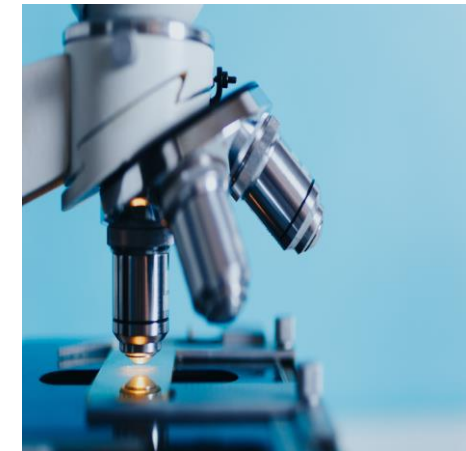


GESTIRE I DATI DELLA RICERCA: tutto ciò che c'è da sapere



LEZIONE 3

Bianca Gualandi
bianca.gualandi4@unibo.it

Data Steward area Umanistica, Area servizi alla ricerca (ARIC), Alma Mater Studiorum - Università di Bologna



19 Ottobre 2023



DURANTE L'INCONTRO DI OGGI...

- Per favore tenete i microfoni spenti durante la presentazione.
- Sentitevi liberi di accendere microfono e videocamera nei momenti di Q&A.
- Sentitevi liberi di porre le vostre domande in chat in qualsiasi momento, risponderemo durante il Q&A!
- Abbiamo previsto all'interno di questa lezione due pause, che seguiranno i momenti Q&A.

GESTIRE I DATI DELLA RICERCA: TUTTO CIÒ CHE C'È DA SAPERE



Modulo formativo diviso in tre incontri.

Unico pre-requisito è avere una propria esperienza di ricerca.



LE TEMATICHE CHE AFFRONTEREMO OGGI

1

IL DATA MANAGEMENT PLAN
Definizioni ed esempi.

2

COSA SCRIVERE IN UN DATA MANAGEMENT PLAN
Come compilare le sezioni che compongono un DMP.

CHI SIAMO? DATA STEWARDS @UNIBO



Il progetto Data Steward @Unibo mira al rafforzamento del supporto alla gestione dei dati della ricerca:

CHI SONO I DATA STEWARDS? Figure di supporto per le tematiche di gestione FAIR dei dati della ricerca e stesura del Data Management Plan ai team ricerca di UNIBO

COSA FANNO? Supportano i ricercatori nella gestione FAIR dei dati (research data management) e nella stesura del Data Management Plan, principalmente nel contesto dei progetti Horizon Europe
Supportano la Governance di Ateneo nella promozione di Open Science

QUAL È IL LORO PROFILO? Background scientifico (esperti di dominio) con esperienza in data management, FAIR principles e conoscenza di pratiche di OS



RIPRENDIAMO LE FILA DEL DISCORSO...

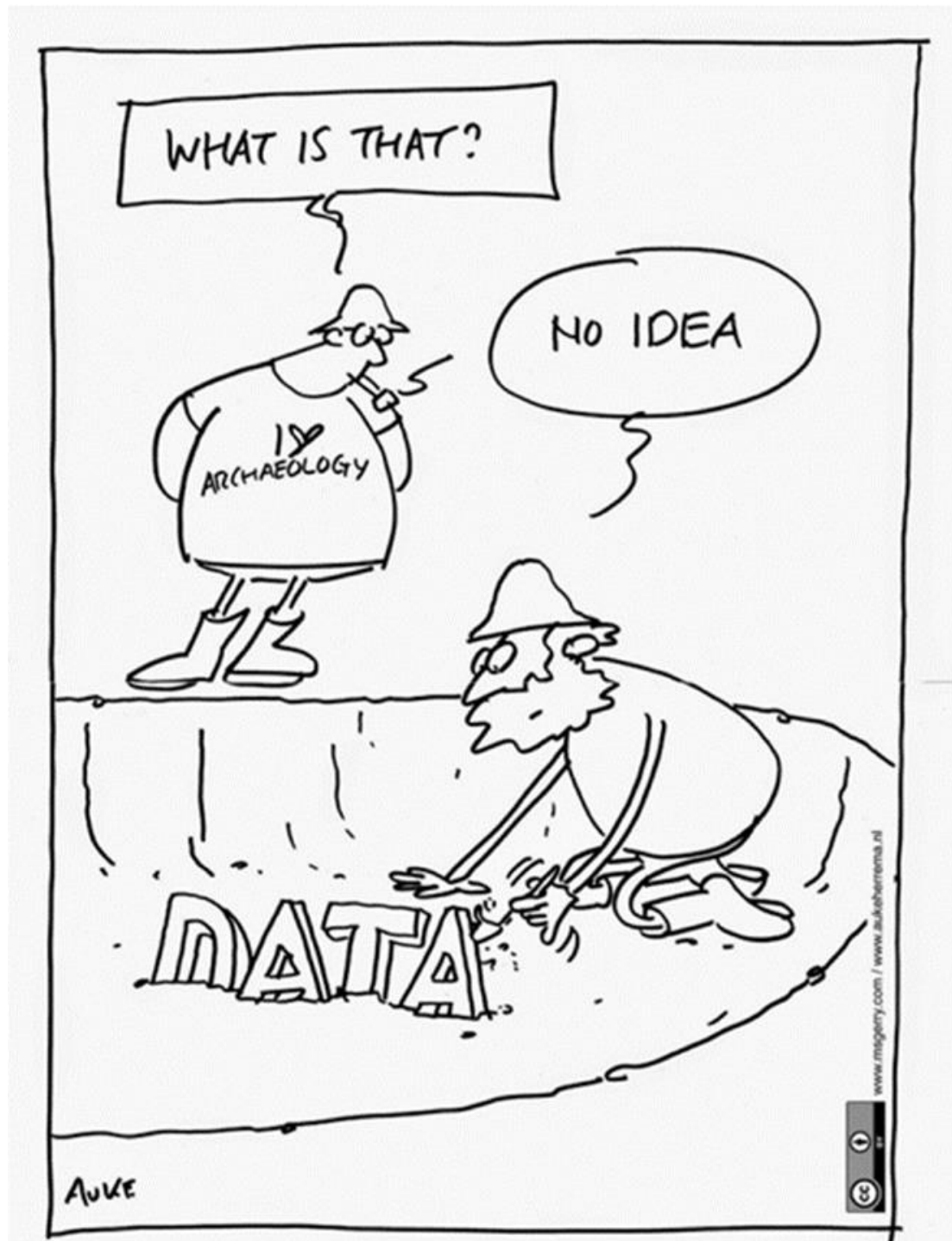


Image credit: <http://aukeherrema.nl> CC-BY

- Esistono diverse definizioni di dato e variano in base alla disciplina di riferimento.
- **Dato è tutto ciò che è alla base** dei ragionamenti a supporto **di una tesi di ricerca**.
- Una collezione di dati accomunati dallo stesso obiettivo è chiamata **dataset**.
- Il dato ha un valore intrinseco come asset della ricerca, per conservarlo deve essere **gestito correttamente**.
- La gestione corretta è un processo che permea tutte le fasi del **ciclo di vita del dato** della ricerca.

RIPRENDIAMO LE FILA DEL DISCORSO...

Abbiamo visto che molti di noi lavorano con il **patrimonio culturale, digitale o digitalizzato** (es. facsimile, modelli 3D, audio, video).



Screenshot da «Broke in China» un cortometraggio comico americano del 1927 diretto da Edward F. Cline e con Ben Turpin. [https://commons.wikimedia.org/wiki/File:Broke_in_China_\(1927\).webm](https://commons.wikimedia.org/wiki/File:Broke_in_China_(1927).webm)



Busto Aldrovandi. 3D model on SketchFab. FrameLAB - Digital storytelling & Multimedia Lab - University of Bologna. <https://sketchfab.com/3d-models/busto-aldrovandi-e100a2c2dc2a44aef94f9d0c5301fe668>



Mafatih al-jinan wa-masabih al-janan. Bologna, Biblioteca Universitaria, ms. 2970. CC BY-NC-ND 4.0. AMS Historica, <https://amshistorica.unibo.it/archivio/0001686/000010.jpg>

RIPRENDIAMO LE FILA DEL DISCORSO...

Altri lavorano con il **linguaggio**, in forma **orale o scritta**.

Voices from Ravensbrück: The Value of Multilingual Oral History

Contributors: Silvia Calamai, Arjan van Hessen, Stefania Scagliola, Christoph Draxler, Henk van den Heuvel

Submitted by Karina Berger on 13 December 2022

The Project

The Voices from Ravensbrück project is the result of the curation and selection of multilingual oral history interviews with survivors from the Ravensbrück concentration camp for women. Organised and funded in the context of [CLARIN's Resource Families](#), this interdisciplinary project brings together multilingual experts in the fields of linguistics, speech technology, speech corpora curation, phonetics and history.

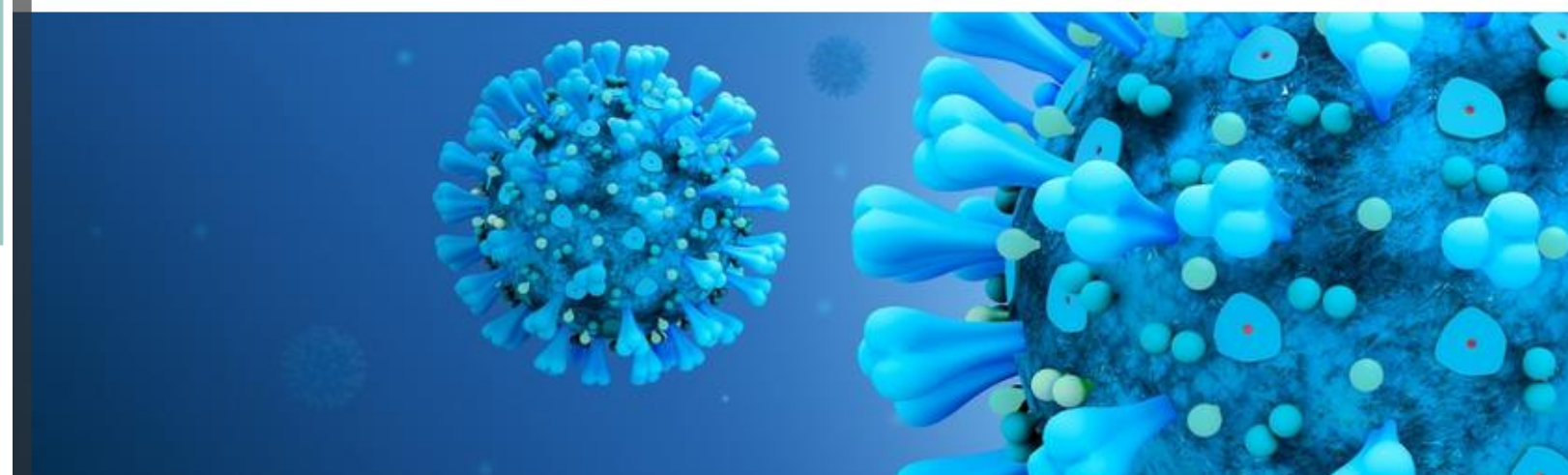
Using a Monitor Newspaper Corpus to Trace Changing Language as a Result of COVID-19

Contributors: Koenraad De Smedt

Submitted by Karina Berger on 9 December 2021

The Project

This project illustrates the possibility to trace, almost in real time, changes in language in response to a crisis using a [monitor newspaper corpus](#). The study 'Contagious "Corona" g by Journalists in a CLARIN Newspaper Monitor Corpus' examines the linguistic occurred in the Norwegian language during the first wave of the COVID-19

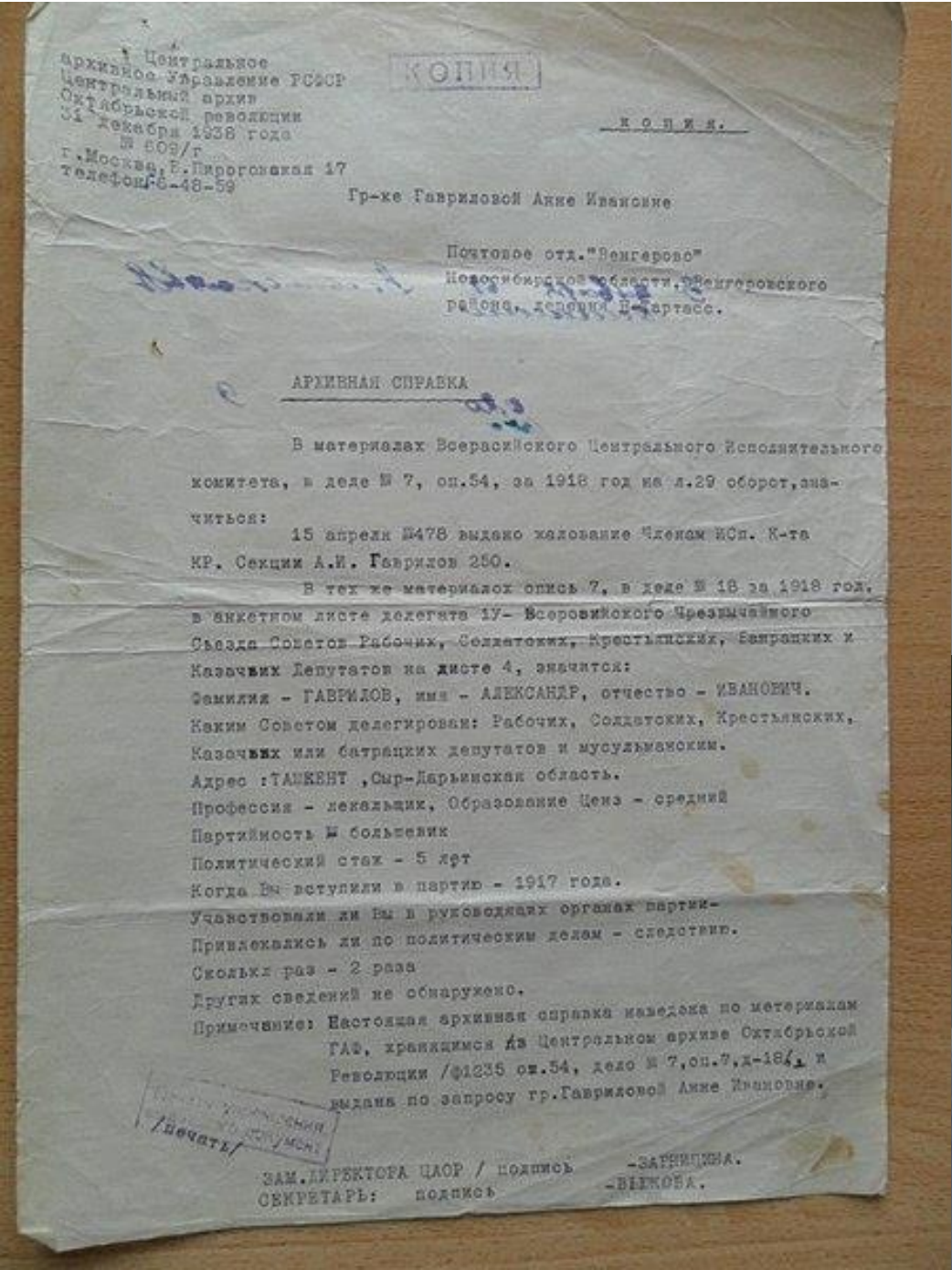


De Smedt, K. Using a Monitor Newspaper Corpus to Trace Changing Language as a Result of COVID-19. CLARIN, Impact Stories. <https://www.clarin.eu/impact-stories/using-monitor-newspaper-corpus-trace-changing-language-result-covid-19>

