

HORIZON
EUROPE

@UNIBO 

Research Data Management & Data Management Plan

Sara Coppini

Data Steward, Research Division (ARIC)
Alma Mater Studiorum – Università di Bologna

Per il ciclo di incontri "Gestione e conservazione dei dati della
ricerca" Nell'ambito del percorso "HorizonEurope@Unibo"



23 maggio 2023

HORIZON
EUROPE

@UNIB 

Indice

- Il contesto
- Il dato
- Il ciclo di vita dei dati e il Research Data Management (RDM)
- Un *FAIR* research data management
- Il Data Management Plan (DMP)
- Il data steward
- Punti chiave

Il contesto

Come ente finanziatore ma anche attraverso iniziative internazionali, l'Unione Europea riconosce la **centralità del dato** e della sua valorizzazione nella ricerca e per la comunità.

La **condivisione** di **buone pratiche e metodologie** di gestione dei dati della ricerca secondo i principi FAIR e dell'Open Science - a partire dall'uso di strumenti come il Data Management Plan - è essenziale per massimizzare il valore e l'impatto della ricerca scientifica.

Anche la nuova figura professionale del **data steward** va oltre al supporto tecnico-amministrativo per l'adempimento degli obblighi contrattuali: è un ricercatore in **dialogo** con altri ricercatori per riflettere sull'importanza del dato e della sua gestione nello specifico **dominio disciplinare** di riferimento.



Why does research data matter?

Research data are a primary output of research.



European Commission

Shaping Europe's digital future

A European Strategy for data

The strategy for data focuses on putting people first in developing technology, and defending and promoting European values and rights in the digital world.

Data is an essential resource for economic growth, competitiveness, innovation, job creation and societal progress in general.

The European strategy for data aims at creating a single market for data that will ensure Europe's global competitiveness and data sovereignty.

Il dato

Your visible output

Publications

Your research data



Informazioni, in particolare fatti o numeri, raccolte per essere esaminate e considerate come base per ragionamenti, discussioni o calcoli.

I record fattuali (numeri, testi, immagini e suoni) utilizzati come fonti primarie per la ricerca scientifica e che sono comunemente accettati dalla comunità scientifica come necessari per convalidare i risultati delle ricerche.

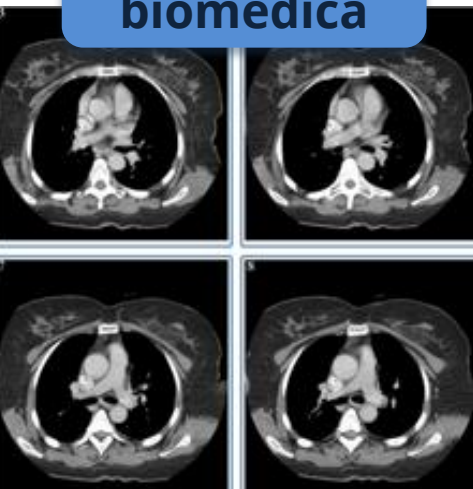
L'intersezione tra scienza dell'informazione e tecnologia ha cambiato la natura e il significato del dato: tutto ciò che si può rappresentare sotto forma di bits è dato.

Il dato

È un **asset di valore** per la ricerca e il ricercatore.

Ogni ricerca produce o riusa dei dati, ma ogni **area disciplinare** ha le sue specificità nella tipologia di dati prodotti e nelle loro strategie di gestione.

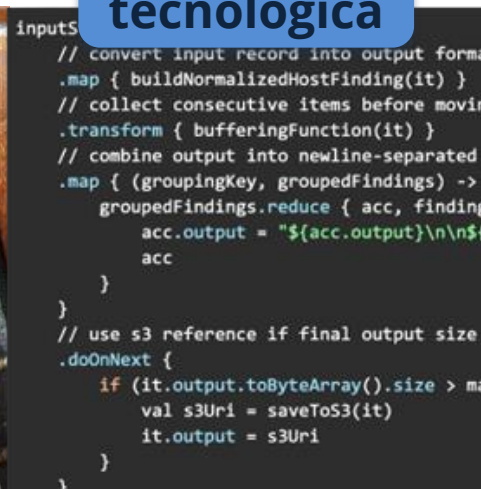
area
biomedica



area
umanistica



area
tecnologica

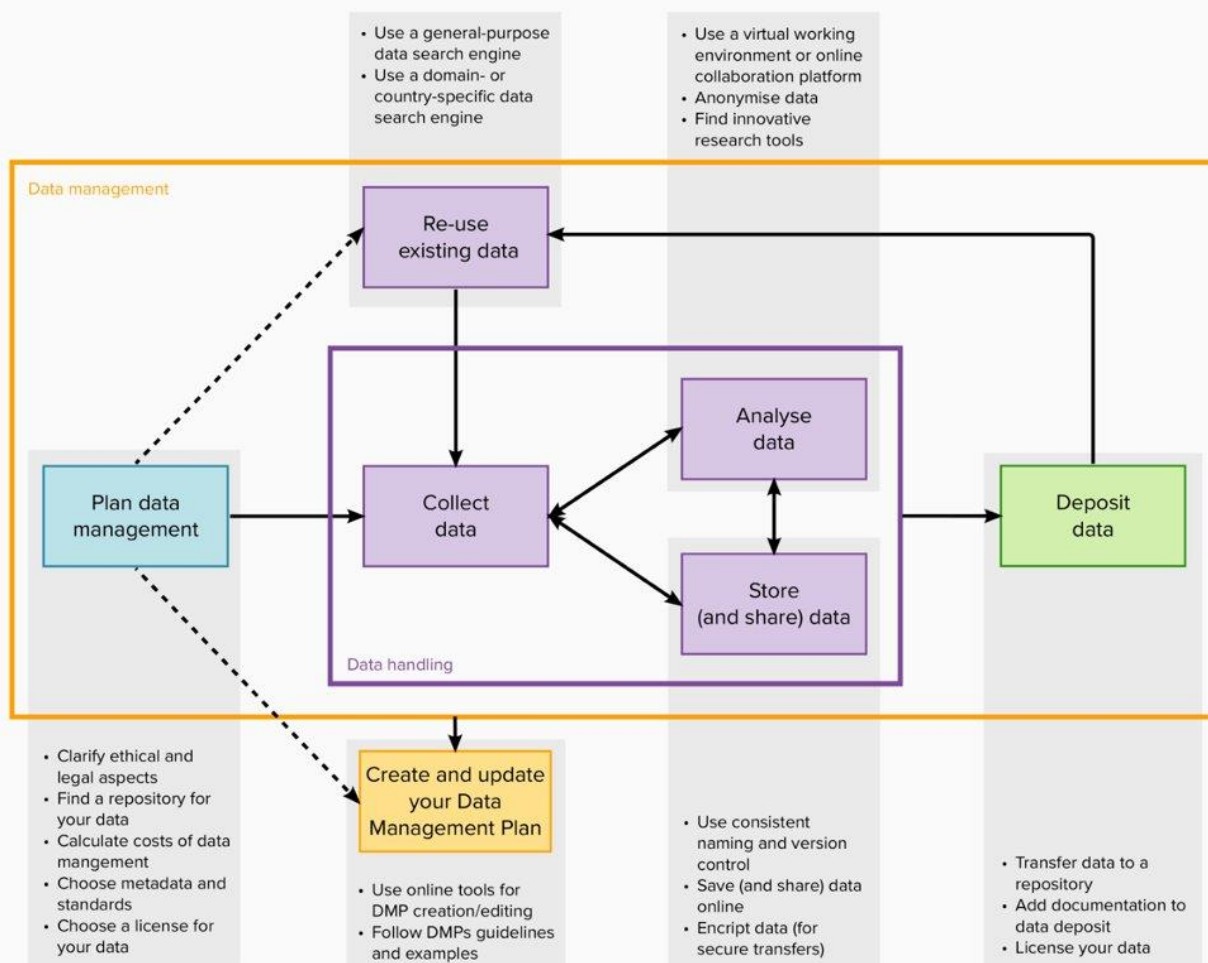


area
sociale



23 maggio 2023

Il ciclo di vita dei dati e il Research Data Management (RDM)



Si riferisce all'organizzazione, allo storage e alla conservazione dei dati creati durante un progetto di ricerca.

