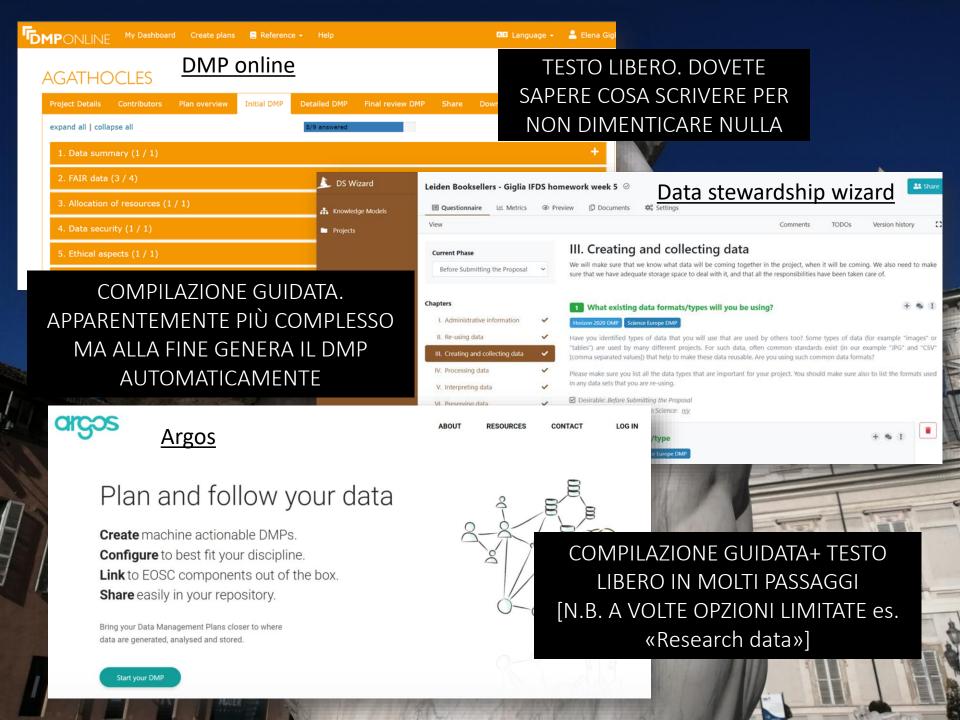
Proposals selected for funding under Horizon Europe will need to develop a detailed data management plan (DMP) for making their data/research outputs findable, accessible, interoperable and reusable (FAIR) as a deliverable by month 6 and revised towards the end of a project's lifetime.

For guidance on open science practices and research data management, please refer to the relevant section of the <u>HE Programme Guide</u> on the Funding & Tenders Portal.

Open science: research data management

The beneficiaries <u>must manage the digital research data generated in the action ('data')</u> 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



# Caratteristiche a confronto

# **MPONLINE**

- TEMPLATE HEU E SCIENCE EUROPE
- ALTRI TEMPLATE PERSONALIZZABILI
- TESTO LIBERO CON POSSIBILITÀ DI INSERIRE TABELLE ED ELENCHI PUNTATI
- GUIDA SINTETICA PUNTO PER PUNTO

-

 SCRITTURA COLLABORATIVA



- KNOWLEDGE MODEL (ELIXIR)
- WIZARD MOLTO
   DETTAGLIATO
   (DIFFICILE OMETTERE)
- APRE I CAPITOLI DEL LIBRO DI BAREND MONS (DO/DON'T)
- SCRITTURA COLLABORATIVA
- CREA TO DO LIST
- MACHINE ACTIONABLE

argos

- TEMPLATE HORIZON EUROPE E SCIENCE EUROPE
- POSSIBILE ASSOCIARE DIVERSI DATASET A UNO STESSO DMP



- MACHINE ACTIONABLE
- CONNESSO A
   OPENAIRE RESEARCH
   GRAPH (LINK FRA DATI
   E PUBBLICAZIONI)







Home

About

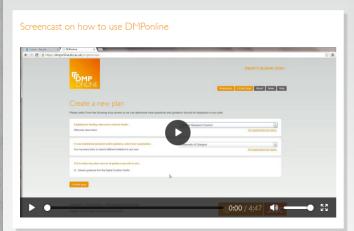
Future plans

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).