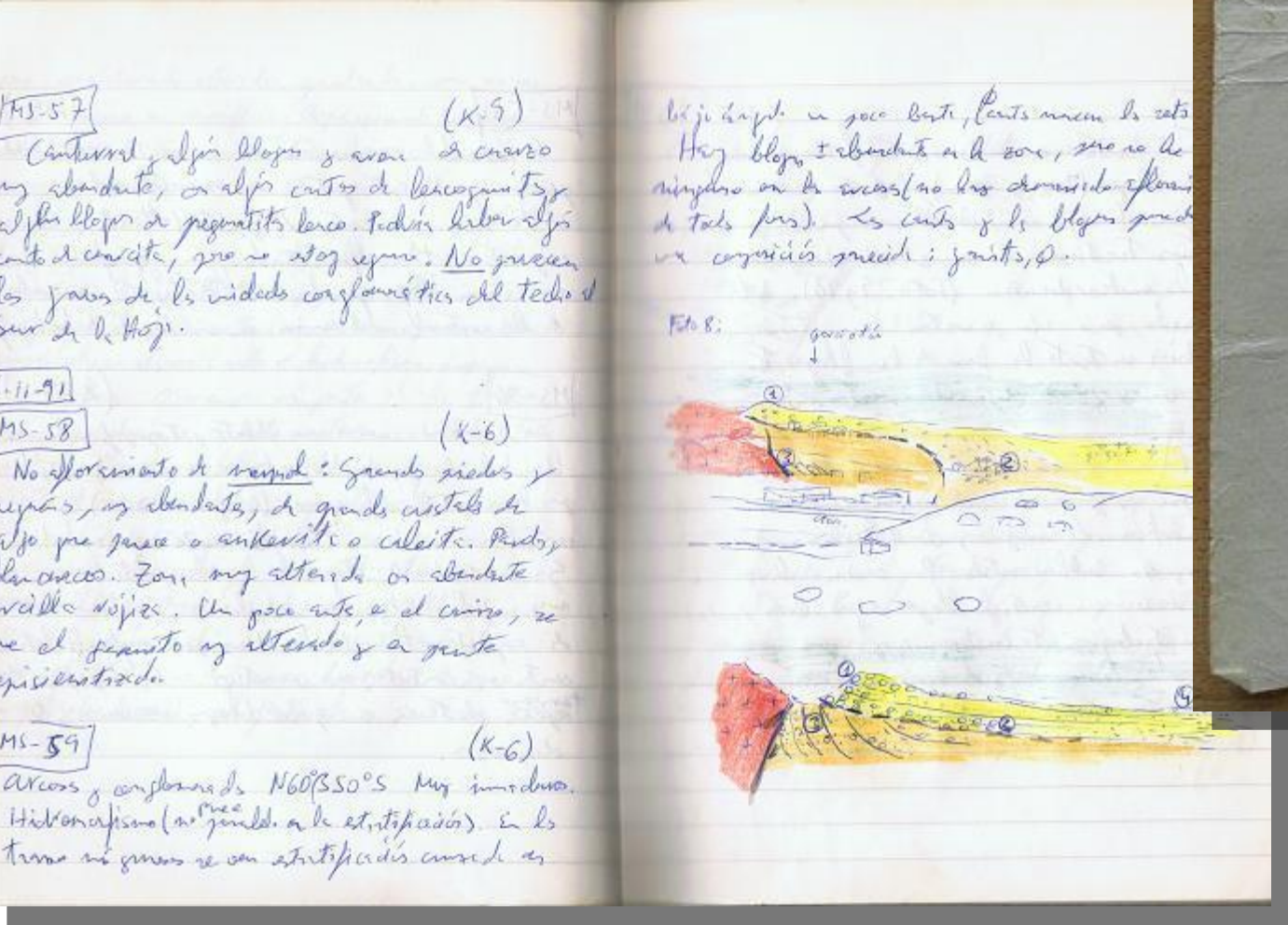
Voices from Ravensbrück: The Value of Multilingual Oral History
Calamai, S., van Hessen, A., S., Scagliola, Draxler, C., van den Heuvel, H. CLARIN, Impact Stories. <https://www.clarin.eu/impact-stories/voices-ravensbruck-value-multilingual-oral-history>

RIPRENDIAMO LE FILA DEL DISCORSO...

Altri ancora con testimonianze personali come **interviste**, **note**, **documenti d'archivio**.



Archival transcript about Alexander Gavrilov. Wikimedia Commons. Public domain, https://commons.wikimedia.org/wiki/File:Archival_transcript_about_Alexander_Gavrilov.jpg

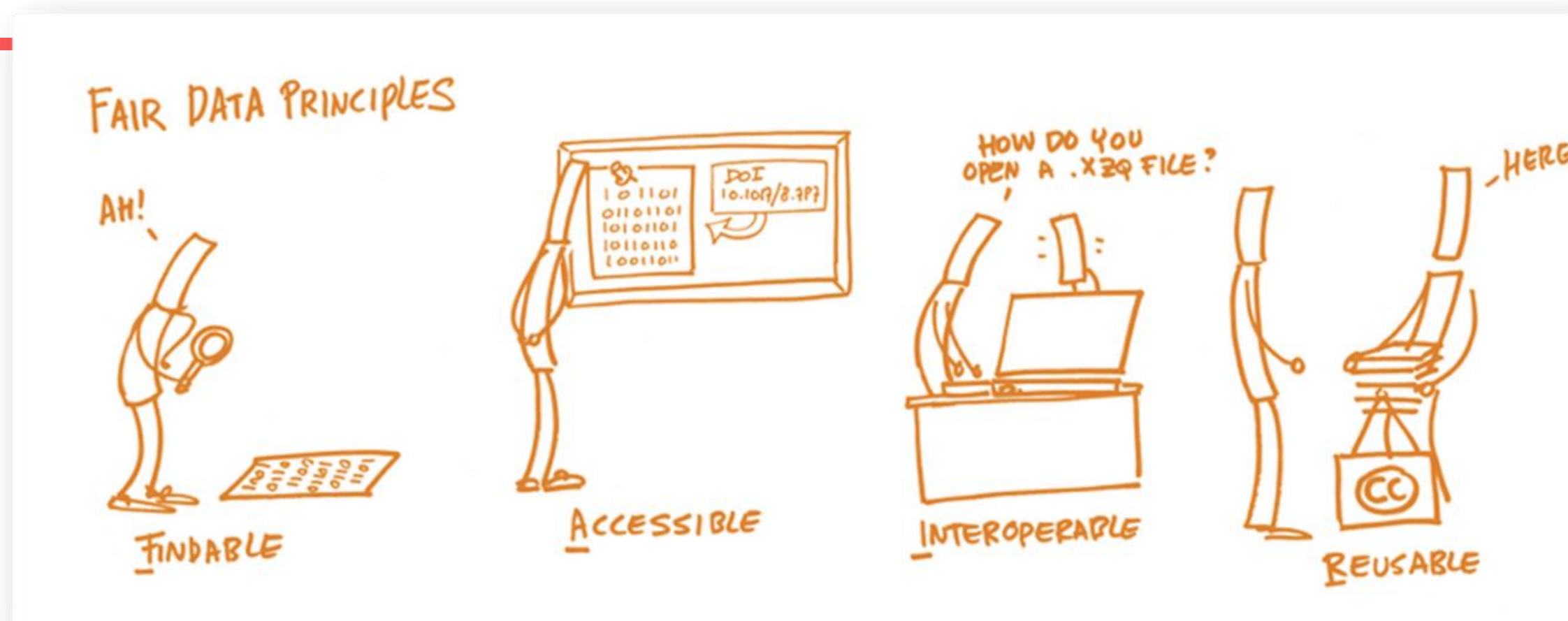


PePeEfe. Geological notebook. Wikimedia Commons. CC BY-SA 3.0, https://commons.wikimedia.org/wiki/File:Geological_notebook.jpg

Elke Wetzig. Live interview with Deutschlandfunk at Wikimania. Wikimedia Commons. CC BY-SA 3.0, https://commons.wikimedia.org/wiki/File:Wikimania_interview_dlf3.jpg



RIPRENDIAMO LE FILA DEL DISCORSO...



- I dati della ricerca devono essere gestiti in modo trasparente: seguire i **principi FAIR** e renderli reperibili, accessibili, interoperabili e riutilizzabili.
- I dati FAIR non sono necessariamente Open Data, ma dati «**as open as possible, as closed as necessary**».
- Un repository per i dati non fa tutto il lavoro per rendere i tuoi dati il più FAIR possibile. Tuttavia, offre un'ottima struttura per ottenere le basi giuste.

RIPRENDIAMO LE FILA DEL DISCORSO...

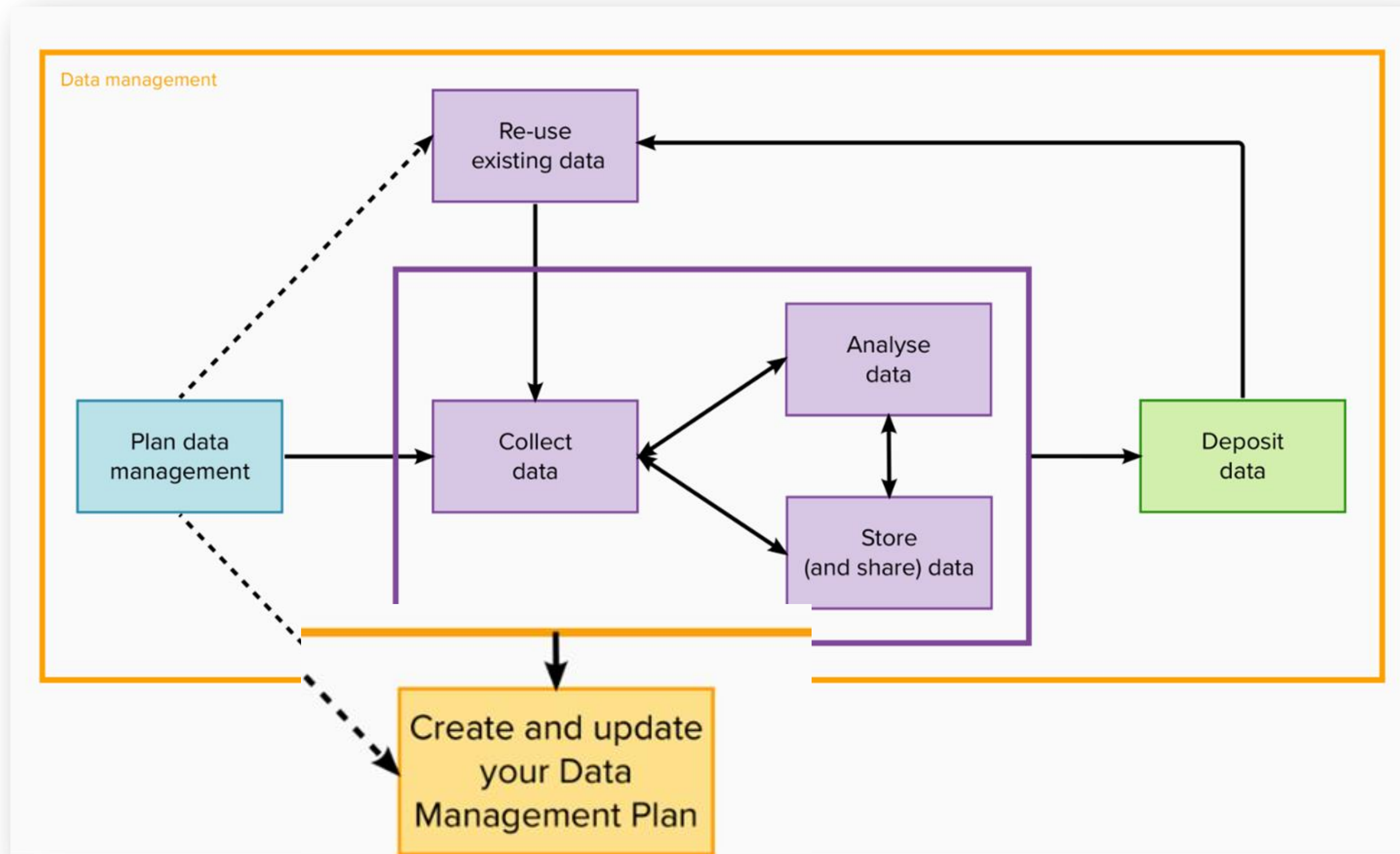


- Gestire attentamente i dati **durante tutte le fasi del ciclo di vita del dato** aiuta a rendere il processo della ricerca più efficiente.
- Le fasi attive della ricerca sono le più concitate: aver **pianificato** può aiutarti a portare avanti la ricerca in modo coerente.
- Non tutti i dati della ricerca sono necessari per la sua comprensione, verifica e per assicurare che la **riproducibilità** dello studio sia garantita.

DATA MANAGEMENT PLAN



CREARE UN DATA MANAGEMENT PLAN: IL PRIMO PASSO PER UNA RICERCA DI QUALITÀ





Data Management Plans (DMPs) are a **key element** of good data management.

A DMP describes the data management life cycle [...]. As part of making research data findable, accessible, interoperable and re-usable (FAIR), a **DMP should include information** on:

- the handling of research data during & after the end of the project
- what data will be collected, processed and/or generated
- which methodology & standards will be applied
- whether data will be shared/made open access and
- how data will be curated & preserved (including after the end of the project).

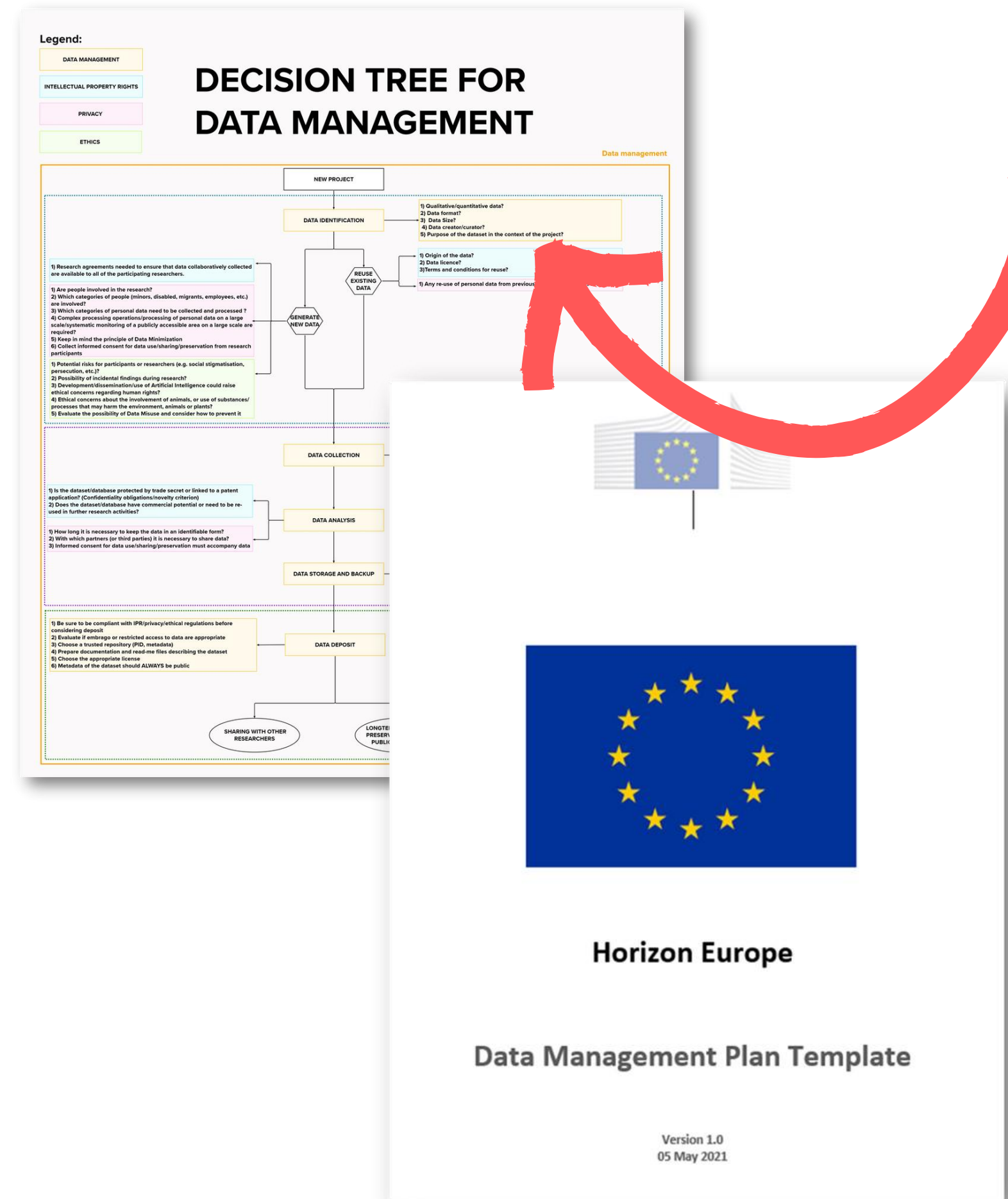


DATA MANAGEMENT PLAN: DOCUMENTA LE RISPOSTE!