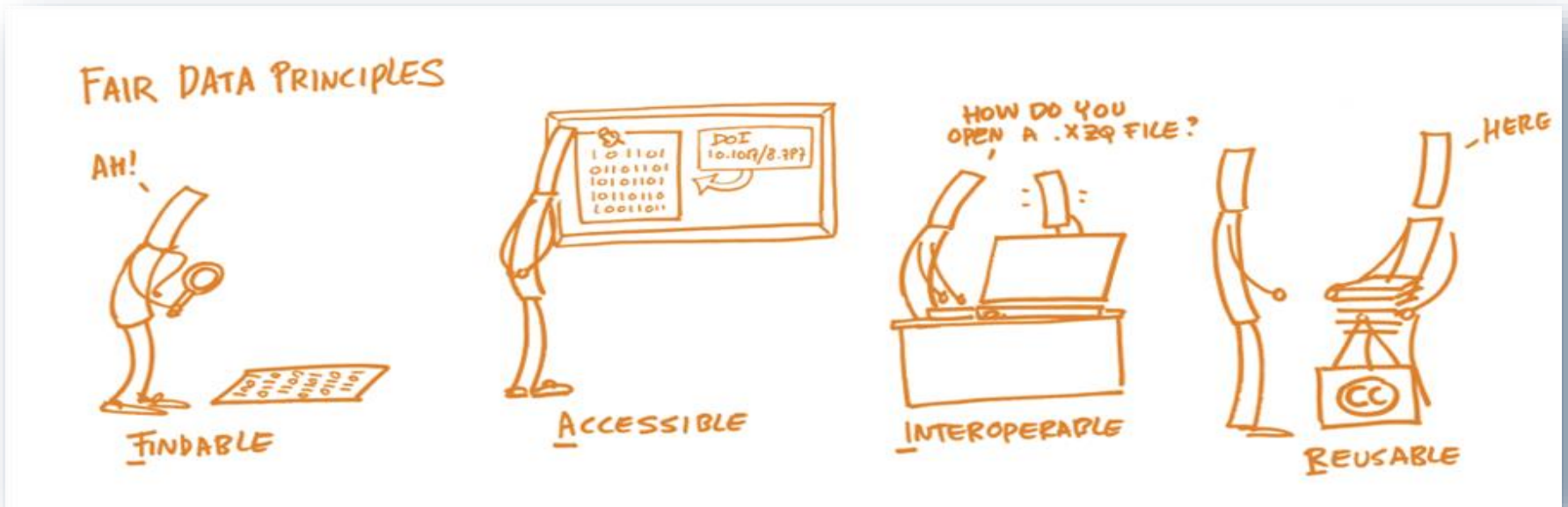
Comprende:

- **Pianificazione** iniziale
- **Gestione** quotidiana
- **Deposito e** condivisione a lungo termine



23 maggio 2023

Un FAIR Research Data Management



FINDABLE

- Ai dataset devono essere assegnati identificatori persistenti (ad esempio, DOI, Handle, URN)
- I dataset devono essere accompagnati da metadati e parole chiave significativi (contenenti il PID)

ACCESSIBLE

- Condizioni di accesso chiare – as open as possible, as closed as necessary
- I metadati devono essere apertamente accessibili indipendentemente

INTEROPERABLE

- Ove possibile, utilizzare formati e linguaggi formali e accessibili
- Includere riferimenti ad altri (meta)dati

REUSABLE

- Usare una chiara e accessibile licenza per l'utilizzo dei dati
- Includere anche la provenienza (metadati o file readme)
- Utilizzare gli standard della comunità scientifica di riferimento

Un *FAIR* Research Data Management

Quali vantaggi?



Per i ricercatori

- Maggiore visibilità e citazioni
- Opportunità di collaborazione
- Riconoscimento per la carriera
- Aiuta a prevenire errori e migliora la qualità delle analisi dei dati
- Diminuisce i rischi di "non-compliance" (etica, privacy, policy degli enti finanziatori)



Per la ricerca

- Facilita la rintracciabilità e il riuso dei dati
- Consente nuove ricerche e nuovi approfondimenti sui dati
- Protezione dei dati di valore
- Sostiene l'integrità e la riproducibilità della ricerca



Per la società

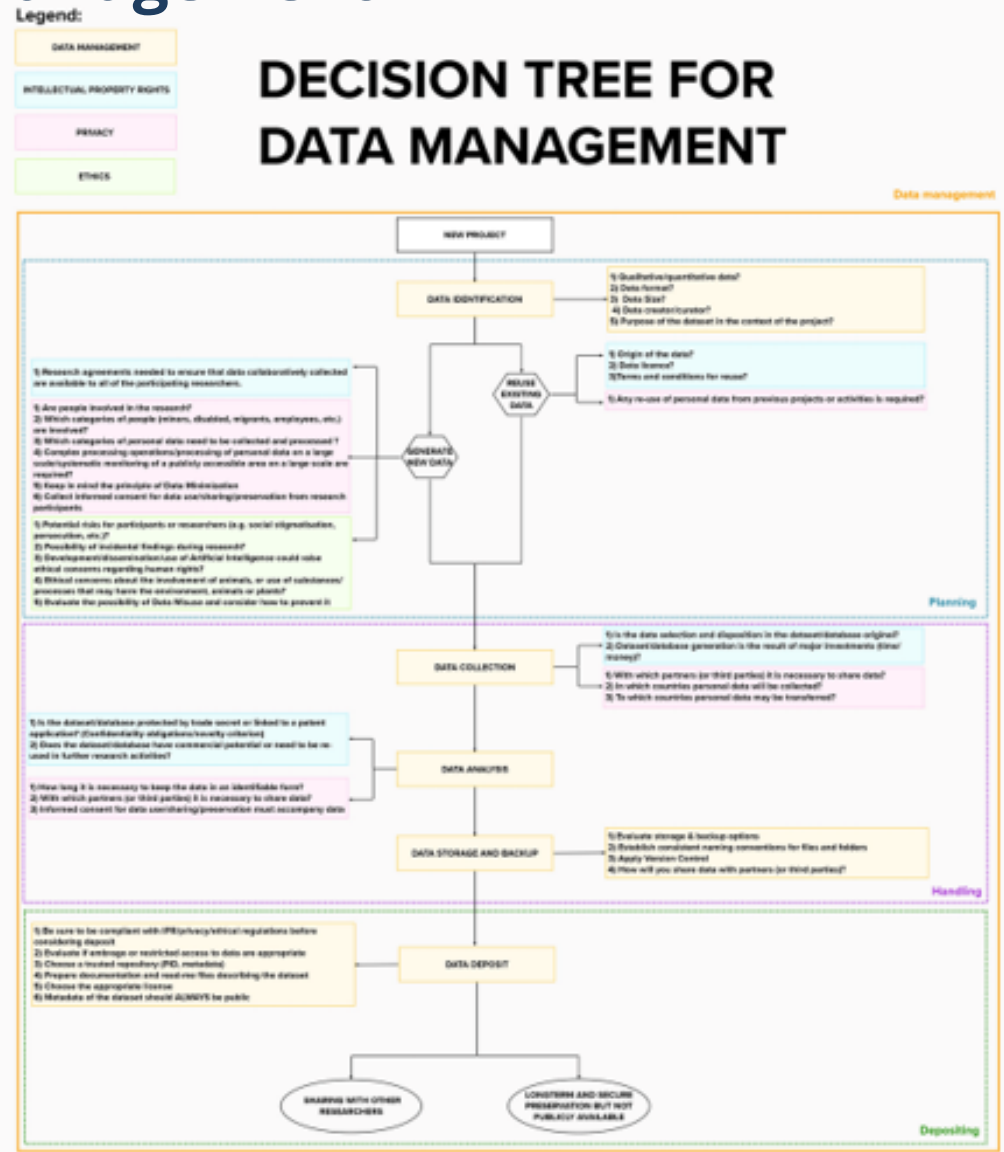
- Uso efficiente delle risorse pubbliche
- Una migliore qualità della ricerca può dare beneficio al processo decisionale
- Opportunità di business
- Opportunità di citizen science
- Maggiore trasparenza e fiducia nella scienza