GRATUITO

BASTA REGISTRARSI; POI SI

ACCEDE E SI TROVANO TUTTI I

PROPRI DMP NEL DESKTOP

## 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?

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/fileformats

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.

Guidance

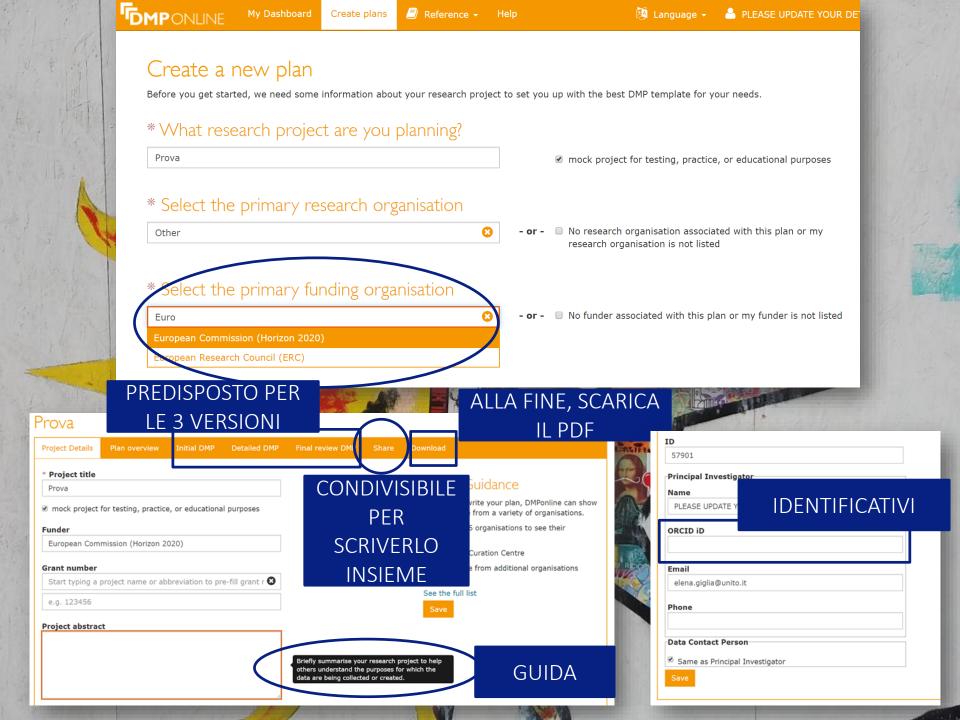
Guidance
Ouestions to consider:

- · What type, format and volume of data?
- Do your chosen formats and software enable sharing and long-term access to
- 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.







**AMOUNT OF DATA**»

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**

- Clearly note what format(s) your data will be in, e.g., plain text (.txt), comma-separated values (.csv), geo-referenced TIFF (.tif,
- · 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.

...E NON «A HUGE

COSTI

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 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?

• What is the purpose of the data generation or re-use and its relation to the objectives of the project?

Perugia Prova

**'DMP**ONLINE

Data Summary

Write Plan expand all I collapse all Will you re-use any existing data and what will you re-use it for? SCRIVETE DIRETTAMENTE the reasons if re-use of any existing as been considered but discarded. (USATE TABELLE E **PUNTI ELENCO)** 

> IN OGNI PASSAGGIO POTETE SALVARE E CONTINUARE DOPO

**USATE FORMATI** STANDARD (CON **ELENCO**)

## PREVIEW DI TUTTE LE **DOMANDE**

#### FAIR data

- 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 met discovery and then potential re-use?
- 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 provid 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