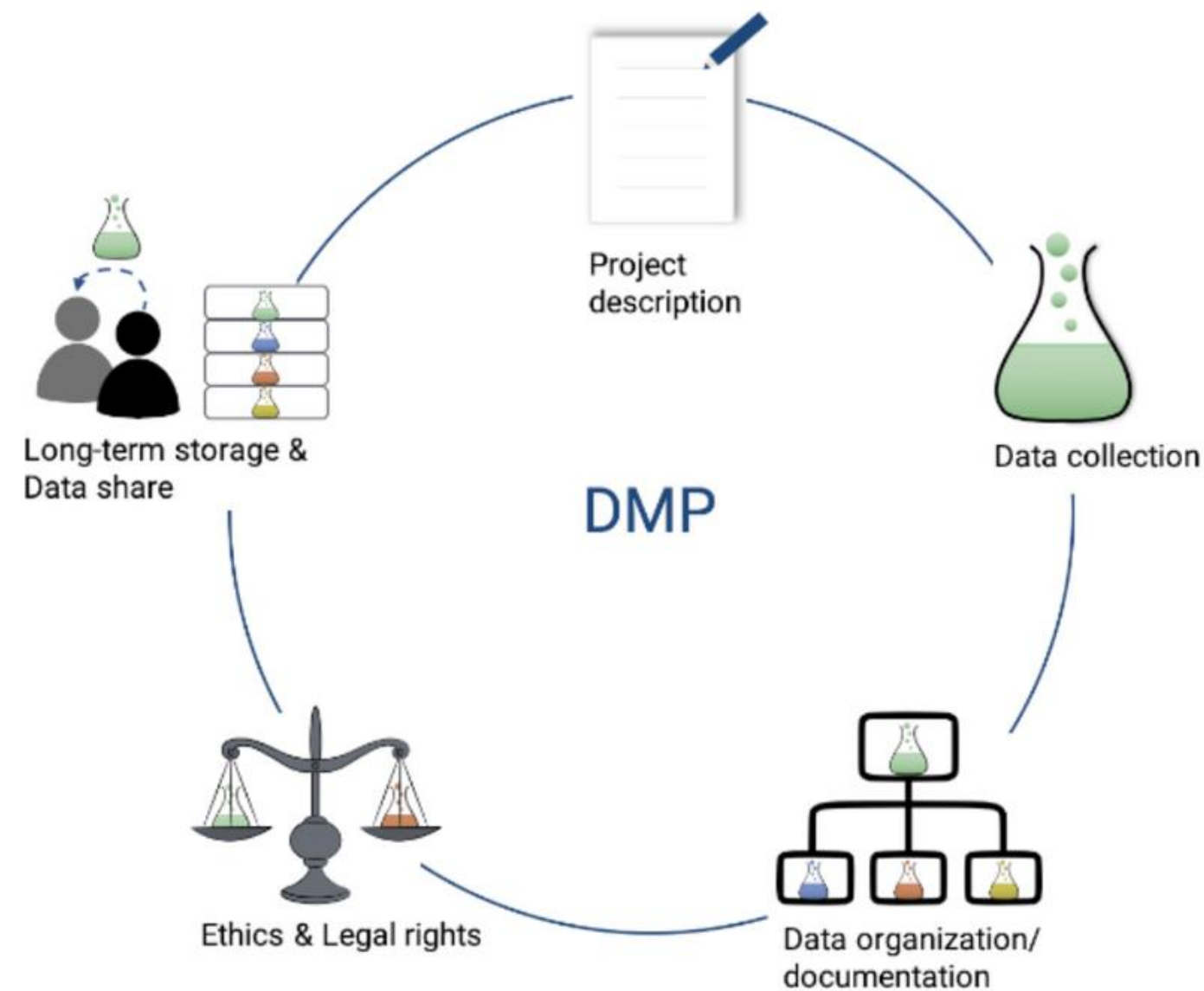
Il DMP viene redatto all'inizio del progetto e aggiornato nel corso del tempo.

Un DMP è una tabella di marcia, con **una serie di domande a cui rispondere** per:

- Gestire i dati **evitando problemi** come duplicazione, perdita di dati, violazioni della sicurezza.
- Sviluppare **procedure efficienti e coerenti** per la gestione dei dati, risparmiando tempo in seguito.
- Essere consapevole degli **adempimenti a tutti gli obblighi** legali ed etici.



UN DATA MANAGEMENT PLAN DEVE INCLUDERE INFORMAZIONI SU:



- **Quali dati** verranno raccolti, processati e/o generati.
- La **gestione dei dati** della ricerca il progetto.
- Quali **metodologie e standard** verranno applicati.
- Come i dati saranno **curati e conservati** dopo la fine del progetto.
- Se i dati saranno **condivisi** ad accesso aperto.



DATA MANAGEMENT PLAN: UN "LIVING DOCUMENT"






Il Data Management Plan ti accompagna durante tutto il progetto.

Inizialmente può documentare delle intenzioni: le tue pratiche di gestione dei dati probabilmente si evolveranno e miglioreranno col procedere della ricerca.

Devi continuare ad **aggiornare il DMP** così da avere, alla fine del progetto, una fotografia delle tue scelte.

- [PERCEIVE project - Deliverable D8.5 "PERCEIVE \(Perception and Evaluation of Regional and Cohesion policies by Europeans and Identification with the Values of Europe\): Data Management Plan. Version \(M6\)". \(deposited 05 Mag 2017 08:23\)](#)
 - [PERCEIVE project - Deliverable D8.5 "PERCEIVE \(Perception and Evaluation of Regional and Cohesion policies by Europeans and Identification with the Values of Europe\): Data Management Plan. Version \(M13\)". \(deposited 15 Nov 2017 09:30\)](#)
 - [PERCEIVE project - Deliverable D8.5 "PERCEIVE \(Perception and Evaluation of Regional and Cohesion policies by Europeans and Identification with the Values of Europe\): Data Management Plan. Version \(M37\)". \(deposited 01 Ott 2019 07:29\)](#) **[Attualmente visualizzato]**

QUALI VANTAGGI PUOI AVERE SCRIVENDO UN DATA MANAGEMENT PLAN?

-  Sviluppare le procedure in anticipo per un lavoro di ricerca efficiente e coerente.
-  Anticipare ed evitare problemi come la duplicazione, la perdita di dati e le violazioni della sicurezza.
-  Assicurarti che i dati siano accurati, completi, affidabili e sicuri.
-  Avere la certezza di rispettare tutti i requisiti legali ed etici.
-  Rendere visibili e citabili i materiali di ricerca per massimizzare l'impatto del tuo lavoro.

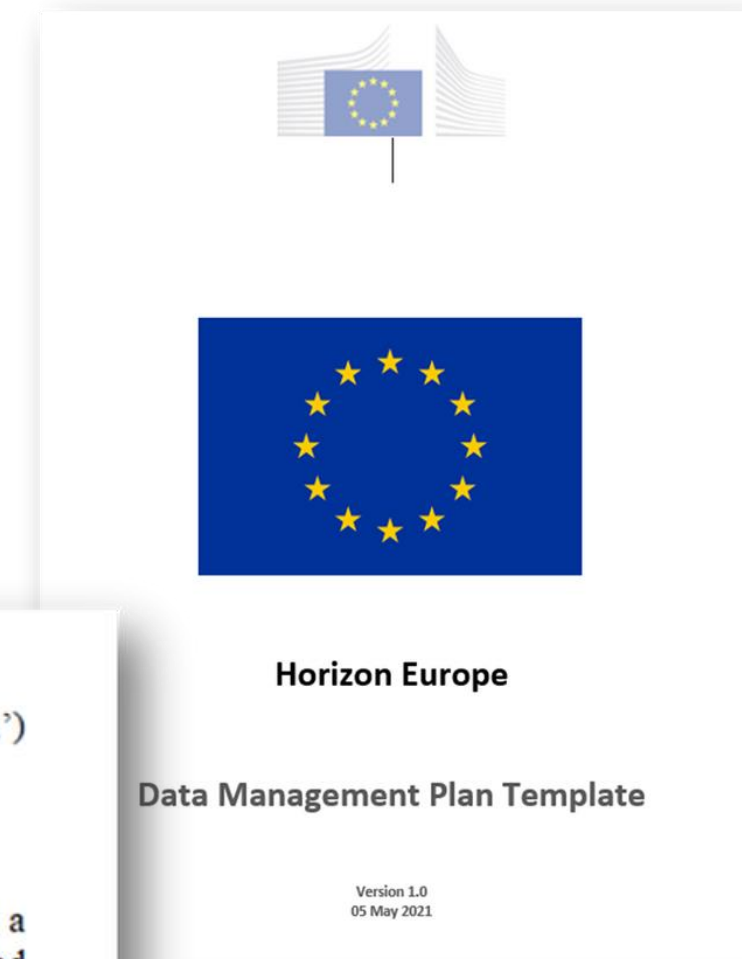
CHI POTREBBE RICHIEDERE UN DATA MANAGEMENT PLAN?

Alcuni enti finanziatori, come la Commissione Europea, richiedono la stesura di un Data Management Plan tra gli obblighi di progetto in Horizon Europe.

Open science: research data management

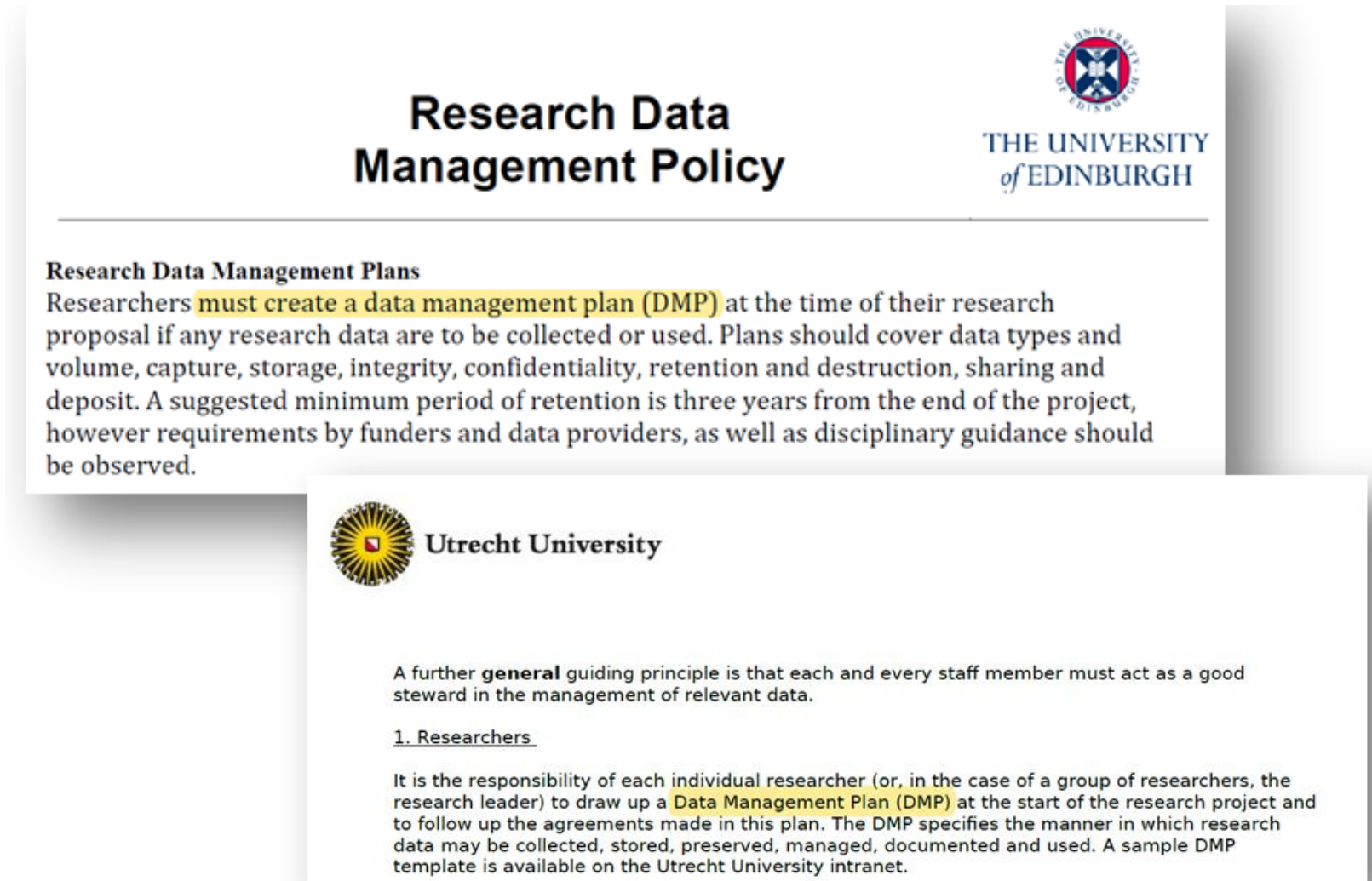
The beneficiaries must manage the digital research data generated in the action ('data') responsibly, in line with the FAIR principles and by taking all of the following actions:

- 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

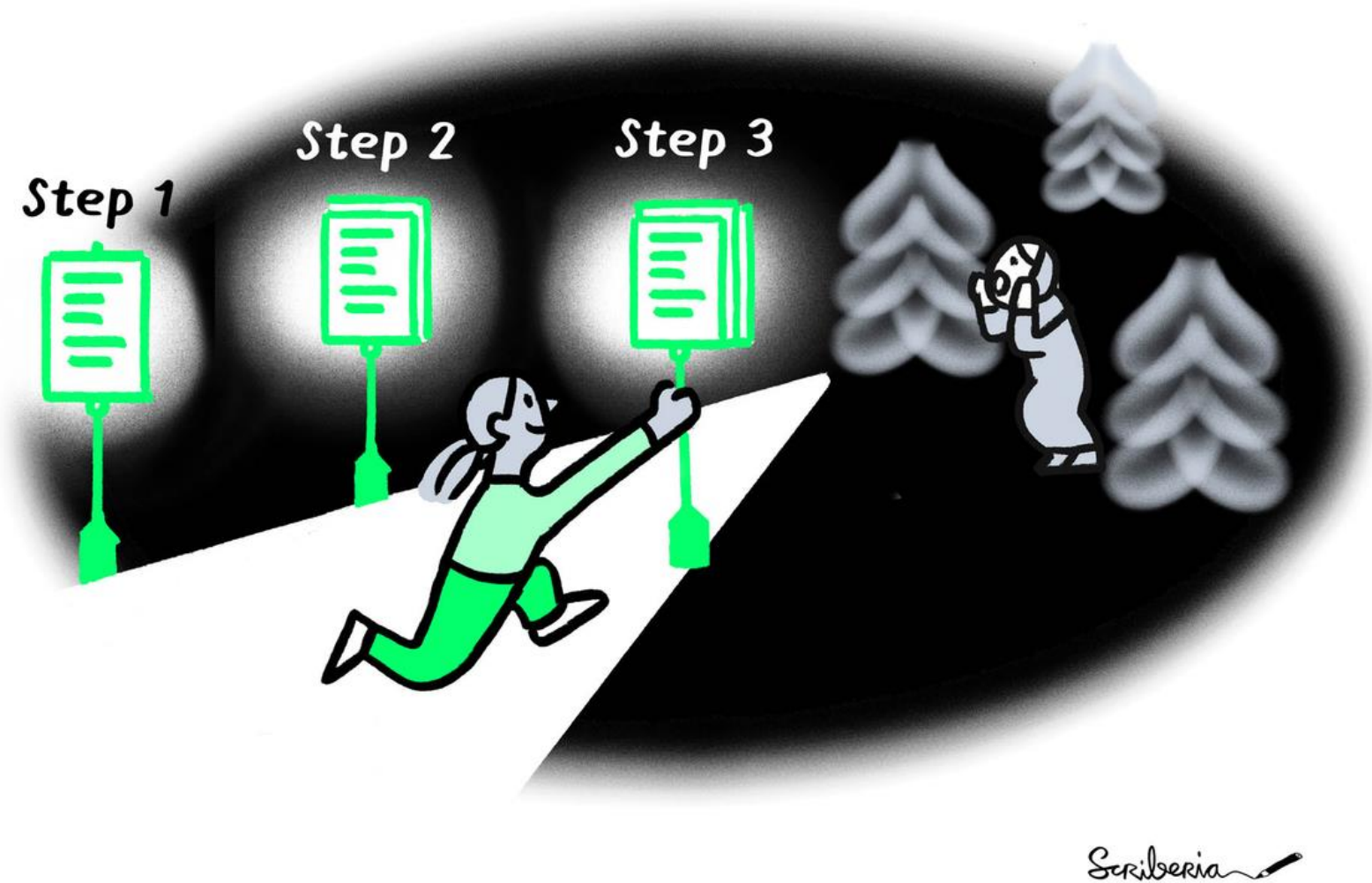


CHI POTREBBE RICHIEDERE UN DATA MANAGEMENT PLAN?