Un FAIR Research Data Management

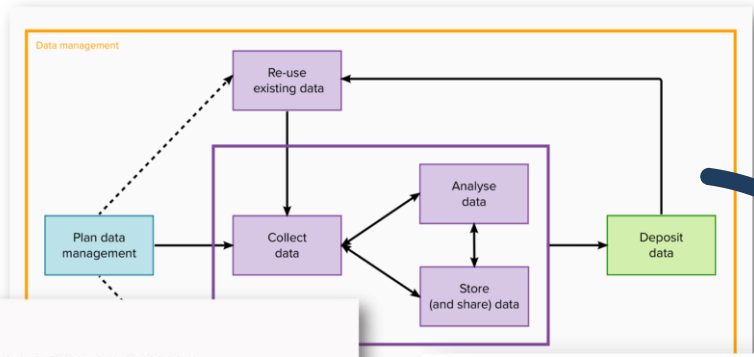
Da dove partire?

Pianificare il Research Data Management in vista del Data Management Plan

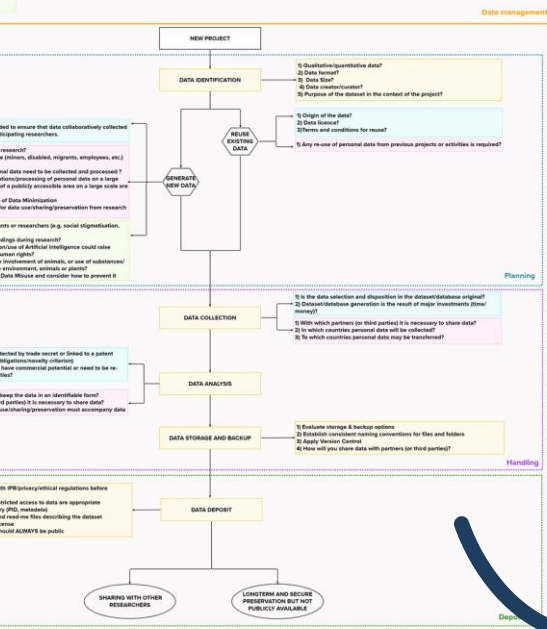
- La gestione del dato va **oltre i principi FAIR**, e riguarda le **questioni etiche e di proprietà intellettuale** per definire le azioni e misure da adottare in merito
- Bisogna farlo per **tutte le fasi del progetto e del ciclo di vita del dato**: dalla pianificazione alla gestione di ogni giorno, al deposito e condivisione
- Uno strumento come l'**albero decisionale** aiuta a porsi le giuste domande per ogni fase



II Data Management Plan (DMP)



DECISION TREE FOR DATA MANAGEMENT



Horizon Europe

Data Management Plan Template

Version 1.0
05 May 2021

1. Data Summary
2. FAIR Data
 1. Making data findable, including provisions for metadata
 2. Making data accessible
 3. Making data interoperable
 4. Increase data re-use
3. Other research outputs
4. Allocation of resources
5. Data security
6. Ethics
7. Other issues

https://ec.europa.eu/info/funding-tenders/opportunities/docs/2021-2027/horizon/temp-form/report/data-management-plan_he_en.docx

Gualandi, Bianca, Caldoni, Giulia, & Marino, Mario. (2022). Research Data Management: Data Lifecycle. Zenodo. <https://doi.org/10.5281/zenodo.7249051>



23 maggio 2023

HORIZONEUROPE@UNIB

Il DMP: contenuti

La descrizione dei dati

Descrizione dei dati: una panoramica dettagliata di ciò che si raccoglie e si genera durante la ricerca.

Informazioni	Consigli pratici: cosa descrivere
Tipologia	In base al loro contenuto (numerico, testuale, audiovisivo...), al formato (immagini, file audio...), al modo di raccolta dei dati (sperimentale, osservativo...), alla natura grezza o elaborata...
Nuovi o riutilizzati	Verificare la presenza di eventuali diritti di proprietà intellettuale da gestire in relazione ai dati che riutilizzi e affrontarli nel DMP
Metodi di generazione	Descrivi la metodologia, gli strumenti e i software (in particolare quelli proprietari) utilizzati per generare i tuoi dati. Considera protocolli computazionali (https://www.protocols.io/)
Nomenclatura	Utilizzare convenzioni di denominazione (ad esempio, nomi brevi ma descrittivi, evitare caratteri speciali e spazi, formattazione della data coerente, numero di versione...)
Formati	Scegli formati aperti per garantire che i tuoi dati rimangano accessibili e riutilizzabili
Volume	Stima necessaria per la scelta di soluzioni di storage e repository: un enorme volume di dati potrebbe comportare la necessità di preventivare costi aggiuntivi
Convalida e verifica	Strategie applicate per evitare dati imprecisi, incompleti o incoerenti. Ad esempio: convalida dell'immissione dei dati, controllo dell'intervallo di dati, controllo di scale di dati coerenti ecc.



23 maggio 2023

Il DMP: contenuti

La descrizione dei dati

Dove inserire la DESCRIZIONE dei dati nel modello DMP di Horizon Europe?

Tipi di dati

Dati appena raccolti o riutilizzati?

Come vengono creati i dati

Nome e descrizione dei dati

Formati di file dei dati

Volume dei dati (stima)

Strategie per la convalida e la verifica dei dati



1. Data Summary

2. FAIR Data

1. Making data findable, including provisions for metadata
2. Making data accessible
3. Making data interoperable
4. Increase data re-use

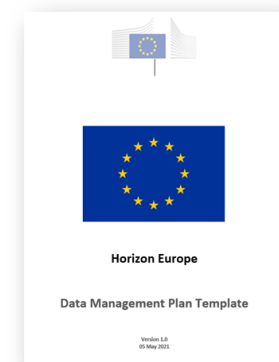
3. Other research outputs

4. Allocation of resources

5. Data security

6. Ethics

7. Other issues



Il DMP: contenuti

Aspetti etici e legali

Sia la ricerca che la gestione dei dati di ricerca possono essere influenzate da:

DIRITTI IPR

- Cosa ti è consentito fare legalmente con i dati creati da altri?
- Come vuoi che gli altri riutilizzino i tuoi dati?
- Il tuo lavoro ha un potenziale di valorizzazione commerciale (ad esempio, trasferimento tecnologico)?
- Gli accordi esistenti con terzi limitano lo sfruttamento o la diffusione dei dati che usi?