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

Guidance

B / ≒ + \ = - 8 | ⊞ -

**GUIDE SPECIFICHE** 

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

 $I \models \neg \models \neg \mid \mathscr{S} \mid \boxplus \neg$ 

#### 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.)?

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

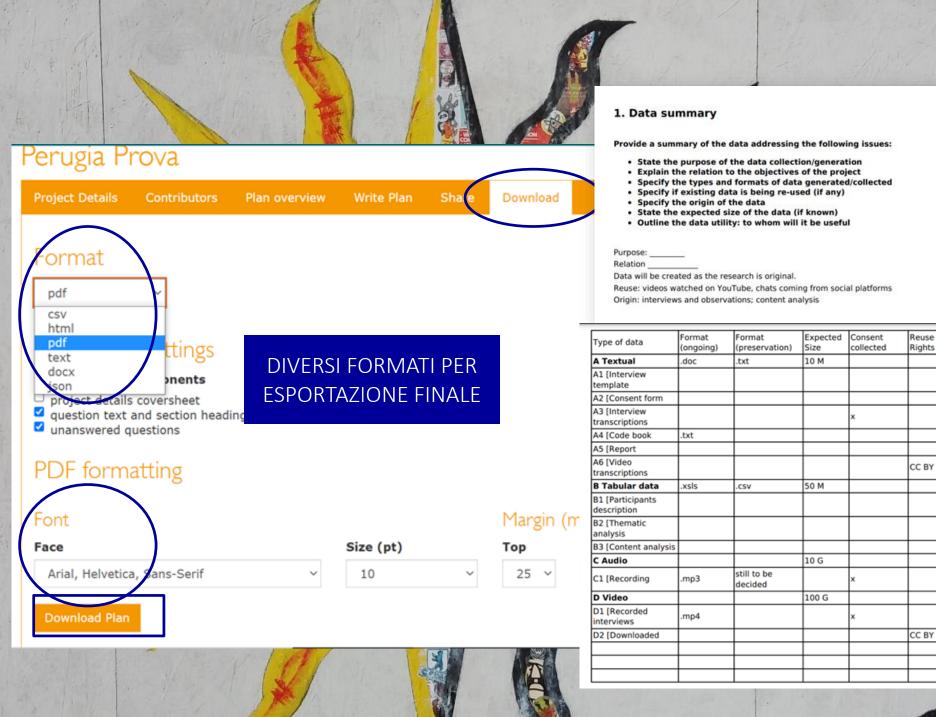
**Costs related to research output management** (HE) – Costs for research output management (e.g. management of research data) are eligible if the eligibility conditions are fulfilled, including open access to peer-reviewed publications (but see the additional eligibility condition referenced immediately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

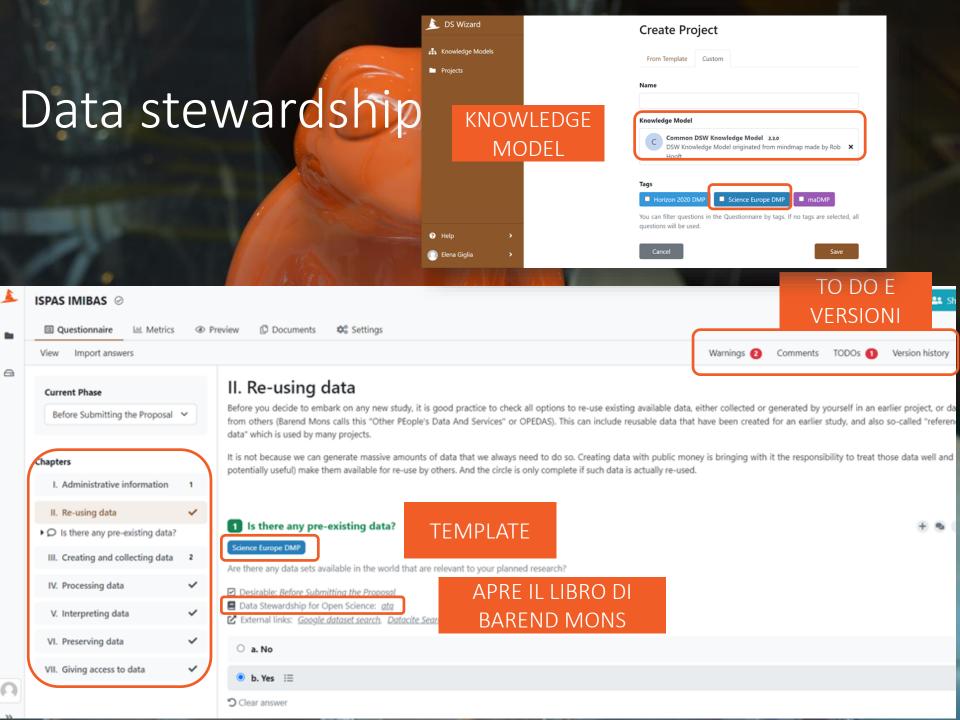
CHI RISPONDE DELLA GESTIONE DEI DATI?