Alcuni atenei europei che si sono dotati di una policy sulla gestione dei dati della ricerca richiedono ai loro ricercatori la stesura obbligatoria di un Data Management Plan nelle fasi iniziali della ricerca.



DATA MANAGEMENT PLAN: SUGGERIMENTI SULLA SCRITTURA




- Non dimenticare che è un **documento operativo**: deve documentare scelte, linee guida, indicazioni.
- Sii il **più specifico possibile** e rendi chiaramente evidenti le informazioni rilevanti. Tabelle, liste, schemi possono essere molto utili.
- Non è necessario spiegare cosa significano "Open Access", "Open Science", "principi FAIR": piuttosto **descrivi le azioni concrete** che intraprenderai!

DATA MANAGEMENT PLAN: SUGGERIMENTI SULLA SCRITTURA



- Non ci sono risposte corrette per tutte le situazioni: **ogni progetto ha le sue specificità.**
- Rifletti sull'approccio più ragionevole per gestire correttamente la tua ricerca.
- **Giustifica scelte e decisioni.**
- Non c'è una ricetta unica alla compilazione, quindi non serve copiare il lavoro degli altri! Tuttavia, dare un'occhiata a DMP esistenti può essere molto d'aiuto.

DOVE TROVARE DEI DMP PUBBLICI?



AMSActa


Institutional Research Repository

by AlmaDL University of Bologna Digital Library

PERCEIVE project - Deliverable D8.5 "PERCEIVE (Perception and Evaluation of Regional and Cohesion policies by Europeans and Identification with the Values of Europe): Data Management Plan. Version (M6)"


A cura di: **Mollona, Edoardo** ; **Pareschi, Luca** ; **Reverberi, Pierre** ; **Brasili, Cristina** (2017) *PERCEIVE project - Deliverable D8.5 "PERCEIVE (Perception and Evaluation of Regional and Cohesion policies by Europeans and Identification with the Values of Europe): Data Management Plan. Version (M6)"*, p. 28. DOI [10.6092/unibo/amsacta/5551](https://doi.org/10.6092/unibo/amsacta/5551).

Salva citazione | Condividi | Citato da



Esiste un'altra versione di questo documento. [Clicca qui per vederla.](#)

Full text disponibile come:



Documento di testo(pdf)
Licenza: [Creative Commons: Attribuzione 4.0\(CC BY\)](#)
[Download \(909kB\)](#) | [Anteprima](#)

URL ufficiale: <http://www.perceiveproject.eu/>

Abstract

This DMP provides details regarding all the research data collected and generated within PERCEIVE project. In particular, licensed and made openly available to the public, and how they will be preserved after the project is completed. This DM project research data cannot be openly shared on account of third-party copyright issues, confidentiality or personal data could jeopardize the project achievements. This DMP reflects the current state of the art of PERCEIVE project (M6). The data vary during the course of research. The variations will be recorded in updated versions of this DMP.

zenodo

data management plan

Up

All versions

Access Right

☐ Open (459702)

☐ Closed (16280)

☐ Restricted (6870)

☐ Embargoed (501)

PDF (19/46)

Found 483353 results.

< 1 2 3 4 5 6 7 8 9 >

Sort by: Best match asc

View

December 18, 2020 (First version) Other Open Access

OH-Harmony-Cap Data Management Plan

Nadia Boisen; Flemming Scheutz;

This is a public deliverable of One Health EJP Joint Research Project, Integrative Action-2.2, OH-HARMONY-CAP: One Health Harmonisation of Protocols for the Detection of Foodborne Pathogens and AMR Determinants <https://onehealth.ejp.eu/jip-oh-harmony-cap/> OH-Harmony-CAP is a 2.5 year project

Uploaded on December 18, 2020

October 24, 2019 (1.0) Data management plan Open Access

D9.1 Data Management Plan

Buresi Pascal; Schoots Fieke; Sesink Laurents;

Executive summary This DMP is a general plan that applies to all research sub- projects in the MIDA project. It is an elaboration of information stated in MIDA's Scientific Proposal, the paragraph on Open Access to Research Data in GRANT Agreement 813547 and the Consortium Agreement for

Uploaded on February 7, 2022

DMPONLINE





Home Public DMPs Funder requirements Help

Public DMPs

Public DMPs are plans created using the DMPonline service and shared publicly by their owners. They are not vetted for quality, completeness, or adherence to funder guidelines.

Q

Search

Project Title	Template	Organisation	Owner	Download
Model - DMP TPM Interviews	TU Delft Data Management Plan template (2021)	Delft University of Technology	Nicolas Dintzner	
DiMo-NEXT	DCC Template	Other	Wolfgang Kremser	
Three dimensional radiological analysis of Hip dysplasia on Computed tomography (CT) scans	UMC Utrecht DMP	UMC Utrecht	Milou Hüskén	
Análisis del impacto de los diseños ópticos de lentes de contacto con desenfoque periférico retiniano en las habilidades visuales de sujetos pediátricos	DCC Template	Other	Hugo Pena Verdeal	

Start new DMP

FAQ

EN

LOG IN

Home

Public DMPs

Public Dataset Desc.

Published DMPs

Sort by: Published

Search DMPs

Published: 29 March 2023

DMP

DIAMOND - Machine-Actionable Data Management Plan

Published Version 0 Grant: Delivering the next generation of open Integrated Assessment M0dels for Net-zero, sustainable Development

Contained Datasets: (1)

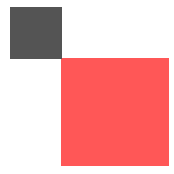
DIAMOND D3.1 - Initial development plan: Surveys

Export

All DMP Versions

<http://amsacta.unibo.it/5551/>; <https://zenodo.org/search?page=1&size=20&q=data%20management%20plan>; https://dmponline.dcc.ac.uk/public_plans; <https://argos.openaire.eu/explore-plans>

19 ottobre 2023

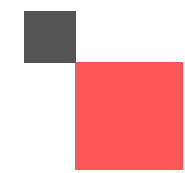


QUALCHE PUNTO CHIAVE



- Il Data Management Plan è un documento strutturato che ti permette di descrivere le scelte fatte durante il progetto di ricerca relativamente alla gestione dei dati.
- Ti permette di tenere traccia delle pratiche da seguire durante le fasi attive della ricerca e di giustificare le tue scelte.
- È un documento che può variare notevolmente da un progetto all'altro perché ogni progetto ha le sue specificità. Non esiste una ricetta unica e giusta per compilarlo!

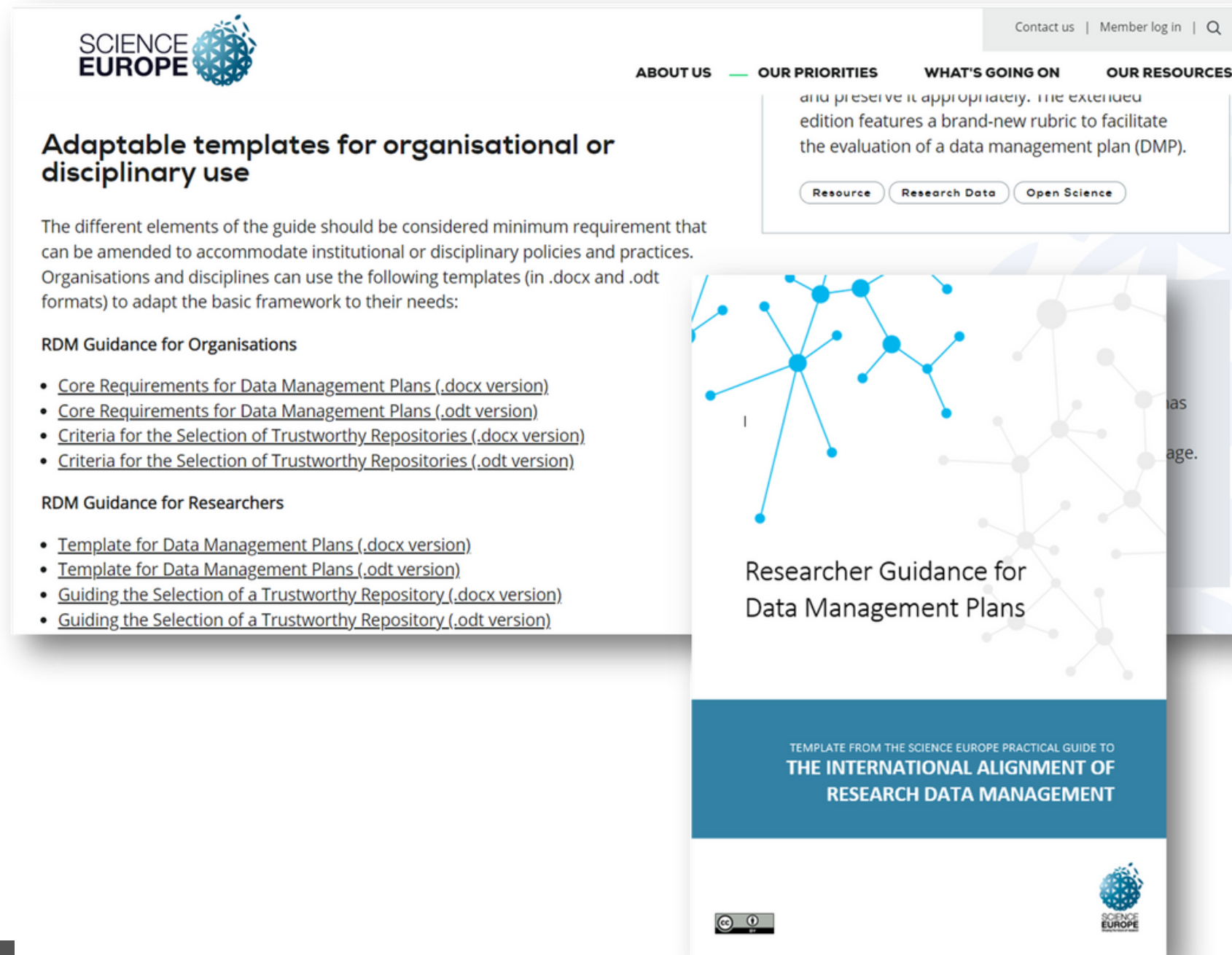




SCRIVERE UN DATA MANAGEMENT PLAN



DATA MANAGEMENT PLAN: I TEMPLATE

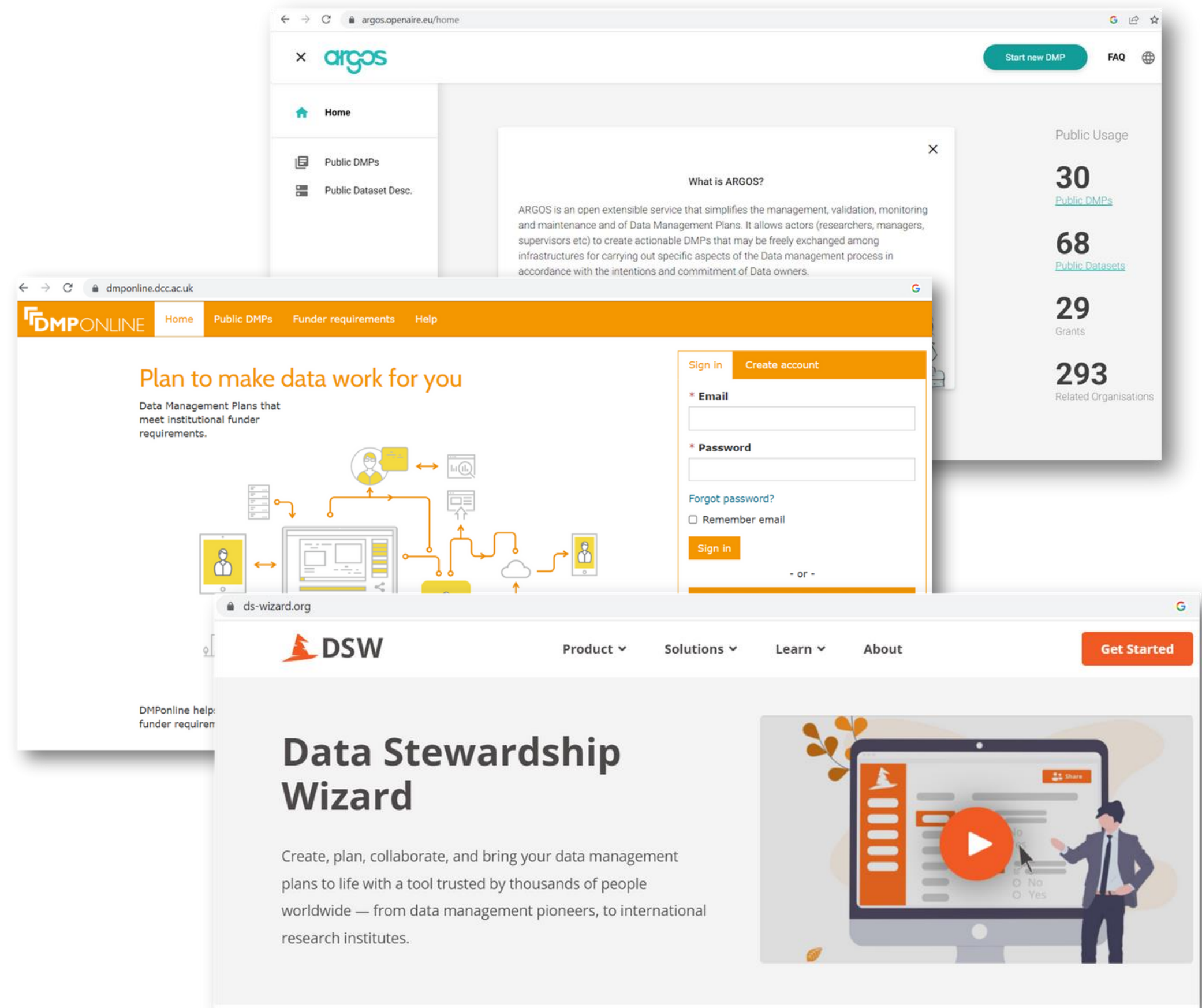


- Un **template** di Data Management Plan ti guida nella compilazione attraverso domande specifiche a cui rispondere.
- Il template di Science Europe raccoglie le informazioni minime da inserire.
- Horizon Europe ha basato il suo template su quello di Science Europe.



DATA MANAGEMENT PLAN: GLI STRUMENTI ONLINE

- Non devi per forza partire da zero per compilare un DMP.
- Esistono dei tool online che ti aiutano in questo scopo.



QUALCHE CONSIGLIO PRATICO PRIMA DI INIZIARE



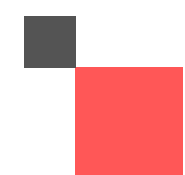
Identifica il prima possibile i tuoi dati di ricerca e le eventuali criticità.



Descrivine dettagliatamente aspetti tecnici e provenienza. Per farlo, segui i suggerimenti dati dal template che stai usando.



Alcuni termini hanno significati specifici e spesso non intuitivi nel contesto del Research Data Management ma ci sono moltissime risorse (glossari, toolkit, guide online) che possono aiutare.



LE SEZIONI DEL DATA MANAGEMENT PLAN





LE SEZIONI DI UN DATA MANAGEMENT PLAN

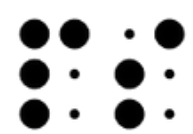
seguendo il template di
Science Europe

- 1. Descrizione dei dati e della loro raccolta/del riutilizzo di dati già esistenti**
2. Documentazione e qualità dei dati
3. Storage e backup durante il processo di ricerca
4. Richieste etiche e legali, codici di condotta
5. Condivisione e preservazione a lungo termine dei dati
6. Responsabilità e costi della gestione dei dati

1. DESCRIZIONE DEI DATI

Cosa deve essere descritto in un Data Management Plan?