PRIVACY

- Raccoglierai dati da esseri umani?
- Come puoi proteggere la privacy dei tuoi soggetti di ricerca?
- Come puoi assicurarti che la tua ricerca rispetti la legislazione sul trattamento dei dati personali (ad esempio, il GDPR)?

ETICA

- Svolgi ricerche su animali?
- La tua ricerca ha applicazioni militari?
- La tua ricerca può influenzare la vita dei tuoi soggetti di ricerca (ad esempio, scoperte accidentali, discriminazioni)?
- Nella tua ricerca svilupperai o utilizzerai Intelligenza Artificiale?

Il DMP: contenuti

Aspetti etici e legali



As open as possible as closed as necessary:
spiegare nel DMP!

Descrivere le **strategie** scelte per affrontare questi problemi (ad esempio, anonimizzazione dei dati personali, sfruttamento dei risultati commerciali, limitazioni al riutilizzo dei dati...) e i **motivi** per cui i dati dovrebbero essere tenuti chiusi o ad accesso limitato (motivi validi: sfruttamento commerciale, riservatezza, copyright ...).

FAIR non significa "necessariamente aperto" e l'accesso aperto non influisce sulla **decisione di sfruttare commercialmente i risultati della ricerca**, ad esempio attraverso brevetti.

Se necessario, definisci un piano di sfruttamento!

Il DMP: contenuti

Aspetti etici e legali

Dove inserire le informazioni relative ad ASPETTI ETICI E LEGALI nel modello DMP di Horizon Europe?

Riutilizzare i dati in base ai diritti di proprietà intellettuale o al diritto d'autore? }

Restrizioni motivate alla disponibilità dei dati aperti }

Dati personali? Dati sensibili? Riutilizzo dei dati in base ai diritti di proprietà intellettuale o al diritto d'autore? Vuoi brevettare? }

1. Data Summary

2. FAIR Data

1. Making data findable, including provisions for metadata

2. Making data accessible

3. Making data interoperable

4. Increase data re-use

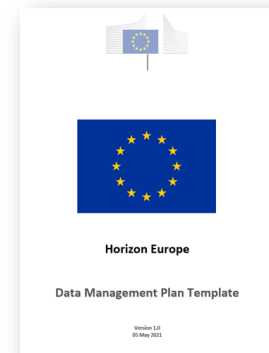
3. Other research outputs

4. Allocation of resources

5. Data security

6. Ethics

7. Other issues



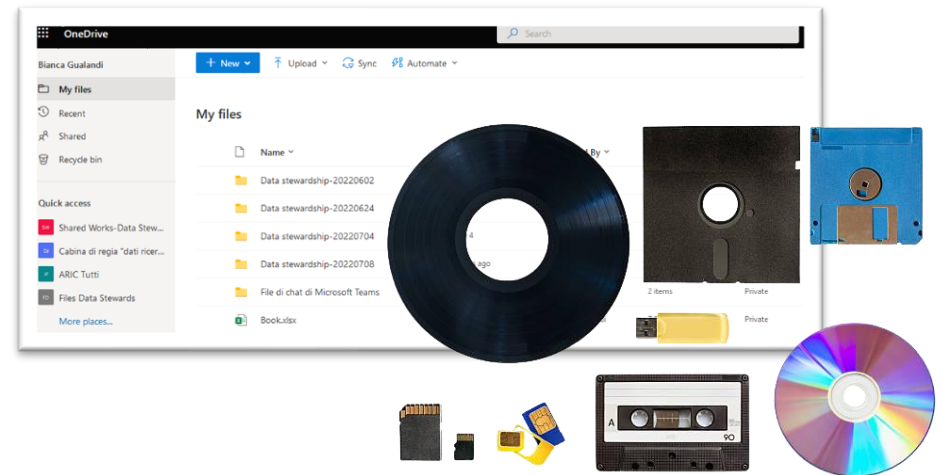
Il DMP: contenuti

Storage vs deposito dei dati

Storage: Immediatamente dopo la raccolta dei dati, può comportare la condivisione con il gruppo di ricerca. Comporta:

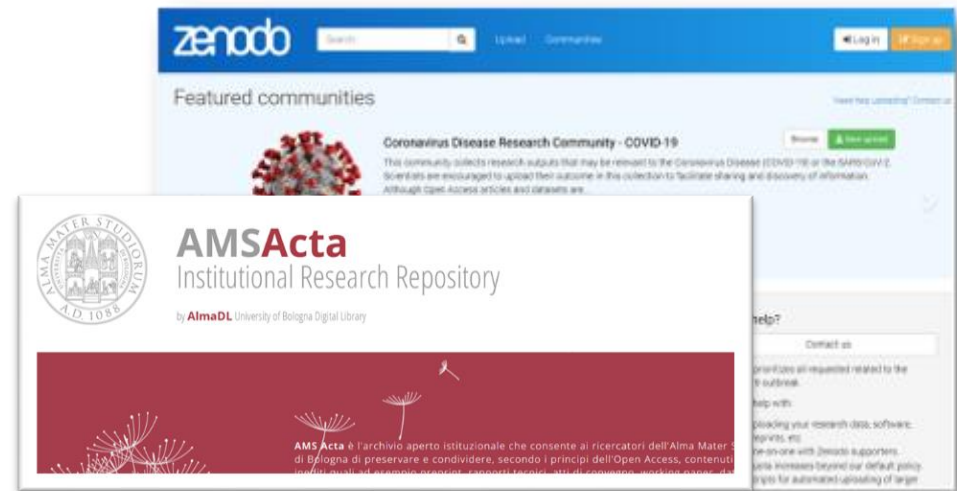
- Backup e sicurezza dei dati
- Organizzazione e denominazione di file e cartelle

→ **Scegli una soluzione di storage**



Deposito: fatto una volta o (preferibilmente) iterativamente. Comporta:

- Passaggio della responsabilità per la conservazione lungo termine dei dati dal ricercatore al repository
- Scelta della licenza e del livello di accesso
- → **Scegliere un repository di dati**



23 maggio 2023

HORIZONEUROPE@UNIB

Il DMP: contenuti

Storage

Sistemi di storage affidabili consentono: procedure di **backup** automatiche; modalità di **accesso controllato** anche in caso di collaborazione; accesso alla **cronologia delle versioni** dei file; e regole di base sulla **denominazione** dei file e sulla **strutturazione** delle cartelle.