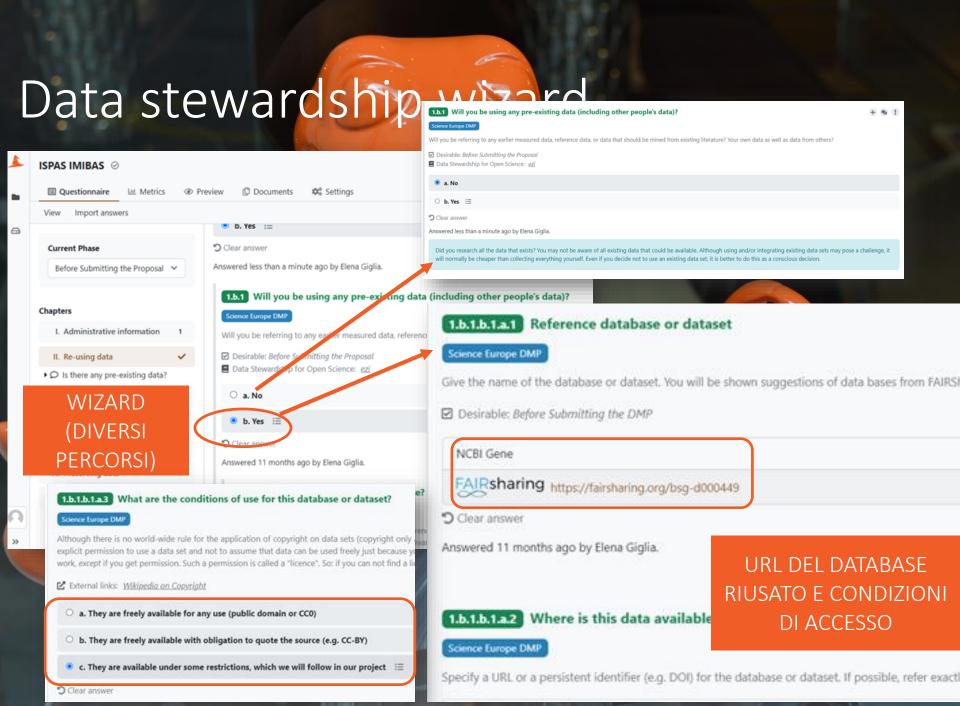
Who will be responsible for data management in your project?