1 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?	<ul style="list-style-type: none">• Explain which methodologies or software will be used if new data are collected or produced.• State any constraints on re-use of existing data if there are any.• Explain how data provenance will be documented.• Briefly state the reasons if the re-use of any existing data sources has been considered but discarded.
1b	What data (for example the kind, formats, and volumes), will be collected or produced?	<ul style="list-style-type: none">• Give details on the kind of data: for <u>example</u> numeric (databases, spreadsheets), textual (documents), image, audio, video, and/or mixed media.• Give details on the data format: the way in which the data is encoded for storage, often reflected by the filename extension (for example pdf, xls, doc, txt, or rdf).• Justify the use of certain formats. For <u>example</u> decisions may be based on staff expertise within the host organisation, a preference for open formats, standards accepted by data repositories, widespread usage within the research community, or on the software or equipment that will be used.• Give preference to open and standard formats as they facilitate sharing and long-term reuse of data (several repositories provide lists of such 'preferred formats').



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



DATASET: raccolta di dati assemblati con uno scopo comune e organizzati per riflettere i risultati di un'attività di ricerca.



RIUTILIZZARE DATI

Durante la pianificazione, **controlla se esistono dati già pubblicati da altri ricercatori** utili per la tua ricerca.

L'elemento più importante da considerare è la **possibilità di riutilizzare** legalmente i dati trovati.

- Sono coperti da diritti di proprietà intellettuale (copyright, diritti sulle banche dati)?
- Se il set di dati è accompagnato da una licenza, questa ti dirà cosa puoi o non puoi fare con esso.
- Se il set di dati non è accompagnato da una licenza, dovrai contattare l'autore originale e chiederlo.

ESEMPI DI REPOSITORY DOVE CERCARE DATI DA RIUTILIZZARE:

DISCIPLINARI

- CESSDA, Data Catalogue: <https://datacatalogue.CESSDA.eu/>
- NASA, EarthData: <https://www.earthdata.nasa.gov/>
- GoTriple: <https://www.gotriple.eu/>

GENERALISTI

- Google Dataset Search: <https://datasetsearch.research.google.com/>
- Harvard University Dataverse: <https://dataverse.harvard.edu/>
- Zenodo: <https://zenodo.org/>

DESCRIZIONE DEI DATI: QUALCHE SUGGERIMENTO

Crea una **panoramica dettagliata di tutto ciò che raccogli e crei** durante la tua ricerca sui seguenti argomenti:

- Tipi di dati, se sono raccolti o riutilizzati e come sono stati creati
- Formati dei file che contengono i dati
- Volume dei dati





TIPI DI DATI

- **Contenuto:** numerico, testuale, audiovisivo...
- **Formato:** fogli di calcolo, database, immagini, file audio...
- **Modalità di raccolta** dei dati: sperimentale, osservativa, derivata/compilata da altre fonti...
- **Natura digitale o non digitale** (ad es. indagini cartacee, appunti...)
- **Natura grezza o elaborata**



FORMATI

Per garantire che i dati rimangano accessibili e riutilizzabili, è possibile esportare tutti i dati da formati proprietari a un formato aperto dopo la fine del progetto.

- **Proprietario/chiuso** (di proprietà di una particolare azienda o di un'altra entità che può stabilire quale software può utilizzare il formato).
- **Aperto** (le specifiche del formato sono apertamente disponibili e chiunque può creare software in grado di utilizzare il formato).

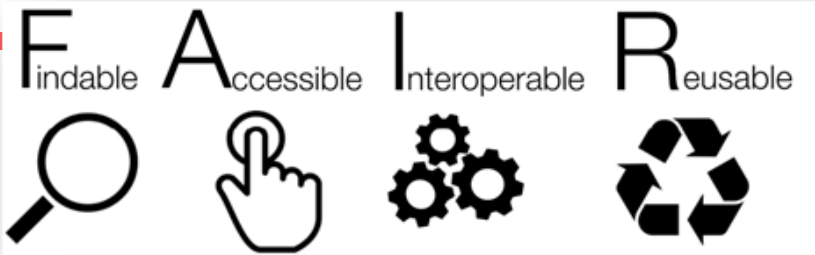


DIMENSIONI

Il volume dei dati diventa rilevante nella scelta delle soluzioni, sia per l'archiviazione attiva che per la conservazione a lungo termine.

Ciò potrebbe determinare la necessità di mettere a budget **costi di archiviazione aggiuntivi**

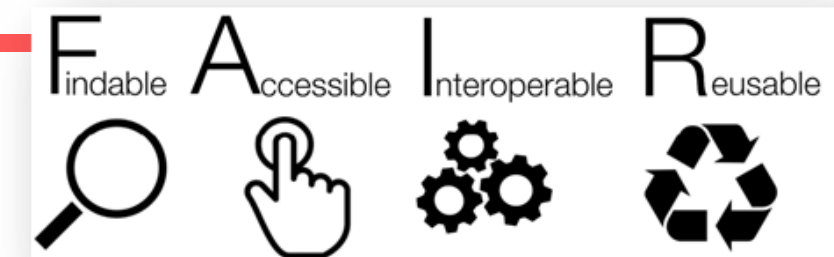
PERCHÈ SCEGLIERE FORMATI APERTI?



L'utilizzo di formati di dati standard o aperti garantisce l'accessibilità e l'usabilità dei dati nel lungo termine.

Type of data	Recommended formats	Other acceptable formats
Qualitative data. Textual.	eXtensible Mark-up Language (XML) text according to an appropriate Document Type Definition (DTD) or schema (.xml). Rich Text Format (.rtf). Plain text data, ASCII (.txt).	Hypertext Mark-up Language (.html). Widely-used formats: MS Word (.doc/.docx). Some software-specific formats: NUD*IST, NVivo and ATLAS.ti.
Digital image data.	TIFF version 6 uncompressed (.tif). Digital Imaging and Communications in Medicine (DICOM) (.dcm, .dcm30) – for CT/MRI data.	JPEG (.jpeg, .jpg) but only if created in this format. TIFF (other versions) (.tif, .tiff). Adobe Portable Document Format (PDF/A, PDF) (.pdf). Standard applicable RAW image format (.raw). Photoshop files (.psd). BMP (.bmp) but only if created in this format. PNG (.png) but only if created in this format.
Digital audio data.	Free Lossless Audio Codec (FLAC) (.flac).	MPEG-1 Audio Layer 3 (.mp3) if original created in this format. Audio Interchange File Format (.aif). Waveform Audio Format (.wav).

PERCHÈ SCEGLIERE FORMATI APERTI?



I dati devono essere combinabili e utilizzabili insieme ad altri dati.
Il formato dei dati deve essere apribile da strumenti e software diversi.

L'elaborazione dei dati deve essere conforme agli standard riconosciuti dalle comunità scientifiche di riferimento.



UN ESEMPIO PRATICO

Data summary

The following table includes an overview of datasets and software solutions developed by SPICE project partners ordered by responsible WP.

Acronyms used: Type = S (Software) / D (Dataset); Status = A (Available) / O (Ongoing) / NA (Not yet available). The DMP identifier is used through the document to identify research outputs.

Type	DMP identifier	Short name	WP	WP Task	Status
D	Workshop	SPICE Participatory Workshop Database	WP2	1-2	O
S	Annotation	Semantic annotation service of social curatorial products	WP3	2	A
S	UM	User Model	WP3	1	O
S	Community	Community visualization tool	WP3	3	O
S	Recomm	Recommender	WP3	4	O
S	LDH	SPICE Linked Data Hub	WP4	1-4	A
S	IMMAPilot	IMMA Pilot Study software	WP7	3	O
D	MuOpinion	Museum visitor opinion dataset	WP7	3	O
S	HaifaPilot	Haifa case study web application	WP7	3	O
D	Demog	End users' demographics, preferences and results from activity analysis	WP7	3	O
D	Ethnog	Autoethnographic and duoethnographic accounts of material culture	WP7	3	O
D	PortableMu	Surveys documenting responses to prototypes	WP7	3	O
S	VRpopup	VR experience of the Pop-up Museum	WP7	3	O
D	FTMdata	User-generated data on the web	WP7	3	O

Table 1 Overview of datasets and software produced in SPICE

Short name	<i>Museum visitors' opinion dataset</i>
DMP identifier	MuOpinion
Identifier	-
Creators	Joel Lanir , University of Haifa
Contributors	-
Rights Holders	University of Haifa, Museum of Hecht
Description and purpose	<p>Description: Hecht Museum visitor's opinions and comments on exhibits at the museum. Tablets will be placed at selected exhibits at the museum presenting information on them and asking for visitor's opinions regarding the exhibits and regarding previous visits. Paid participants will be asked to visit the museum, go to the selected exhibits and enter their comments. Regular museum visitors will also be asked to enter comments.</p> <p>Purpose: Classify museum visitors' opinions according to sentiment according to topics in dedicated tasks in WP3. Evaluate results in WP7.</p>
Data utility	Pilot in WP7
Data types and formats	Qualitative, collected and generated, raw and derivative data. Formats: txt, csv

Table 16 Museum visitors' opinion dataset

NON SOLO ATTENZIONE ALLA
RIPRODUCIBILITÀ DEL DATO

Software di ricerca: comprende
sia il codice sorgente che gli
eseguibili utilizzati nell'ambito
del processo di ricerca.



The FAIR4RS Principles are:

F: Software, and its associated metadata, is easy for both humans and machines to find.
F1. Software is assigned a globally unique and persistent identifier. <ul style="list-style-type: none">F1.1. Components of the software representing levels of granularity are assigned distinct identifiers.F1.2. Different versions of the software are assigned distinct identifiers. F2. Software is described with rich metadata. F3. Metadata clearly and explicitly include the identifier of the software they describe. F4. Metadata are FAIR, searchable and indexable.
A: Software, and its metadata, is retrievable via standardized protocols.
A1. Software is retrievable by its identifier using a standardized communications protocol. <ul style="list-style-type: none">A1.1. The protocol is open, free, and universally implementable.A1.2. The protocol allows for an authentication and authorization procedure, where necessary. A2. Metadata are accessible, even when the software is no longer available.
I: Software interoperates with other software by exchanging data and/or metadata, and/or through interaction via application programming interfaces (APIs), described through standards.
I1. Software reads, writes and exchanges data in a way that meets domain-relevant community standards. I2. Software includes qualified references to other objects.
R: Software is both usable (can be executed) and reusable (can be understood, modified, built upon, or incorporated into other software).
R1. Software is described with a plurality of accurate and relevant attributes. <ul style="list-style-type: none">R1.1. Software is given a clear and accessible license.R1.2. Software is associated with detailed provenance. R2. Software includes qualified references to other software. R3. Software meets domain-relevant community standards.

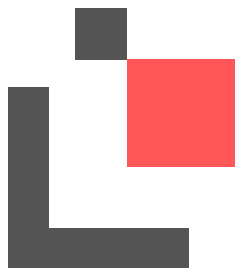


Table 1: The FAIR Principles for Research Software

GESTIRE IL SOFTWARE IN MODO FAIR



Ma GitHub è una «forge», quindi non ha garanzie sulla sostenibilità e sulla conservazione a lungo termine del software!

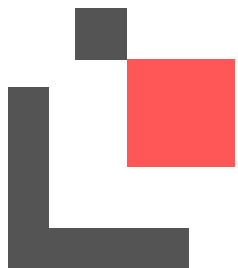


GESTIRE IL SOFTWARE IN MODO FAIR



Affidarsi ad archivi di software per il deposito e la pubblicazione come Software Heritage:

- Pensato per la conservazione a lungo termine del codice sorgente;
- Indice e catalogo del software esistente su diverse piattaforme online;
- Sistema di identificazione criptato (SWHID - PID specifico);
- Archiviazione di elementi contestuali che consentono la comprensione e la ricostruzione del codice;
- Consente la citazione corretta e specifica del software.



DESCRIZIONE DEI DATI: QUALCHE SUGGERIMENTO FINALE

- Se ancora non hai materialmente i dati, è comunque utile **iniziare a pianificare** quali hai intenzione di trattare e come.
- Delinea fin dalle prime fasi la **struttura dei dataset che produrrai**.
- Includere nel DMP finale una **lista di tutti i dataset** che hai prodotto ti sarà molto utile.





LE SEZIONI DI UN DATA MANAGEMENT PLAN

seguendo il template di
Science Europe

1. Descrizione dei dati e della loro raccolta/del riuso di dati già esistenti
- 2. Documentazione e qualità dei dati**
3. Storage e backup durante il processo di ricerca
4. Richieste etiche e legali, codici di condotta
5. Condivisione e preservazione a lungo termine dei dati
6. Responsabilità e costi della gestione dei dati

2. DOCUMENTAZIONE E QUALITÀ DEI DATI



METADATI: informazioni strutturate sui dati. Esistono schemi standard di metadati, sia generici che specifici per disciplina.



DOCUMENTAZIONE: informazioni che rendono i dati più comprensibili agli altri, permette di capirli e interpretarli anche molto dopo la raccolta.

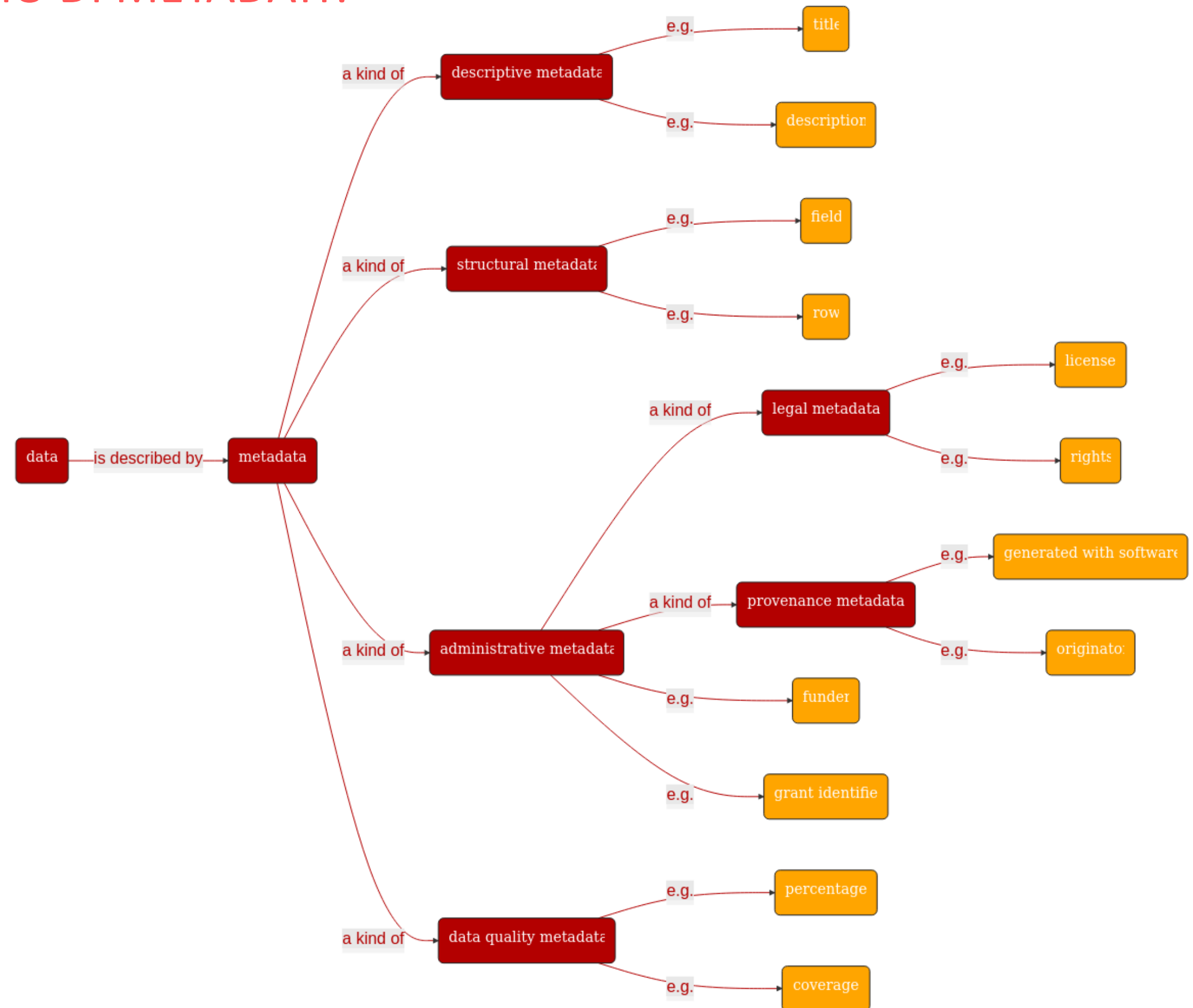
2 Documentation and data quality

- | | |
|------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 2a What metadata and documentation (for example the methodology of data collection and way of organising data) will accompany the data? | <ul style="list-style-type: none">• Indicate which metadata will be provided to 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. |
| 2b What data quality control measures will be used? | <ul style="list-style-type: none">• 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. |

COSA INTENDIAMO QUANDO PARLIAMO DI METADATI?