Consigli pratici:

- Se hai bisogno di collaborare con altri ricercatori, scegli uno strumento adatto alla **collaborazione**;
- Scegli **più di un sistema** di storage: non memorizzare i dati solo sul tuo computer;
- Quando si memorizzano **informazioni sensibili o dati personali**, definire le condizioni di accesso, per quanto tempo conservarle e quali misure saranno messe in atto per proteggere questi dati (ad esempio, "AMNESIA" per l'anonimizzazione dei dati, "VERACRYPT" per criptare...);
- Considerare i **costi** per lo storage e il backup dei dati durante il progetto di ricerca e come saranno coperti.



no title, CC-BY, by Auke Herrema

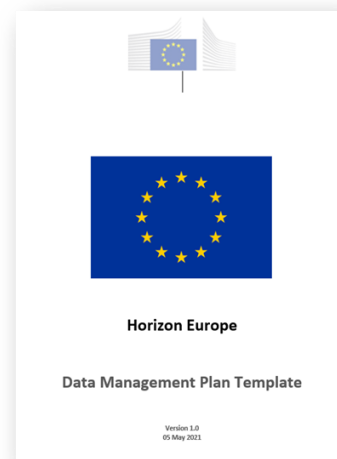
Il DMP: contenuti

Storage

Dove inserire le informazioni relative allo STORAGE SICURO DEI DATI nel modello DMP di Horizon Europe?

1. Data Summary
2. FAIR Data
 1. Making data findable, including provisions for metadata
 2. Making data accessible
 3. Making data interoperable
 4. Increase data re-use
3. Other research outputs
4. Allocation of resources
5. **Data security**
6. Ethics
7. Other issues

Soluzioni scelte per lo storage e il backup, misure di sicurezza



Il DMP: contenuti

Il deposito dei dati

I repository considerati «trusted» sono progettati per la gestione e la conservazione dei dati secondo i principi di interoperabilità e accessibilità:

- **Prevedono norme per l'interoperabilità:** ad es. OAI-PMH
- Al momento del deposito, forniscono in associazione ai dati:
 - **Identificatori persistenti (PID)**
 - **Metadati e documentazione**
 - **Licenza e condizioni di accesso**

DOI:

DOI 10.3389/fdata.2019.00043

Keyword(s):

open science, open data, open access, open scholarship,
European science policies, artificial intelligence

Published in:

Frontiers in Big Data: 2 (43).

License (for files):

 Creative Commons Attribution 4.0 International

```
"description": "Codemeta defines a 'JSON-LD' format for describing software metadata.\n",\n"name": "codemeta: Generate CodeMeta Metadata for R Packages",\n"codeRepository": "https://github.com/codemeta/codemeta",\n"issueTracker": "https://github.com/codemeta/codemeta/issues",\n"license": "https://spdx.org/licenses/MIT",\n"version": "0.1.0",\n"programmingLanguage": {\n  "@type": "ComputerLanguage",\n  "name": "R",\n  "version": "3.4.0",\n  "url": "https://r-project.org"\n},\n"runtimePlatform": "R version 3.4.0 (2017-04-21)",\n"author": [\n  {\n    "@type": "Person",\n    "givenName": "Carl",\n    "familyName": "Boettiger",\n    "email": "cboettig@gmail.com",\n    "@id": "http://orcid.org/0000-0002-1642-628X"\n  }\n]
```

Il DMP: contenuti

Il deposito dei dati

Un **identificatore persistente (PID)** è un riferimento univoco e unico di lunga durata a una risorsa. La risorsa può essere:

- un **oggetto** (pubblicazioni, dati, software...) per i PID solitamente sono: **URN, DOI, ARK, Handle**
- una **persona** (ricercatori, autori, collaboratori...) per cui i PID solitamente sono: **ORCID, ISNI**
- un **ente** (di ricerca, finanziatore, un'organizzazione...) per cui il PID solitamente è **ROR**



Lo scopo principale del PID è fornire le informazioni necessarie per **identificare, verificare e localizzare** in modo affidabile la risorsa.

Molti archivi assegnano un PID quando un oggetto viene **depositato**.

Un PID può essere collegato a un insieme di **metadati** che descrivono l'oggetto.

<https://doi.org/10.5281/zenodo.5018216>

Resolver Service

Directory indicator

Prefix

Suffix



23 maggio 2023

HORIZONEUROPE@UNIB

Il DMP: contenuti

Il deposito dei dati

Metadati: informazioni strutturate che descrivono, localizzano e facilitano l'uso o la gestione di un'altra risorsa - sia per le persone che per le macchine.

Mentre i metadati **descrittivi** forniscono informazioni sul **contenuto** di un oggetto digitale (ad esempio, le parole chiave per descrivere la trama di un libro), i metadati **strutturali** sono utilizzati per specificare le **relazioni formali** tra i componenti di un oggetto digitale (struttura interna) e tra diversi oggetti digitali (struttura esterna) (ad esempio, i capitoli di un libro).

I metadati possono essere **generalisti** o **specifici** del dominio disciplinare:

- Dublin Core è un esempio di schema di metadati generico molto diffuso;
- Per trovare metadati specifici di area esistono diverse risorse online: FAIRsharing, Research Data Alliance, DCC Metadata Standard Selector



Il DMP: contenuti

Il deposito dei dati

Oltre ai metadati, potrebbe essere utile depositare anche altra **documentazione** che fornisca informazioni sulla **provenienza** dei dati, sui **metodi** di raccolta, riuso o generazione dei dati e sui processi di **garanzia di qualità** nella generazione dei dati.

Un esempio di documentazione è il **file README**, che solitamente contiene informazioni sulla struttura e organizzazione dei dataset depositati, oltre a informazioni sulle metodologie.



Regola generale: per migliorare l'interoperabilità e il riutilizzo dei dati è necessario raccogliere e depositare *tutte le informazioni di cui gli utenti avrebbero bisogno per comprendere e riutilizzare i dati* (senza aver partecipato al progetto di ricerca!).

Il DMP: contenuti

Il deposito dei dati

Rilasciare le pubblicazioni e i dati con una **licenza** chiara e permissiva permette:

- agli **altri utenti** di trovare, leggere e riutilizzare i materiali;
- **ai proprietari dei dati** di controllare il riuso dei dati e stabilirne le condizioni.

CREATIVE COMMONS LICENSES		COPY & PUBLISH	ATTRIBUTION REQUIRED	COMMERCIAL USE	MODIFY & ADAPT	CHANGE LICENSE
	PUBLIC DOMAIN	✓	✗	✓	✓	✓
	CC BY	✓	✓	✓	✓	✓
	CC BY-SA	✓	✓	✓	✓	✗
	CC BY-ND	✓	✓	✓	✗	✗
	CC BY-NC	✓	✓	✗	✓	✓
	CC BY-NC-SA	✓	✓	✗	✓	✗
	CC BY-NC-ND	✓	✓	✗	✗	✗

Legend:

- You can redistribute (copy, publish, display, communicate, etc.)
- You have to attribute the original work
- You can use the work commercially
- You can modify and adapt the original work
- You can choose license type for your adaptations of the work.

I dati sono **sempre coperti da copyright**, anche quando le informazioni sulla licenza non sono incluse nei dati stessi.

Se i dati non sono accompagnati da una licenza è necessario **ottenere l'autorizzazione** del proprietario per copiarli, condividerli o pubblicarli in altro modo.

Suggerimento pratico:

Usare un "Licence Chooser" può servire a determinare quale licenza è più adatta al proprio dataset, ad esempio: [Creative Commons License Chooser](https://creativecommons.org/choose/), [Ufal Public License Selector](https://ufal.github.io/public-license-selector/), choosealicense.com, [Conformant Licenses to Open Definition](https://conformantlicenses.com/)

Il DMP: contenuti

Il deposito dei dati

La **scelta di un buon repository per i dati** è il primo passo per rendere i dati FAIR e facili da trovare sia per i ricercatori che per i motori di ricerca.