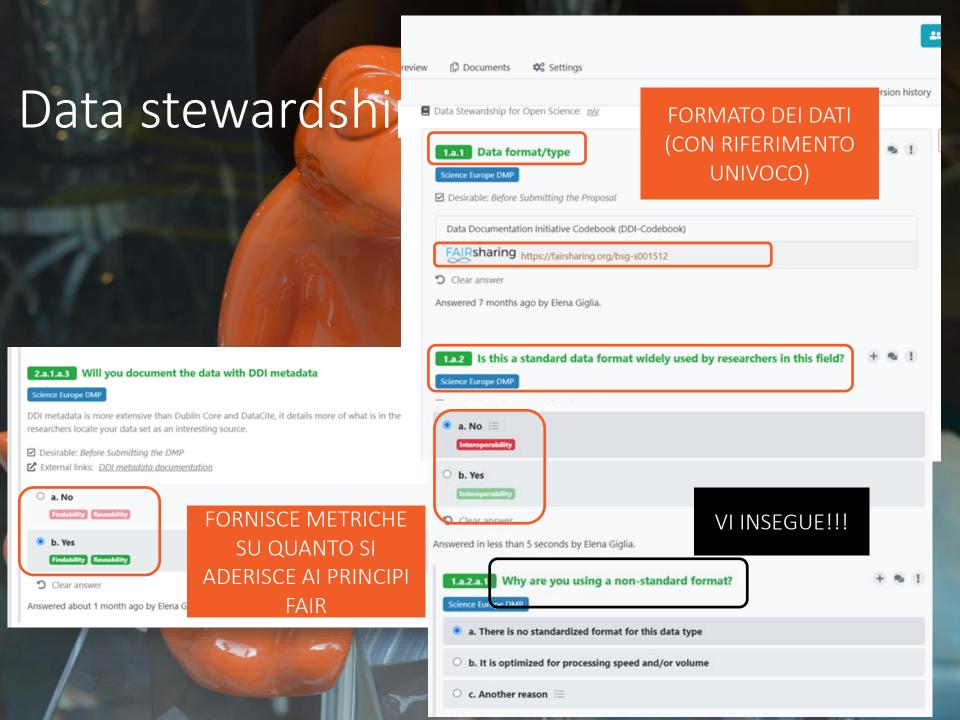
#### Roles & responsibilities

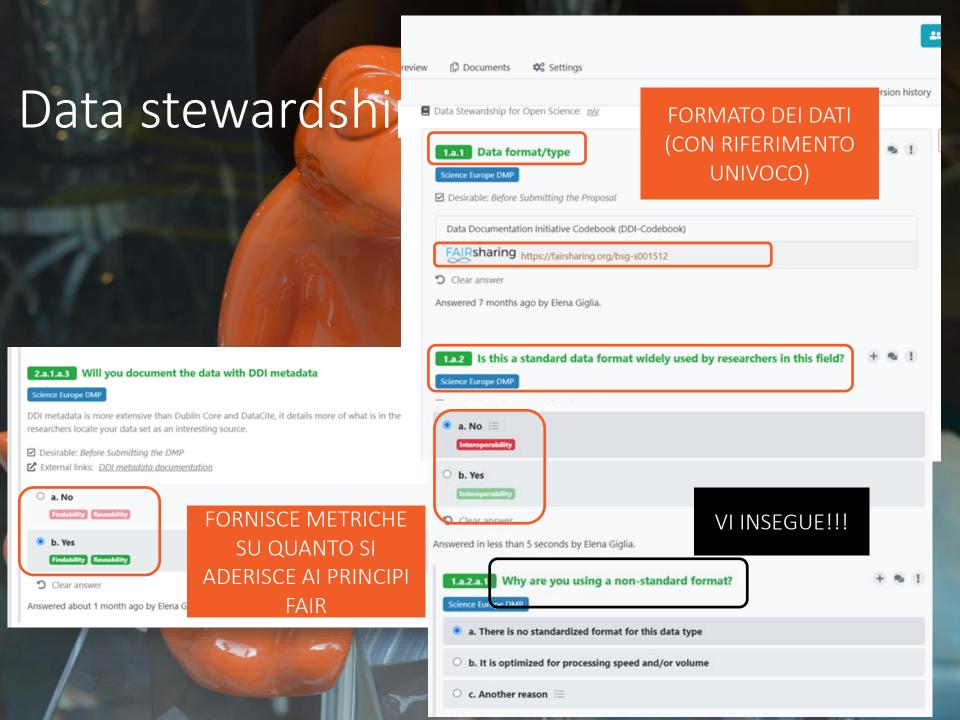
- Outline the roles and responsibilities for all activities, e.g., data capture, metadata production, data quality, storage and backup, data archiving & data sharing. Individuals should be named where possible.
- For collaborative projects you should explain the coordination of data management responsibilities 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 stewardship wizard



We will use the following reference datasets:

databae COVID del Ministero salute
 Chttps://www.salute.gov/it/portale/nuovocoronavirus/dettaglioPubblica

Chttps://www.salute.gov.it/portale/nuovocoronavirus/dettaglioPubblicazioniNu lingua-italiano&id-3147)

We will use version "bollettino 9/12/2021 (ver.2.2)" of this dataset. If a new version becomes available during the project, we will stay with the old version.