Informazioni strutturate sui dati:

- Li descrivono, spiegano, localizzano
- Ne facilitano l'uso o la gestione
- Possono essere aggiunti manualmente o automaticamente
- Esistono standard specifici per disciplina





METADATI

- Informazioni che aggiungono struttura ai dati per renderli *machine-readable*.
- Esistono schemi standard di metadati, sia generici che specifici per disciplina.
- È possibile utilizzare vocabolari controllati.

DOCUMENTAZIONE

- Informazioni che rendono i dati più comprensibili agli altri, permette di capirli e interpretarli anche molto dopo la raccolta.
- *Human-readable*.
- Molti ricercatori accompagnano i dati con un README file, che spiega come sono stati raccolti i dati, cosa significano i metadati etc., ma può anche essere "in-file", come i commenti nel codice.

CC

COS'È UN READ ME FILE?

```
README(1).txt - Blocco note di Windows
File Modifica Formato Visualizza ?
README file

Dataset title: What do we mean by “data” in the arts and humanities? Interview transcripts (University of Bologna, FICLIT) and GTM coding

Dataset author/s:
Bianca Gualandi (University of Bologna), https://orcid.org/0000-0001-8202-8493
Luca Pareschi (University of Rome Tor Vergata), https://orcid.org/0000-0002-4402-9329
Silvio Peroni (University of Bologna), https://orcid.org/0000-0003-0530-4305

Dataset contact person: Bianca Gualandi (University of Bologna), biancagualandi15@gmail.com

Publication year: 2022

Dataset contents:

The dataset consists of:
- this README file in .txt format

- a folder named "Interview_trascripts".
This folder contains 19 files in .txt format, each corresponding to 1 interview. All interviews were recorded and transcribed, and the transcript was then anonymised.
The participants to the study all worked at the department of Classical Philology and Italian Studies (FICLIT) at the University of Bologna at the time of the interview. Before the start of each interview, the participant gave their explicit consent to being recorded. After the interview they gave their explicit consent, via email, to the publication of the anonymised interview transcripts as part of this dataset.
The interviews were conducted by Bianca Gualandi between the 9th of November to the 1st of December 2021 both in person and online (according to the interviewee's preference). Whenever the name "Bianca" appears in the transcript, it is a direct reference to the interviewer.

- a folder file named "Analysis_GTM.qda"
This folder contains the qualitative data analysis carried out according to Grounded Theory Methodology as formalised by Cathy Urquhart (2013) "Grounded Theory for Qualitative Research: A Practical Guide". Los Angeles: SAGE Publications. https://doi.org/10.4135/9781526402196
For the coding, QualCoder version 2.9 was used: Curtain, C. (2022) QualCoder 2.9 [Computer software]. Retrieved from https://github.com/ccbogel/QualCoder. QualCoder is written in python3 and distributed under the MIT LICENSE. It uses a SQLite database to store coding data. The authors of this dataset cannot guarantee that the .qda file in this folder is compatible with other qualitative data analysis applications.

Dataset documentation:

Abstract
This dataset contains the anonymised transcripts of the interviews conducted between November and December 2021 at the department of Classical Philology and Italian Studies (FICLIT) at the University of Bologna. It further includes the qualitative data analysis of the interviews, carried out using the Grounded Theory Methodology and the open source software QualCoder version 2.9.

Dataset license
This dataset is released under Creative Commons Zero v1.0 Universal (CC0 1.0)
```

README file

Data Set Title: “[insert title as defined in the DMP]”

Data Set Author/s: Name Surname (Affiliation), ORCID (if available);
[Add one or more creators, if present]

Data Set Contributor/s: Name Surname (Affiliation), ORCID (if available);
[Add one or more contributors, if present. Otherwise, cancel this line]

Data Set Contact Person/s: Name Surname (Affiliation), ORCID (if available), email;
[Add one or more contact person]

Data Set License: this data set is distributed under a [INSERT LICENSE]
[Insert the chosen license as indicated in the DMP: e.g., “this data set is distributed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license, <https://creativecommons.org/licenses/by/4.0/>”]

Publication Year: [insert YEAR]

Project Info: [INSERT PROJECT ACRONYM] ([INSERT PROJECT FULL TITLE]), funded by European Union, Horizon 2020 Programme. Grant Agreement num. **[INSERT G.A. NUMBER]**; **[INSERT LINK TO PROJECT WEBSITE]**.

Data set Contents

The data set consists of:
[Indicate the files that compose the dataset and their name and format.

WE STRONGLY SUGGEST YOU TO FOLLOW THE EXAMPLES PROVIDED FOR THE FILE NAMING, MATCHING THE DATASET FILENAME WITH THE README ONE

In the following **examples** the data sets were composed by only one file. In case the dataset consists of more **files** you can name them as described and put them in a compressed folder. In this **case** **readme** file name should match the compressed folder name]

EXAMPLE1

- 1 textual qualitative file saved in .rtf format
“PROJECTxyz_WP1_T1-2_FocusGroups_20181108_v01.rtf”
[structure of the filename “PROJECT ACRONYM insert WP number insert Task number, e.g. T1.2 (if needed)_ insert Content Describing Keywords insert date YYYYMMDD insert version, if needed format”
Suggested format:
-for textual qualitative data .rtf or .txt
-for tabular quantitative and qualitative data .csv
avoid proprietary formats such as .doc/.docx and .xls/.xlsx]
- 1 README file
“README_PROJECTxyz_WP1_T1-2_FocusGroups_20181108_v01.rtf”
[Same naming as the dataset file. Preferred format .rtf/.txt, allowed format .pdf]

EXAMPLE2

- 1 tabular quantitative file saved in .csv(/.ods) format

Data set Documentation

Abstract

....
[Insert a brief abstract describing the content of the dataset]

Content of the files:

- file **[Insert filename]** contains ...
[Provide a brief description of the content of the file/s. This is an example of how you could start]
- file **[Insert filename]** contains ...
- ...

File specifics

....
[Provide useful info regarding file conversion etc... (Optional)
Please indicate instruction/technical info in order to allow potential users to correctly visualize and reuse your data (e.g. specific software).
In case of data converted in open **formats** it could be useful to provide some further information. For example if you deposit for **long term** preservation a .csv file derived from an excel you can describe the conversion. Here is an example of description of conversion using libre office calc software:

To create the .csv files, “LibreOffice Calc” version: 5.1.4.2 (portable) was used, with the following specifics:
•Character set Europa occidentale (Windows-1252/WinLatin1)
•Field delimiter “,” (comma)
•Text delimiter “” (quotes)]

Notes

... [Related to the whole dataset or to single files of a multi-file dataset (Optional)]

Data sources

... [Optional]

Methodologies

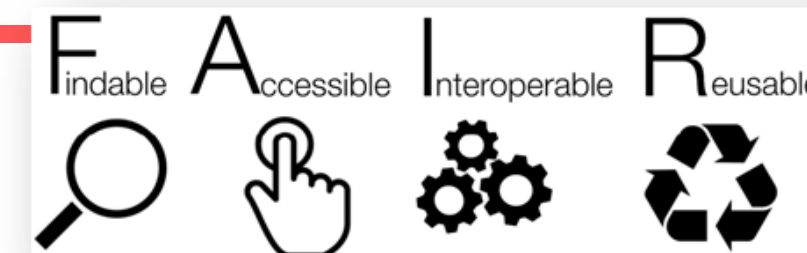
... [If necessary to understand how to reuse data]

Codebook of variables

... [If necessary to understand the meaning of the variables]

Instructions, examples and footnotes in grey should be deleted from final version

PERCHÈ METADATI E DOCUMENTAZIONE SONO FONDAMENTALI?



I dataset devono essere accompagnati da metadati e parole chiave significative per essere rintracciabili online tramite browser e per essere indicizzati da aggregatori di vario tipo.



I metadati utilizzano un linguaggio standardizzato e condiviso a livello internazionale. Inoltre, la documentazione permette di a terzi di comprendere meglio i tuoi dati ed eventualmente combinarli ad altri.



I dati ben descritti e documentati sono di più alta qualità e possono essere riutilizzati più facilmente in contesti diversi.

DOCUMENTAZIONE E QUALITÀ DEI DATI: QUALCHE SUGGERIMENTO

- Nelle fasi iniziali, dire se è presente uno **schema di metadati già noto** per la propria disciplina da utilizzare.
- Verso la fine afferma esplicitamente qual è lo **schema di metadati selezionato** (può dipendere dal repository!).
- Nella descrizione dei dataset dire sempre chiaramente qual è la **documentazione** che li accompagnerà e in che formato sarà disponibile.

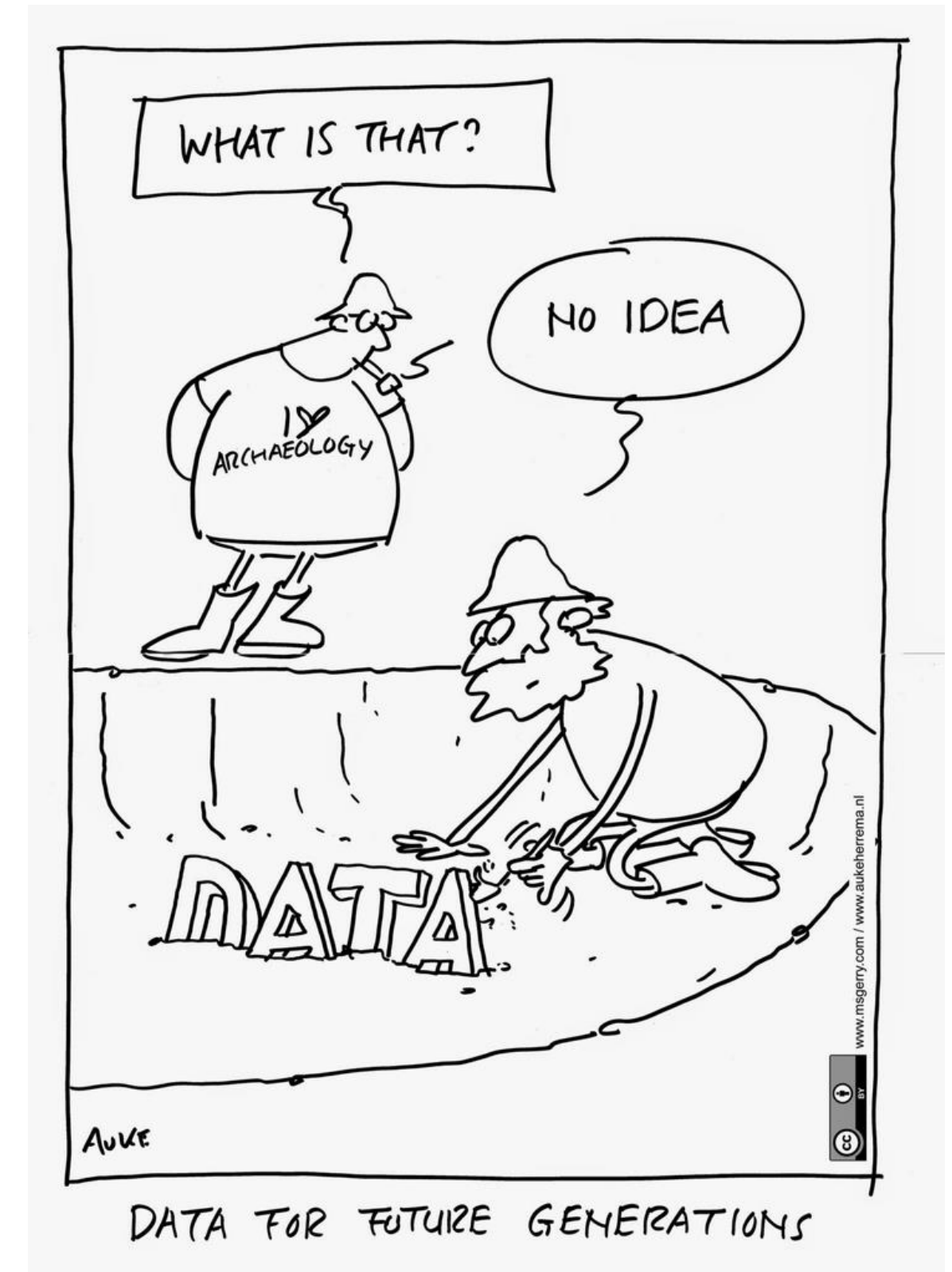


Image credit: <http://aukeherrema.nl> CC-BY

UN ESEMPIO PRATICO

DMP identifier	Provision of metadata and documentation	Reuse or development of metadata standards	Persistent identifiers
Workshop	README file	Zenodo Metadata set, CLAUDIA Metadata set	Zenodo (DOI)
Annotation	Internal documentation	-	-
UM	Javadoc	Zenodo Metadata set	Zenodo (DOI)
Community	README file and user manual	Zenodo Metadata set	Zenodo (DOI)
Recomm	Javadoc	Zenodo Metadata set	Zenodo (DOI)
LDH	README file and interactive documentation	Zenodo Metadata set	Zenodo (DOI)
LDHReg	README file, JSON-LD context	Zenodo Metadata set	Zenodo (DOI)
IGmuseum	README file, JSON-LD context	Zenodo Metadata set	Zenodo (DOI)
IGclassifier	README file, Jupyter notebook	Zenodo Metadata set, AMSActa Metadata set	Zenodo (DOI), AMSActa (DOI)
IUcitizen	README file and interactive documentation	Zenodo Metadata set	Zenodo (DOI)
Reasoner	README file and Javadoc	Zenodo Metadata set	Zenodo (DOI)
Ontologies	Online ontology documentation	Zenodo Metadata set, AMSActa Metadata set	Zenodo (DOI), AMSActa (DOI)



LE SEZIONI DI UN DATA MANAGEMENT PLAN

seguendo il template di
Science Europe

1. Descrizione dei dati e della loro raccolta/del riuso di dati già esistenti
2. Documentazione e qualità dei dati
- 3. Storage e backup durante il processo di ricerca**
4. Richieste etiche e legali, codici di condotta
5. Condivisione e preservazione a lungo termine dei dati
6. Responsabilità e costi della gestione dei dati

3. STORAGE E BACKUP DURANTE LA RICERCA

Storage: immediatamente dopo la raccolta dei dati, può comportare la condivisione con il team/partner.

Comporta:

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

È un aspetto di rilevanza fondamentale durante le fasi attive della ricerca.

3 Storage and backup during the research process

- | | |
|---------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 3a How will data and metadata be stored and backed up during the research? | <ul style="list-style-type: none">• 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. |
| 3b How will data security and protection of sensitive data be taken care of during the research? | <ul style="list-style-type: none">• Explain how the data will be recovered in the event of an incident.• Explain who will have access to the data during the research and how access to data is controlled, especially in collaborative partnerships.• Consider data protection, particularly if your data is sensitive for example containing personal data, politically sensitive information, or trade secrets. Describe the main risks and how these will be managed.• Explain which institutional data protection policies are in place. |

STORAGE DEI DATI: COSA SCEGLIERE?



Hard drive locale (PC, laptop)

- Sempre disponibile

- Rischio di perdita/furto/danneggiamento
- Backup manuali
- Nessun controllo di versione integrato
- Difficile da condividere con altri

Hard drive esterno (hard disk, USB key)

- Portatile
- Economico
- Facile da condividere con altri

- Rischio di perdita/furto/danneggiamento
- Backup manuali
- Nessun controllo di versione integrato
- (Di solito) capacità limitata

Servizi di cloud storage

- Accessibile da più dispositivi
- Facile da condividere
- Collaborativo
- (Di solito) controllo di versione integrato

- Disponibile solo online
- Possibili errori di sincronizzazione
- Backup manuale
- (Solitamente) dipende da un'azienda privata, no garanzia a lungo termine, soprattutto se servizio gratuito

Archiviazione di file su server

- Molto sicuro
- (Di solito) backup automatici
- (Di solito) controllo di versione integrato

- Costoso

BACKUP: PERCHÈ È FONDAMENTALE?