Per cercare un repository adatto: **Re3data**, <https://www.re3data.org>

Un recente lavoro di censimento dei repository di dati esistenti può aiutare a sceglierne uno in base ai criteri più adatti: <https://doi.org/10.5281/zenodo.7728016>



Per facilitare la reportistica continua dei dati e la pubblicazione, verificare se il repository scelto è indicizzato dall'aggregatore OpenAIRE+, <https://www.openaire.eu>



AMSActa
Institutional Research Repository

by **AlmaDL** University of Bologna Digital Library

[AMSActa](https://www.amsacta.unibo.it) è il repository istituzionale di Ateneo

Il DMP: contenuti

Il deposito dei dati

Dove inserire le informazioni relative ai REPOSITORY DATI nel template DMP di Horizon Europe?

Attribuzione di identificatori persistenti, metadati standard.]

Standard, ontologie e vocabolari, documentazione dei dati.]

Scelta di una licenza, documentazione dei dati.]

1. Data Summary

2. FAIR Data

1. Making data findable, including provisions for metadata

2. Making data accessible

3. Making data interoperable

4. Increase data re-use

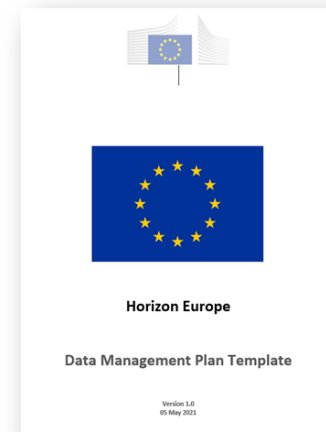
3. Other research outputs

4. Allocation of resources

5. Data security

6. Ethics

7. Other issues



Il DMP: contenuti

Altri output della ricerca

Risultati della ricerca diversi dalle pubblicazioni scientifiche e dai dati della ricerca.

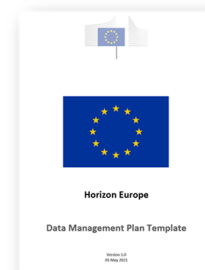
- Sia digitali (ad esempio, workflow, protocolli, modelli, notebook elettronici, ecc.), che **fisici** (es. nuovi materiali, hardware, modelli 3D, opere d'arte, libri, ecc.) o **intangibili** (ad esempio, eventi)
- Gestiti, condivisi, resi disponibile in linea con i principi FAIR (ad esempio, anche se un oggetto di ricerca tangibile non può essere digitalizzato, la metodologia per svilupparlo potrebbe essere digitalizzata e depositata)



Software

- Repository dedicati: Software Heritage (fornisce PID, consente l'archiviazione del software, la raccolta periodica da GitHub/GitLab)
- Licenze specifiche (non Creative Commons)

1. Data Summary
2. FAIR Data
 1. Making data findable, including provisions for metadata
 2. Making data accessible
 3. Making data interoperable
 4. Increase data re-use
3. **Other research outputs**
4. Allocation of resources
5. Data security
6. Ethics
7. Other issues



23 maggio 2023

Il DMP: contenuti

Allocazione delle risorse

I **costi monetari e i costi di tempo / persone** devono essere presi in considerazione per la gestione dei dati FAIR e la stesura del DMP.

Per attività come: **pulizia dei dati; trascrizione; anonimizzazione/pseudoanonimizzazione; traduzione; preparazione della documentazione del dataset; storage temporaneo e a lungo termine.**

Esistono **strumenti** per la valutazione dei costi:



<https://www.openaire.eu/how-to-comply-to-h2020-mandates-rdm-costs>

UK Data Service



<https://www.ukdataservice.ac.uk/manage-data/plan/costing>



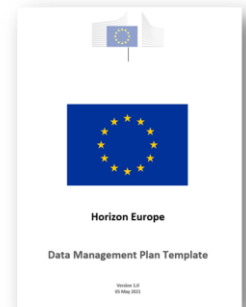
<https://www.tudelft.nl/en/library/research-data-management/rplan/data-management-costs>



National Coordination Point
Research Data Management

https://www.lcrdm.nl/files/lcrdm/2020-04/RDM%20and%20Costs_v20160218_EN.pdf

1. Data Summary
2. FAIR Data
 1. Making data findable, including provisions for metadata
 2. Making data accessible
 3. Making data interoperable
 4. Increase data re-use
3. Other research outputs
4. **Allocation of resources**
5. Data security
6. Ethics
7. Other issues



Utrecht University

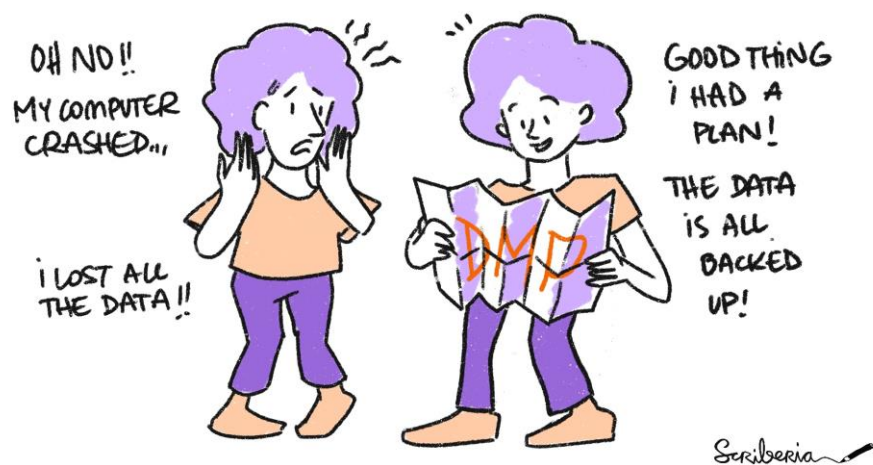
<https://www.uu.nl/en/research/research-data-management/guides/costs-of-data-management>



23 maggio 2023

HORIZONEUROPE@UNIB

Il Data Management Plan (DMP)



Il Data Management Plan è una **tabella di marcia** per **identificare** i dati di ricerca e gestirli correttamente, sviluppando procedure in anticipo per un lavoro di ricerca efficiente e coerente, risparmiando tempo in seguito (es. identificando i ruoli all'interno di un gruppo). Aiuta anche a rendere i materiali di ricerca visibili e citabili per massimizzare l'impatto del lavoro.

Si tratta di documentare le tue **intenzioni** su:

- Come **evitare problemi** come duplicazione, perdita di dati e violazioni della sicurezza
- Come stai adempiendo a tutti i tuoi **obblighi** legali ed etici

È un documento «vivente»: se le tue scelte e le tue pratiche di ricerca cambiano, si può sempre aggiornare!

Dai dati al DMP: gli obblighi di Horizon Europe

- **Gestione** dei dati in linea con i **principi FAIR**
 - Depositare il prima possibile i dati della ricerca in quelli che sono considerati "*repository trusted*" seguendo il principio **"as open as possible, as closed as necessary"**
 - I dati dovrebbero essere depositati con una licenza **Creative Commons CC BY o CC0**
 - I dati depositati dovrebbero essere accompagnati da adeguati **metadati aperti** (obbligatoria la licenza **CC0**)
- **Definire e aggiornare** un Data Management Plan (**DMP**)

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 (CC 0) 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**
- provide information via the repository about any research output or any other tools and instruments needed to re-use or validate the data.