#### Data formats and types

We will be using the following data formats and types:

tabellari

It is a standardized format. This is not a suitable format for long-term archiving; however, we plan to convert it to a suitable format before the end of the project. We expect to have 30 GB of data in this format.

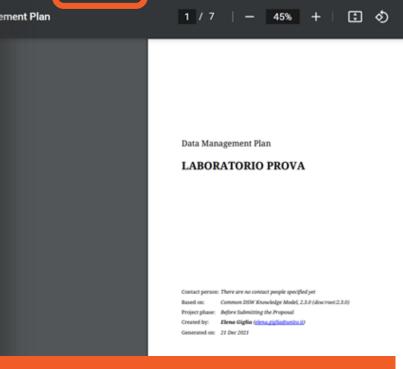
testuali

#### 2. How will the data be collected or created?

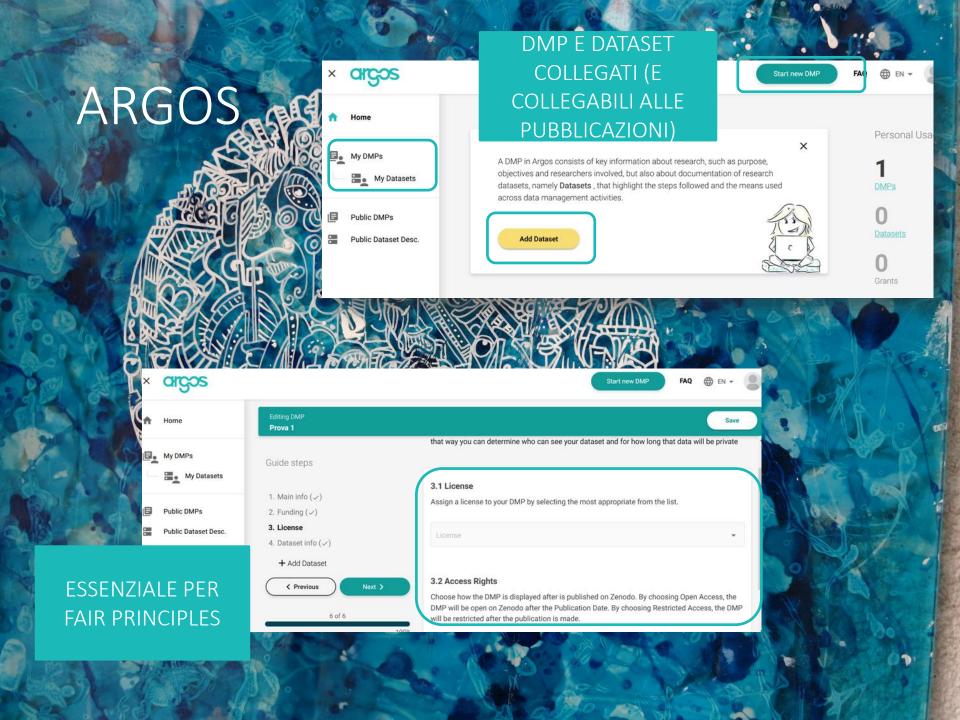
There will be no instrument dataset in this project.