Image credit: <http://aukeherrema.nl> CC-BY

Per gestire i dati in sicurezza ed **evitare** di incorrere in problemi come **la perdita parziale o totale** dei dati!

Buone pratiche:

- Non archiviare i dati solo sul portatile,
- Utilizzare una soluzione di archiviazione con backup automatici,
- Fai backup di tutto il contenuto del laptop/PC,
- Considera soluzioni di storage che permettono di recuperare versioni precedenti di un file.

DATA MANAGEMENT HORROR STORIES

«Ho comprato un portatile nuovo di zecca solo per le mie ricerche (...). Lo usavo per caricare i dati della ricerca (interviste e diari sul campo) e tutto ciò che riguardava la ricerca (...).

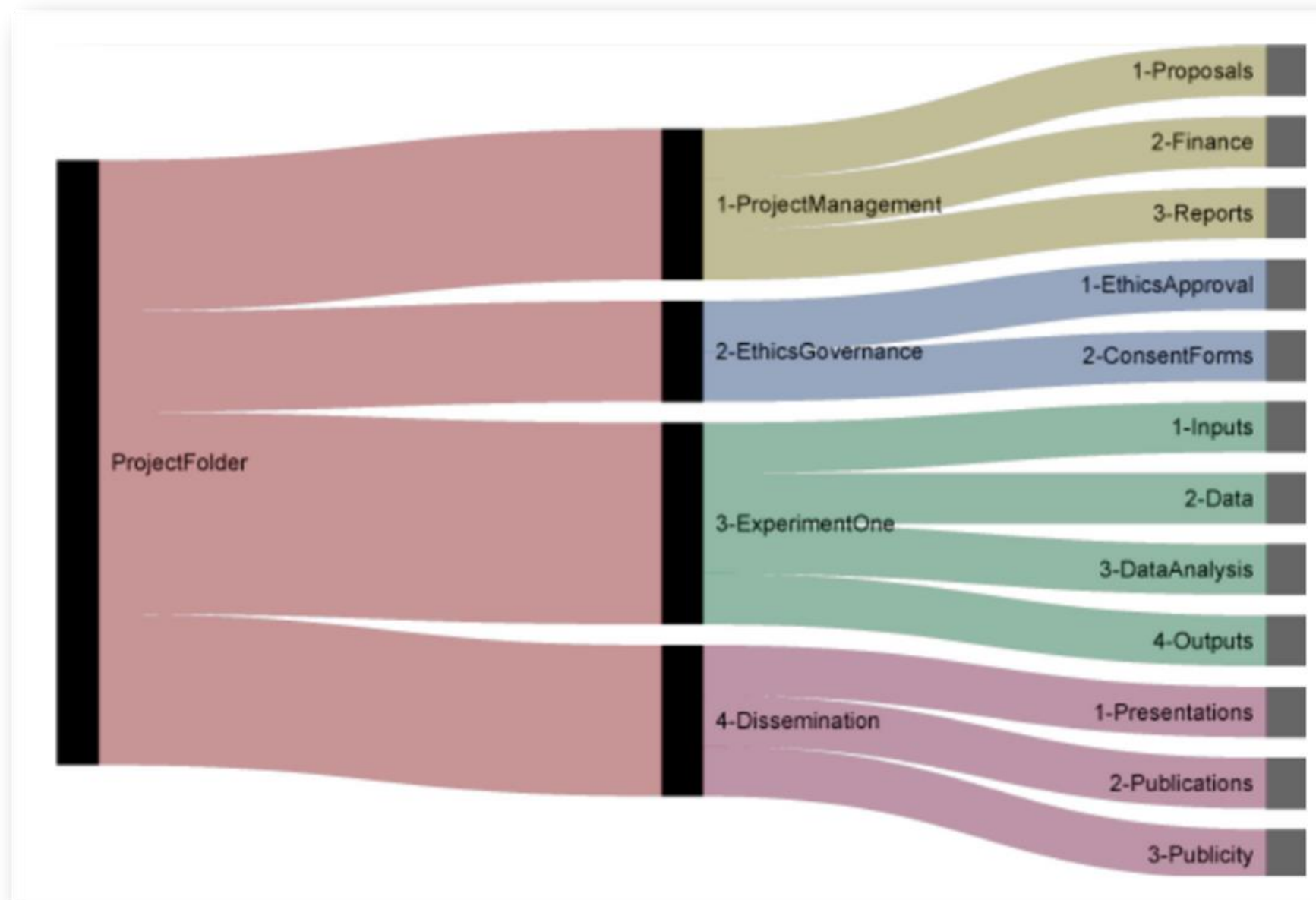
Ho dimenticato il portatile da qualche parte (...). I dati erano sensibili e non erano memorizzati altrove (...).

Il collega ricercatore ha comunque osato chiedermi se avessi fatto un Data Management Plan e una valutazione dei rischi. Mi avrebbe aiutato ad anticipare i problemi».



ORGANIZZAZIONE FILE E CARTELLE

Per organizzare i dati, è bene creare una **struttura di cartelle chiara e organizzata**.



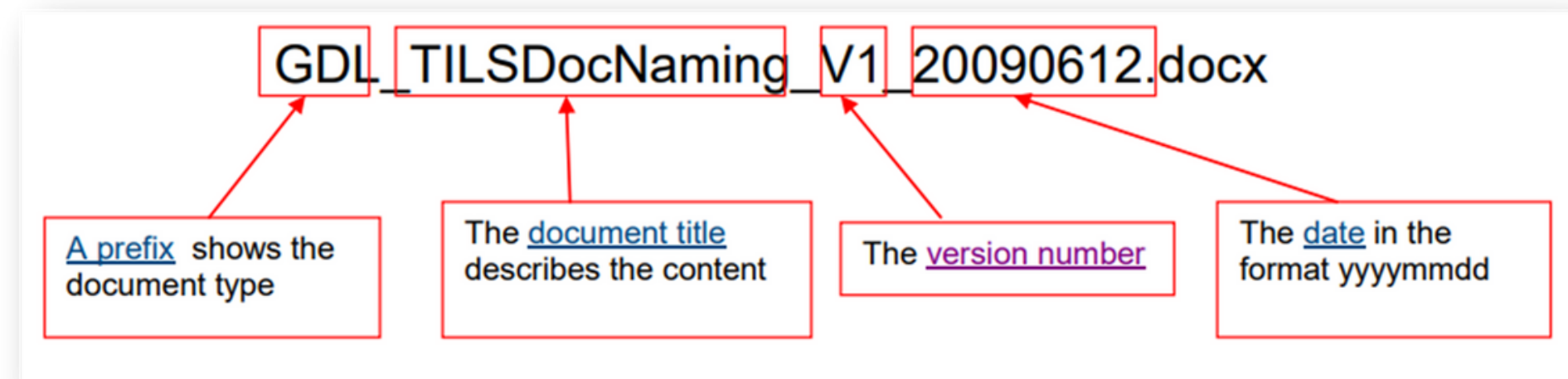
È possibile strutturare le cartelle:

- Per sotto-progetto,
- In base al metodo/attrezzatura di raccolta e analisi,
- In base al tipo di dati,
- In base alla persona che ha generato i dati,
- Cronologicamente.

NOMENCLATURA DI FILE E CARTELLE

Avere **regole di nomenclatura chiare** è essenziale per garantirne l'accessibilità e l'interoperabilità.

- Mantieni un nome breve ma descrittivo
- Per rendere i nomi *machine-readable*, evita spazi e caratteri speciali
- Usa maiuscole o trattini bassi per delimitare le parole



- Utilizzare una formattazione coerente della data, ad esempio ISO 8601: AAAA-MM-GG
- Includere un numero di versione, se applicabile

STORAGE E BACKUP: QUALCHE SUGGERIMENTO

- Decidi subito come organizzare i tuoi dati: puoi farlo per caso studio, per work package, per obiettivo...
- Dedica una sezione a descrivere come terrai sicuri i tuoi dati, dove li terrai salvati e quali procedure di backup utilizzerai
- Una tabella può essere utile nel caso di molti tipi di dati diversi.



UN ESEMPIO PRATICO

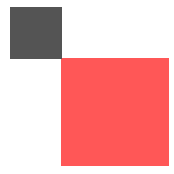
Data security and reuse

All documents relevant to the project are stored in a dedicated **GDPR-compliant remote sharing platform (Microsoft 365)** provided by the University of Bologna, which bought a license that is available to any member of the institution and to external guests that have a Microsoft account. In this private storage area, access is granted only to project partners after requesting permission. In case sensitive data are here stored, these are uploaded only after anonymization procedures, so as to ensure no misuse by other partners is possible. Data recovery and secure storage is guaranteed by the service provider.

Moreover, the **SPICE Linked Data Hub (LDH)** developed in WP4 will allow partners to upload their datasets and have full control over the definition of rules for data reuse, including the definition of copyright restrictions and licenses to any asset they manage. Specifically, the SPICE LDH implements an access control mechanism that allows dataset owners to define their own policies for accessing the dataset. Users can subscribe to the dataset and the owner can grant read/write access to it. This mechanism revolves around the concept of key. A key is a sequence of characters (i.e., a password) that enables a user to perform an operation (i.e., read or write) on a given dataset on the basis of the policy configured by the dataset owner. Keys can be:

- *Read only.* The user can only read the dataset;
- *Write-only.* Dataset owners may wish to grant write access to users or applications that need to contribute data to the dataset without having read access to other data potentially submitted by other users;
- *Read and Write.*

Finally, LDH also ensures data recovery and data backup on a regular basis. The data will be stored in the LDH for the duration of the project. The dataset owner can then decide whether to keep or delete the data. Long-term secure storage of final versions of datasets and software is guaranteed for all aforementioned research output by **Zenodo**.



QUALCHE PUNTO CHIAVE



- Le informazioni che inserisci nel DMP rispecchiano il grado di maturità del progetto: possono essere modificate.
- Identifica tutti i materiali di ricerca che userai/produrrai e descrivili accuratamente nel Data Management Plan.
- Descrivi le strategie di documentazione e metadatazione che decidi di mettere in atto.
- Dedica una sezione a descrivere come terrai sicuri i tuoi dati, dove li terrai salvati e quali procedure di backup utilizzerai.





LE SEZIONI DI UN DATA MANAGEMENT PLAN

seguendo il template di
Science Europe

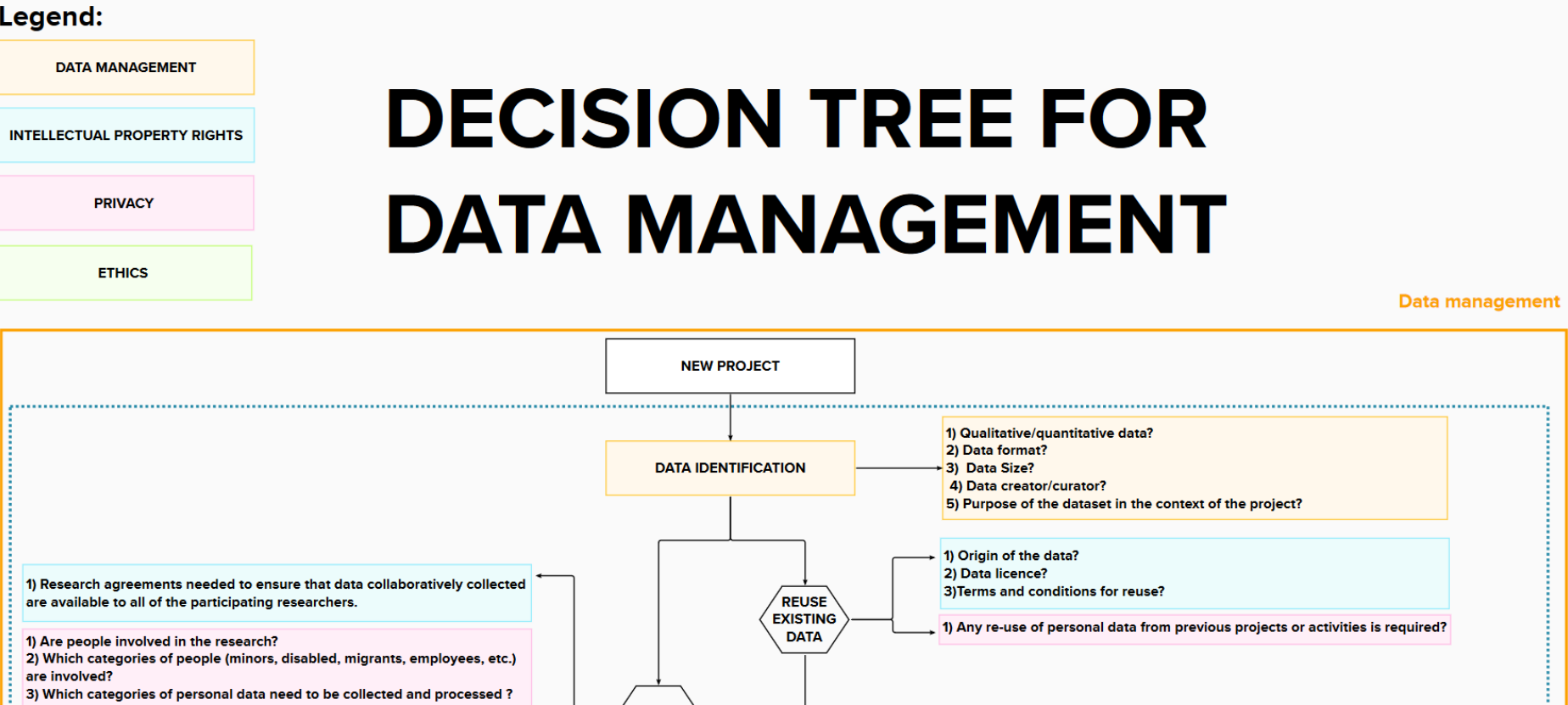
1. Descrizione dei dati e della loro raccolta/del riutilizzo di dati già esistenti
2. Documentazione e qualità dei dati
3. Storage e backup durante il processo di ricerca
- 4. Richieste etiche e legali, codici di condotta**
5. Condivisione e preservazione a lungo termine dei dati
6. Responsabilità e costi della gestione dei dati

4. RICHIESTE ETICHE E LEGALI, CODICI DI CONDOTTA

Quando inizia una ricerca, è opportuno prendere in considerazione gli aspetti che possono condizionarla.

In particolare le **responsabilità etiche e legali nella ricerca** influenzano le modalità di raccolta, gestione, conservazione e condivisione dei dati e possono essere di diversa natura.

4 Legal and ethical requirements, codes of conduct		
4a	If personal data are processed, how will compliance with legislation on personal data and on security be ensured?	<ul style="list-style-type: none">• Ensure that when dealing with personal data protection laws (for example GDPR) are complied with:<ul style="list-style-type: none">➢ Gain informed consent for preservation and/or sharing of personal data.➢ Consider anonymisation of personal data for preservation and/or sharing (truly anonymous data are no longer considered as personal data).➢ Consider pseudonymisation of personal data (the main difference with anonymisation is that pseudonymisation is reversible).➢ Consider encryption which is seen as a special case of pseudonymisation (the encryption key must be stored separately from the data, for instance by a trusted third party).➢ Explain whether there is a managed access procedure in place for authorised users of personal data.
4b	How will other legal issues, such as intellectual property rights and ownership, be managed? What legislation is applicable?	<ul style="list-style-type: none">• Explain who will be the owner of the data, meaning who will have the rights to control access:<ul style="list-style-type: none">➢ Explain what access conditions will apply to the data? Will the data be openly accessible, or will there be access restrictions? In the latter case, which? Consider the use of data access and re-use licenses.➢ Make sure to cover these matters of rights to control access to data for multi-partner projects and multiple data owners, in the consortium agreement.• Indicate whether intellectual property rights (for example Database Directive, sui generis rights) are affected. If so, explain which and how will they be dealt with.• Indicate whether there are any restrictions on the re-use of third-party data.
4c	What ethical issues and codes of conduct are there, and how will they be taken into account?	<ul style="list-style-type: none">• Consider whether ethical issues can affect how data are stored and transferred, who can see or use them, and how long they are kept. Demonstrate awareness of these aspects and respective planning.• Follow the national and international codes of conducts and institutional ethical guidelines, and check if ethical review (for example by an ethics committee) is required for data collection in the research project.



COMPLIANCE ETICA

È necessario riconoscere questioni e criticità di rilievo etico in relazione a **obiettivi, metodologia e impatti** del progetto di ricerca.