Metadata of deposited data must be open under a Creative Commons Public Domain Dedication (CC 0) or equivalent (to the extent legitimate interests or constraints are safeguarded), in line with the FAIR principles (in particular machine-actionable) and **provide information at least about the following: datasets (description, date of deposit, author(s), venue and embargo); Horizon Europe or Euratom funding; grant project name, acronym and number; licensing terms; persistent identifiers for the dataset, the authors involved in the action, and, if possible, for their organisations and the grant. Where applicable, the metadata must include persistent identifiers for related publications and other research outputs.**

Communication, dissemination, open and visibility – article 17, annex 5, grant agreement

Technical reports: quali **scelte** prese riguardo questi obblighi

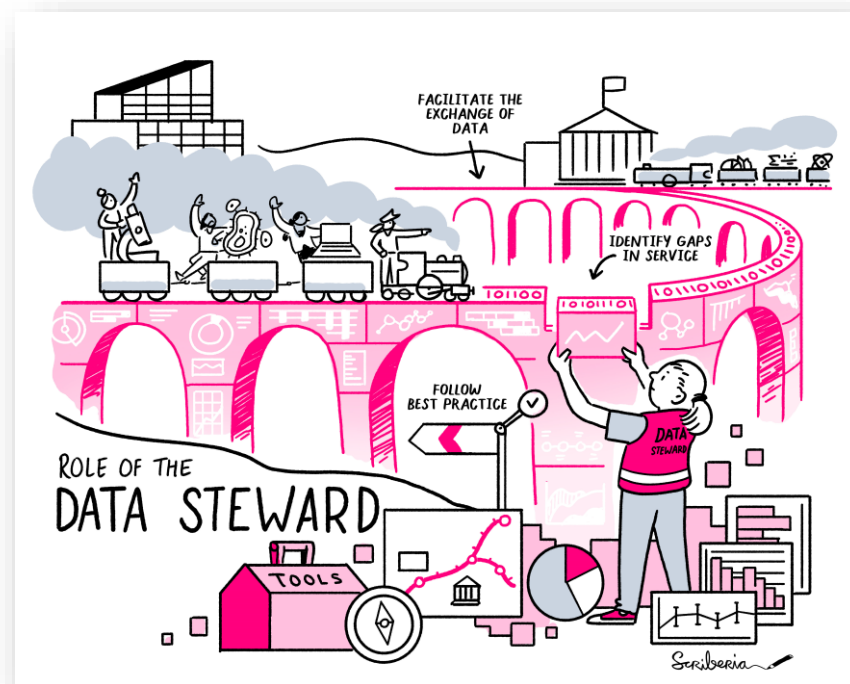


23 maggio 2023

Il data steward

Supporto consulenziale per i ricercatori nella gestione dei **dati** secondo principi FAIR.

Figura professionale che nasce nel **contesto accademico europeo** dal 2007 circa.



The Turing Way Community, & Scriberia. (2023). Illustrations from The Turing Way: Shared under CC-BY 4.0 for reuse. Zenodo. <https://doi.org/10.5281/zenodo.7587336>

Il data steward: chi è, cosa fa e perché è fondamentale, 23 dicembre 2021, Valentina Pasquale <https://www.icdi.it/it/news/121-il-data-steward-chi-e-cosa-fa-e-perche-e-fondamentale>

Wildgaard, Lorna, Vlachos, Evgenios, Nondal, Lars, Larsen, Asger Væring, & Svendsen, Michael. (2020).

National Coordination of Data Steward Education in Denmark: Final report to the National Forum

for Research Data Management (DM Forum) (Version 1). Zenodo. <https://doi.org/10.5281/zenodo.3609516>

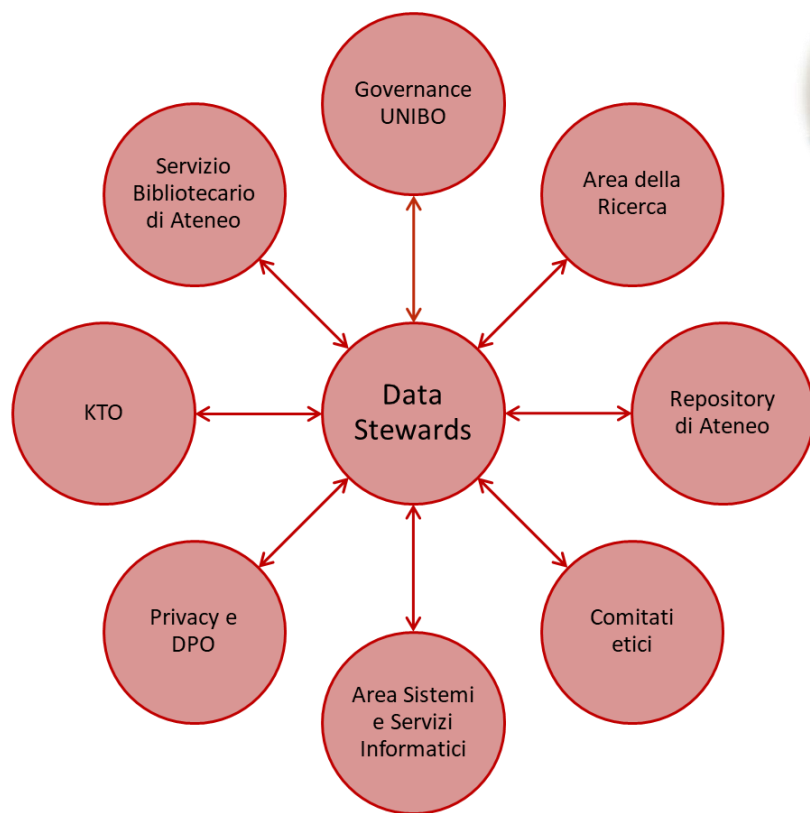


23 maggio 2023

HORIZONEUROPE@UNIB

Data Steward in Unibo

Figura professionale



GIULIA CALDONI

Area biomedica



SARA COPPINI

Area tecnologica



BIANCA GUALANDI

Area umanistica



MARIO MARINO

Area Sociale

- Esperto di **dominio**
- Professionista competente nella **pianificazione e gestione responsabile** dei dati durante tutto il ciclo di vita
- Punto di **contatto unico** per ricevere indicazioni sulla gestione dei dati della ricerca
- Servizio centralizzato presso l'Area della Ricerca (**ARIC**)
- Per qualsiasi dubbio: **aric.datasteward@unibo.it**

Data Steward in Unibo

Principali azioni di supporto

- Supporto ai ricercatori nel contesto dei progetti finanziati (Horizon Europe in primis) e in tutte le attività di ricerca con condivisione di **materiale informative e strumenti di supporto**
 - nella gestione **FAIR** dei dati ed **RDM**
 - nella stesura del **Data Management Plan**
- Supporto in **coordinamento con gli altri servizi di Ateneo** inerenti (Privacy, KTO, Repository di Ateneo, Biblioteche, etc.)