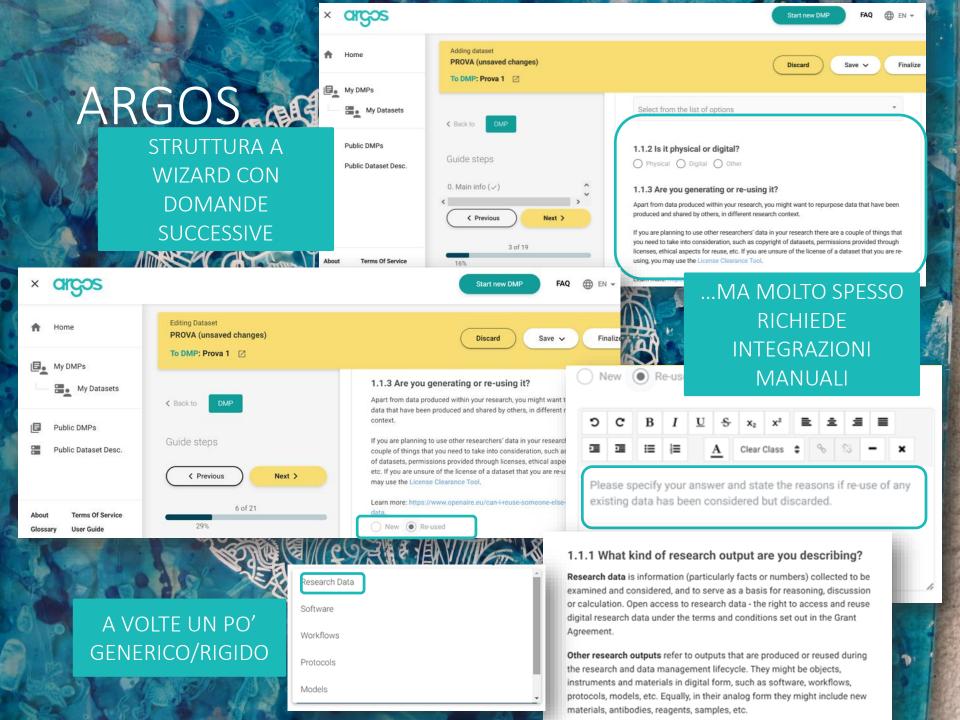
#### Storage and file conventions

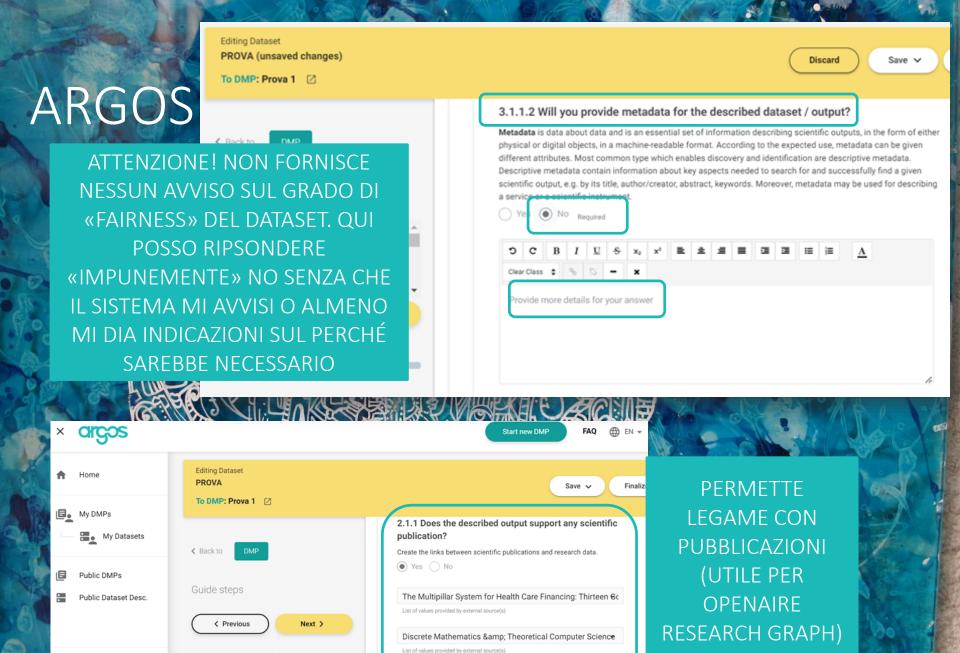
We will use a filesystem with files and folders with the following folder



AL TERMINE SCARICA IL DMP... SENZA CHE NOI ABBIAMO SCRITTO NULLA, ABBIAMO SOLO FORNITO RISPOSTE CHE IL SISTEMA RIELABORA E ASSEMBLA OVE RILEVANTE







ZENODO

6 of 21

29%

About

Contact Support

Terms Of Service

User Guide



ULTIMATE - Data Management Plan

Aitor Corchero (orcid:0000-0002-8463-4128), Joep van den Broeke (orcid:0000-0002-5707-740X)

Funder European CommissionIIEC

Grant
ULTIMATE: indUstry
water-utiLiTy symblosis
for a sMarter wATer
society/ No 869318

NEL PRODOTTO FINALE SI
VEDE LA DIFFERENZA FRA
TESTO ESTRAPOLATO DAL
SISTEMA E TESTO INSERITO
MANUALMENTE

#### Datasets

Title: Case Studies Dataset Template: Horizon 2020

The Case Studies in ULTIMATE are collecting experimental data from laboratory experiments and pilot scale water treatment installations. This dataset contain the public data collected as part of the Case Studies.

#### **Dataset Description**

- 1.1 Data Summary
  - 1.1.1 What is the purpose of the data collection/generation and its relation to the objectives of the project?
  - . To keep on record
  - + To develop a product
  - 1.1.2 What are the types of the described generated/collected data?
  - · sample or specimen data
  - · observational (e.g.
  - · sensor data
  - · data from surveys)
  - + experimental (e.g.
  - . gene sequencing data)
  - + simulation (e.g.
  - climate modeling data).
  - 1.1.3 What are the formats of the described generated/collected data?
  - + Text files
  - \* Numerical
  - Models

1.1.3 What are the formats of the described generated/collected data?

- · Text files
- Numerical

Batch datasets in CSV and JSON formats

1.1.4 What is the origin of the described data?

Primary data

1.1.5 What is the expected size of the described data?

MB (megabyte)

The size of data is around 100 to 500 MB including electricity, gas and meteorological datasets  $\,$ 

1.1.6 To whom might it be useful ('data utility')?

- Researchers
- Research communities

SEZIONE SUL RIUSO

Watercolor painting

#### OA@unito.it

Come scrivere un DMP

TROVATE I CORSI COMPLETI E LA PAGINA SUI DMP

#### Seminari

#### 2022

- Open Science dalla A alla Z, Unviersità Bocconi, 25/1;
- 2. Open Science come e perché / biblioteche UniTO (pro
- Laboratorio Open Science prof. Paccagnella nov-dic 2
- Open Science come e perché / biblioteche civiche (pre
- Open Science why and how / LTTA event (progetto CE)
- Open Science why and how, MSCA candidates, Univer
- Open Science, questa sconosciuta, Dipartimento di N
- 8. Come fare Open Access (con un pizzico di Open Scien
- Open Science come e perché, Università di Trieste, 05
- Open Science why and how, SISSA, Trieste, 05/03
- 11. Lavorare su Open Science alla luce di Horizon Europe
- 12. Open Science is here to stav. Digital Humanities course

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

Come scrivere un Data Management Plan

producono i dati, come li si conserverà e come li si condividerà (se possibile).

Deve essere

Open science it

## **OPEN-SCIENCE.IT**

La scienza condivisa

## Technology - MIP Politecn

- 14. Il futuro è Open: come cam letterarie - la Sapienza, Ro
- 15. Open Science A to Z, Phd S
- 16. FAIR data basics, ISPAS pro
- 17. Open Science dalla A alla Z
- 18. Principi FAIR (con un pizzio

## 13. Open Science, la ricerca al E voglio informazioni su:



Gestione dei dati della ricerca

- 19. Open Science A to Z, ISPAS project, IMIBAS, 01/17-19
- 20. FAIR data basics ISPAS project, University of Girona, 01/11-13



Ricercatore

Sono un:

I dati dalla ricerca tra protezione e licenze per il riuso

Perché è importante Editori e Politiche Open Access (EPOcA) Eventi Corsi e formazione

- Gestire i dati, un compito fondamentale per se stessi e per gli altri
- Che cos'è il Data Management Plan
- Cosa sono i principi FAIR
- Data Management plan: strumenti e risorse utili

In Unito

Regolamento di Ateneo

Open Access in pratica

Open Data

Il Data Management Plan (DMP) è un documento strutturato, vivo, che cresce con il progetto. Serve a dichiarare come si