Il processo di scrittura del DMP: un esempio di pianificazione col supporto *strong* dei data steward

I) Presentazione obblighi HE su FAIR principles e Open Science e introduzione a DMP e strumenti utili

II) Identificazione dei dataset (tramite analisi *proposal* e GA), individuazione delle problematiche principali legate a questi dati

III) Raccolta delle informazioni sui dataset dai partner per i progetti coordinati

IV) Eventuali ulteriori incontri, presentazioni, chiarimenti e raccolta di informazioni

V) Revisione della bozza del DMP, eventuali ulteriori chiarimenti e correzioni

VI) Approvazione e consegna del DMP

DMP obbligatorio:

- Mese **6**
- Progetti coordinati:
 - Metà progetto (mese **18+**)
 - Fine progetto (mese **36+**)
- Progetti monobeneficiari: l'officer può rendere aggiornamenti in qualsiasi momento



23 maggio 2023

Come vi possiamo supportare

Durante il primo incontro, che di solito viene fatto con il PI e con la persona che si occuperà redigere il DMP, vengono mostrati e successivamente forniti i seguenti materiali:

1. il **template** per il DMP HE fornito dalla Commissione Europea, per avere un'idea della struttura e delle richieste specifiche del documento
2. una **presentazione di Power Point** in cui sono riassunti gli obblighi contrattuali di HE, insieme ad un focus su obiettivi e contenuti del DMP
3. **link a un Microsoft Form** che ogni partner dovrà compilare per ogni dataset che prevede di produrre/ha prodotto, contenuto in un file Word con la proposta per la mail da inviare ai partner (SOLO PROGETTI COORDINATI)
4. un **modello di DMP Unibo sul template HE commentato** dai Data Steward come guida nella stesura del documento
5. spiegazione e link all'**Albero Decisionale**

Horizon Europe: Open Science Requirements for Publications and Data

Alma Mater Studiorum – Università di Bologna

ARIC – Area Servizi alla Ricerca

ABIS – Area Biblioteche e Servizi allo Studio



1. Data Summary

[Insert project acronym] aims at ...

The project will reuse data from the following source ... Reuse has been favored as ...

[Insert project acronym] will produce different types of data by using different methodologies. Research teams have agreed to convert research data from proprietary formats to well-known and documented open formats in order to facilitate accessibility and reusability (Table 1).

Table 1 - Summary of data formats

Type of data	Formats used during data processing	Formats for sharing, reuse and preservation

Data sets to include in DMP - form for project partners

In order to create the best possible Data Management Plan (DMP) for our project, we kindly ask you to fill in the attached questionnaire, providing details on the data that you collect and/or generate as a partner. Detailed guidance is provided for each question below.

Please bear in mind:

- A data set is defined as a **named collection of data units with the same focus and scope**
- **Each dataset needs to be described separately**; when you finish answering all questions, you will find a button to start the questionnaire again.
- If you have **datasets that are similar** to each other, after filling out the form the first time click on **"Print or get PDF of answers"** (at the end of the form) -> **"Save as PDF"** to get a trace from which to copy/paste the same fields and save time.
- We are aware that many questions might appear too detailed if you are at the initial stages of data collection. Nevertheless, we kindly ask you to be **as specific as the project stage allows** when describing what you **plan to do**.

Thank you!

Autore
Descrivere brevemente lo scopo del progetto

Autore
DA QUESTIONARIO:
18. The data set is:
17. Why (not) re-using existing data?

Autore
DA QUESTIONARIO:
20. Original provenance of data
21. Data types
22. Data formats

(descrivere discorsivamente le informazioni ricapitolate in table 1)

Autore
QUESTIONARIO:
Data types
Data formats



23 maggio 2023

HORIZONEUROPE@UNIB

Il processo di scrittura del DMP: un esempio di pianificazione col supporto *light* dei data steward

I) Presentazione obblighi europei su FAIR principles e Open Science e introduzione a DMP e strumenti utili

II) Prima bozza di DMP con solo parti partner Unibo

III) Modifiche e revisioni solo per il gruppo Unibo nel caso HE partner con DMP, e nel caso di altri progetti non HE

**Il supporto è
incentrato sul
gruppo UniBO**



23 maggio 2023

HORIZONEUROPE@UNIBO

Punti chiave: il dato e RDM

- **Ogni tipo di ricerca genera o riutilizza dati:** che cos'è il dato dipende dall'area disciplinare a cui appartiene il progetto di ricerca.
- Il **dato** è un asset di valore per il ricercatore e la comunità se gestito e valorizzato correttamente attraverso i **principi FAIR** e un buon **Research Data Management**.
- I dati FAIR **non sono necessariamente dati open**, basta seguire la regola «*as open as possible as closed as necessary*».
- La scelta di un **repository trusted** con le caratteristiche adatte per la conservazione a lungo termine del tuo tipo di dati specifico, è un ottimo punto di partenza per essere in linea principi FAIR.
- I dati depositati hanno un valore per le tue future ricerche e per la comunità scientifica solo se accompagnati da una **documentazione** sulla provenienza, uso e metodologie di generazione e gestione dei dati stessi.
- Oltre ai principi FAIR, nella gestione del dato bisogna sempre considerare le questioni che riguardano **etica, privacy e proprietà intellettuale**, in ogni fase del ciclo di vita dei dati e del progetto.



23 maggio 2023

Punti chiave: il DMP

- Non è solo un obbligo ma anche uno **strumento di pianificazione**, una *roadmap* per evitare problemi ed essere sicuri di ottemperare agli obblighi legali ed etici.
- È uno strumento utile per tenere traccia delle scelte fatte in ambito della ricerca e per svolgere un lavoro ottimale **anche se non richiesto** come obbligo dall'ente finanziatore.
- È un documento in continua evoluzione: documenta le **strategie** e le tue **intenzioni**, può essere sempre aggiornato se qualche cosa cambia!
- Nei DMP di inizio progetto l'obiettivo non è avere le risposte a tutte le domande, ma di **considerare le domande giuste da porsi** per la corretta gestione del dato.
- È un documento utile specialmente per i **progetti collaborativi**, perché permette di mettere in chiaro ruoli e responsabilità di tutti i ricercatori e partner coinvolti.



23 maggio 2023

Punti chiave: i data steward

- Il servizio dei data steward è pensato per supportare i ricercatori Unibo
 - nella **gestione del dato della ricerca**, sia in linea con i principi FAIR che nel rispettare gli obblighi in ambito etica, privacy e diritti di proprietà intellettuale
 - per semplificare e promuovere la **compilazione di DMP**
- Con **diverse modalità di supporto**: ci siamo per un supporto strutturato e pianificato, ma anche solo per rispondere a qualche domanda!
- Sempre all'interno di una **rete di supporto e interazione** tra data steward, project manager, research manager, principal investigator, ricercatori e le varie expertise di ateneo (biblioteche, ufficio privacy, KTO...)
- ... E sviluppando **strumenti** per il supporto, come il Data Life Cycle e il Decision Tree.



Il supporto disponibile a UniBo



Per supporto riguardo la gestione dei dati potete rivolgervi ai data stewards che lavorano in ARIC – area della ricerca: aric.datasteward@unibo.it



Per supporto all'open access delle pubblicazioni potete rivolgervi al servizio di supporto OA della vostra Biblioteca di riferimento: <https://sba.unibo.it/it/almadl/open-access-e-open-science/servizio-di-supporto-delle-biblioteche>



Per supporto riguardo la proprietà intellettuale dei risultati di ricerca UniBo potete rivolgervi al Knowledge Transfer Office: kto@unibo.it



Per supporto sulle questioni relative alla privacy potete scrivere all'indirizzo: privacy@unibo.it

Per approfondimenti su Horizon Europe

[Horizon Europe reference documents](#)

[Program Guide of Horizon Europe](#)

[Annotated Model Grant Agreement \(AGA\)](#)

[ERC Managing your project > Open Science](#)

[MSCA Work Programme](#)



23 maggio 2023

HORIZONEUROPE@UNIB



ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Grazie per l'attenzione

sara.coppini2@unibo.it

HORIZONEUROPE@UNIB 