Utilizza il DMP per prevedere (e dimostrare) il rispetto dei principi etici e delle normative a fronte delle criticità etiche rilevate. Queste possono riguardare:

- **ricerca con esseri umani**, in particolare:
 - privacy (che vedremo tra poco)
 - potenziali rischi per partecipanti o ricercatori
 - *unexpected findings*
- **intelligenza artificiale** (sviluppo, diffusione, uso)
- coinvolgimento di **animali**
- uso di sostanze/processi che possono danneggiare **l'ambiente**.

- 1) Potential risks for participants or researchers (e.g. social stigmatisation, persecution, etc.)?
- 2) Possibility of incidental findings during research?
- 3) Development/dissemination/use of Artificial Intelligence could raise ethical concerns regarding human rights?
- 4) Ethical concerns about the involvement of animals, or use of substances/processes that may harm the environment, animals or plants?
- 5) Evaluate the possibility of Data Misuse and consider how to prevent it

TUTELA DEI DATI PERSONALI

1) Are people involved in the research?
2) Which categories of people (minors, disabled, migrants, employees, etc.) are involved?
3) Which categories of personal data need to be collected and processed ?
4) Complex processing operations/processing of personal data on a large scale/systematic monitoring of a publicly accessible area on a large scale are required?
5) Keep in mind the principle of Data Minimization
6) Collect informed consent for data use/sharing/preservation from research participants

1) With which partners (or third parties) it is necessary to share data?
2) In which countries personal data will be collected?
3) To which countries personal data may be transferred?

Dato personale:

Qualsiasi dato riferibile direttamente o indirettamente ad una persona fisica

Dato particolare (sensibile):

- origine razziale o etnica
- opinioni politiche, appartenenza sindacale
- convinzioni religiose o filosofiche
- dati genetici, dati biometrici
- dati sulla salute
- dati relativi all'orientamento/vita sessuale

DATA MANAGEMENT HORROR STORIES



«Abbiamo raccolto dati di ricerca ricchi e diversificati.(...). Poi il progetto è terminato e ci è venuto in mente che potevamo depositare i dati per un ulteriore utilizzo (...).

Hanno iniziato a farci domande difficili come: avete chiesto ai partecipanti alla ricerca il permesso di archiviare i dati e avete menzionato l'archiviazione nell'informativa sulla privacy? (...) Tutti argomenti che ci erano sfuggiti perché avevamo fretta di iniziare il lavoro sul campo. Di conseguenza, abbiamo perso l'opportunità di aprire i dati ad altri e persino di utilizzarli nei nostri progetti futuri».

TUTELA DEI DATI PERSONALI: QUALCHE SUGGERIMENTO

- Pianifica e ottieni **un consenso informato** all'inizio della ricerca **per tutti i tipi di usi** che vuoi fare con i dati (compresa la disseminazione)
- Proteggi i documenti che contengono i dati personali; se li carichi in un cloud, considera se cifrarli (esistono tool appositi) oppure **pseudonimizzali** subito
- Appena possibile **anonimizza i dati**, in modo che possano essere trattati come normali dati di ricerca



TUTELA DEI DATI PERSONALI: QUALCHE SUGGERIMENTO

- Raccogli dati personali **solo se sono strettamente necessario** agli scopi della ricerca
- Consenti l'accesso a questi dati solo a **persone autorizzate**
- Descrivi nel DMP tutte le strategie che metti in atto



NORME DI RIFERIMENTO:

- Codice Privacy: Dlgs. 196/2003
- Regolamento (UE) 2016/679 ("General Data Protection Regulation - GDPR")
- Regolamento di Ateneo DR 171/2013
- Pronunce / atti del Garante della Privacy



UN ESEMPIO PRATICO

- As stated in SPICE Deliverable D9.2, where the collection of personal data is desirable, it will be limited to the scope of the partner acting as Data Controller and with the sole purpose of allowing the data owners to act on their rights with relation to GDPR.
- It is a recommended policy to use one or more pseudonymisation techniques to enable the exchange of data between partners and where it is not necessary to know the identity of the data subject.
- Anonymisation and pseudonymisation techniques adopted include:
 - *Data Suppression*. Removing identifier references such as data subjects' name, address, location data, online identifier, postcode information.
 - *Data Generalisation*. Modify the precision of the gathered variables, where possible and according to the type of data, represent data using scales. For example, age range (0-3, 4-6), partial information such as year of birth rather than full date.
 - *Data Encryption*. A technique that uses a secret key to obtain pseudonyms from an identifier, it uses an encryption algorithm.
 - *Data Tokenisation*. A non-mathematical approach that generates randomly values that replace data identifiers with non-sensitive information.
 - *Data swapping*. Rearranging records so that they do not match the initial information.
 - *Data perturbation*. changing the original data by using rounding methods and random noise.

TUTELA DELLA PROPRIETÀ INTELLETTUALE

Insieme di diritti legali volti ad assicurare la tutela delle creazioni della mente umana in campo scientifico, industriale e artistico.

- 1) Origin of the data?
- 2) Data licence?
- 3) Terms and conditions for reuse?

IP used



IP created

- 1) Is the data selection and disposition in the dataset/database original?
- 2) Dataset/database generation is the result of major investments (time/money)?

- 1) Is the dataset/database protected by trade secret or linked to a patent application? (Confidentiality obligations/novelty criterion)
- 2) Does the dataset/database have commercial potential or need to be re-used in further research activities?

RICERCA

IL DIRITTO D'AUTORE

Protezione che copre le **opere letterarie, scientifiche e artistiche**, pubblicate e non, fissate in forma materiale o tangibile (ad esempio su carta, pellicola, registrazione sonora); è un diritto non registrato che **inizia automaticamente quando l'opera viene effettivamente creata**.



- Requisiti: originalità, creatività, novità
- Esiste automaticamente con la creazione dell'opera
- Non richiede una registrazione
- Durata: vita dell'autore + 70 anni

ESEMPIO: FONTI PROTETTE DAL DIRITTO D'AUTORE



Genoni, Rosa. *Per una moda italiana: modelli saggi schizzi di abbigliamento femminile: 1906-1909*. Milano : Alfieri & Lacroix ©
Podreider. <https://amshistorica.unibo.it/archivio/000322/000019.jpg>

Chi lavora con fonti contemporanee ha spesso a che fare con fonti protette dal diritto d'autore, solitamente con «tutti i diritti riservati».

Questo significa che è **necessario chiedere il permesso agli autori** o a chi per loro **per qualsiasi tipo di riutilizzo** dell'opera.

Questa protezione generalmente scade **70 anni dopo la morte degli autori**. Alla scadenza della protezione, l'opera diventa di pubblico dominio.

Chi lavora con fonti antiche e moderne non avrà a che fare con fonti protette dal diritto d'autore, tuttavia...

[UNA PARENTESI: FONTI NON PIÙ PROTETTE DAL DIRITTO D'AUTORE]

...se la fonte si trova in un istituto di conservazione (biblioteca, museo, archivio), è bene mettersi in contatto con quest'ultimo il prima possibile, per **definire le condizioni di riutilizzo** delle riproduzioni del patrimonio culturale.

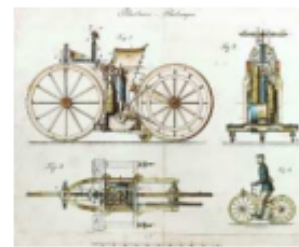
Il **“Codice dei beni culturali”** (2004) regola l'utilizzo del patrimonio culturale e dà **un forte controllo agli istituti di conservazione** anche **per quanto riguarda le rappresentazioni digitali dei beni culturali** che hanno in consegna.

Il Ministero della Cultura ha recentemente (maggio 2023) emanato delle [contestate](#) Linee Guida per determinare canoni e corrispettivi minimi per la **concessione d'uso dei beni culturali** statali.

D'altro canto, la nuova **direttiva europea sul diritto d'autore e sui diritti connessi nel mercato unico digitale** (recepita in Italia nel 2021) stabilisce che, alla scadenza della durata di protezione di un'opera, **il materiale derivante da un atto di riproduzione di tale opera non è soggetto al diritto d'autore** o a diritti connessi, a meno che non costituisca una creazione intellettuale propria.

I BREVETTI

Assicurano che un'invenzione diventi proprietà di una persona o di un'organizzazione, che l'invenzione possa essere acquistata, venduta, affittata o noleggiata e che **il titolare abbia il diritto, per un periodo limitato, di impedire ad altri di realizzare, utilizzare o vendere l'invenzione senza il permesso del proprietario.**



Invenzione

Soluzione nuova
ed originale ad
un problema
tecnico



- Requisiti:
 - Novità
 - Originalità
 - Applicazione industriale
 - Sufficiente descrizione
 - Liceità
- Durata: 20 anni

TUTELA DELLA PROPRIETÀ INTELLETTUALE: QUALCHE SUGGERIMENTO

- Valuta fin dall'inizio se i dati costituiscono un'invenzione e possono servire per un brevetto; in quel caso, **tienili segreti fino alla brevettazione**
- Nel caso non vi sia interesse a brevettare, **considera la possibilità di rilasciare i tuoi dati in maniera aperta**, per promuoverne il massimo riuso (anche commerciale)
- Alternativamente, puoi valutare di vietarne il riutilizzo commerciale **attraverso l'uso di licenze specifiche** (e.g. CC BY-NC)





LE SEZIONI DI UN DATA MANAGEMENT PLAN

seguendo il template di
Science Europe

1. Descrizione dei dati e della loro raccolta/del riutilizzo di dati già esistenti
2. Documentazione e qualità dei dati
3. Storage e backup durante il processo di ricerca
4. Richieste etiche e legali, codici di condotta
- 5. Condivisione e preservazione a lungo termine dei dati**
6. Responsabilità e costi della gestione dei dati

5. CONDIVISIONE E PRESERVAZIONE A LUNGO TERMINE DEI DATI

Deposito: fatto una volta o (preferibilmente) in modo iterativo.

Comporta:

- Passaggio della responsabilità per la preservazione a lungo termine dal ricercatore al **repository**
- Scelta della licenza e del livello di accesso

NB. Con «deposito dei dati» non si intende il loro caricamento su siti web personali, siti web degli editori, servizi di cloud storage (Dropbox, Google drive, ecc.), piattaforme di ricerca (Academia.edu, ResearchGate)!

5 Data sharing and long-term preservation		
5a	How and when will data be shared? Are there possible restrictions to data sharing or embargo reasons?	<ul style="list-style-type: none">• Explain how the data will be discoverable and shared (for example by deposit in a trustworthy data repository, indexed in a catalogue, use of a secure data service, direct handling of data requests, or use of another mechanism).• Outline the plan for data preservation and give information on how long the data will be retained.• Explain when the data will be made available. Indicate the expected timely release. Explain whether exclusive use of the data will be claimed and if so, why and for how long. Indicate whether data sharing will be postponed or restricted for example to publish, protect intellectual property, or seek patents.• Indicate who will be able to use the data. If it is necessary to restrict access to certain communities or to apply a data sharing agreement, explain how and why. Explain what action will be taken to overcome or to minimise restrictions.
5b	How will data for preservation be selected, and where data will be preserved long-term (for example a data repository or archive)?	<ul style="list-style-type: none">• Indicate what data must be retained or destroyed for contractual, legal, or regulatory purposes.• Indicate how it will be decided what data to keep. Describe the data to be preserved long-term.• Explain the foreseeable research uses (and/or users) for the data.• Indicate where the data will be deposited. If no established repository is proposed, demonstrate in the data management plan that the data can be curated effectively beyond the lifetime of the grant. It is recommended to demonstrate that the repositories policies and procedures (including any metadata standards, and costs involved) have been checked.
5c	What methods or software tools are needed to access and use data?	<ul style="list-style-type: none">• Indicate whether potential users need specific tools to access and (re-)use the data. Consider the sustainability of software needed for accessing the data.• Indicate whether data will be shared via a repository, requests handled directly, or whether another mechanism will be used?
5d	How will the application of a unique and persistent identifier (such as a Digital Object Identifier (DOI)) to each data set be ensured?	<ul style="list-style-type: none">• Explain how the data might be re-used in other contexts. Persistent identifiers should be applied so that data can be reliably and efficiently located and referred to. Persistent identifiers also help to track citations and re-use.• Indicate whether a persistent identifier for the data will be pursued. Typically, a trustworthy, long-term repository will provide a persistent identifier.

SELEZIONA I DATI DA PRESERVARE



Non tutti i dati che utilizzi nella tua ricerca sono necessari per la sua comprensione, verifica e riproducibilità. Considera quali è utile preservare.



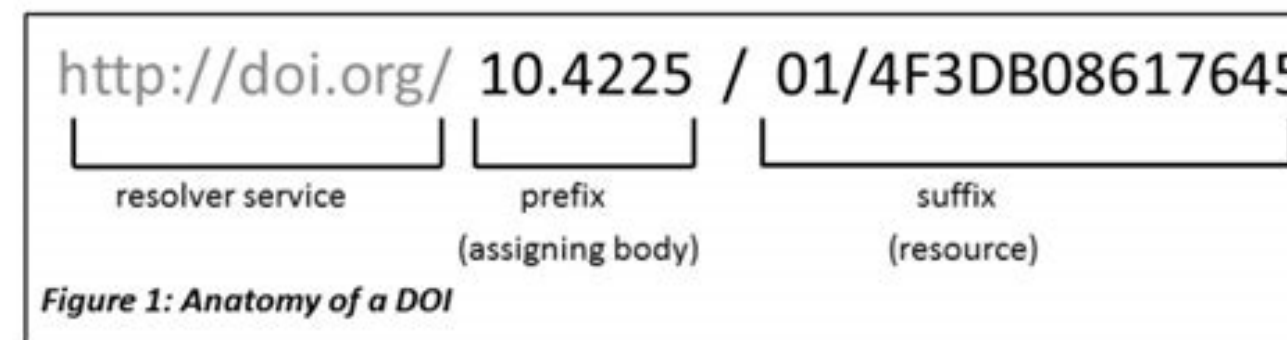
Tieni i dati che potrebbero esserti utili per progetti di ricerca futuri e quelli che puoi raccogliere in una sola occasione o che hai ottenuto dopo notevoli sforzi personali.



In alcuni casi i dati non possono essere depositati per (mancanza di) accordi precedentemente presi con parti terze.

I REPOSITORY: IDENTIFICATORI PERSISTENTI

- I repository di dati permettono di assegnare un **identificatore persistente** (PID) al tuo dataset
- Lo scopo principale del PID è **fornire le informazioni necessarie per identificare, verificare e localizzare** la risorsa in modo affidabile (non si rompono come i normali URL)
- Un PID può essere collegato a un insieme di metadati che descrivono un elemento piuttosto che all'elemento stesso.



IDENTIFICATORI PRETISTENTI (PID)

Il PID è insomma un **riferimento duraturo a una risorsa**, che può essere:

- una persona (ricercatori, autori, collaboratori: e.g. ORCID, ISNI)
- un'organizzazione (e.g. ROR)
- un «oggetto» digitale oppure no (pubblicazioni, dati, software: e.g. DOI, ISBN, Handle)



Publication date:
October 12, 2022

DOI:
[DOI 10.5281/zenodo.7190005](https://doi.org/10.5281/zenodo.7190005)

Keyword(s):
FAIR Research Data Management Privacy Ethics
Intellectual Property Storage Preservation

License (for files):
[Creative Commons Attribution 4.0 International](#)

Versions

Version	Date
Version 1	Oct 12, 2022

[10.5281/zenodo.7190005](https://doi.org/10.5281/zenodo.7190005)

Cite all versions? You can cite all versions by using the DOI [10.5281/zenodo.7190004](https://doi.org/10.5281/zenodo.7190004). This DOI represents all versions, and will always resolve to the latest one. [Read more.](#)

PERCEIVE project - Deliverable D8.5 "PERCEIVE (Perception and Evaluation of Regional and Cohesion policies by Europeans and Identification with the Values of Europe): Data Management Plan. Version (M6)"

A cura di: Mollona, Edoardo ; Pareschi, Luca ; Reverberi, Pierre ; Brasili, Cristina (2017) *PERCEIVE project - Deliverable D8.5 "PERCEIVE (Perception and Evaluation of Regional and Cohesion policies by Europeans and Identification with the Values of Europe): Data Management Plan. Version (M6)".* p. 28 DOI [10.6092/unibo/amsacta/5551](https://doi.org/10.6092/unibo/amsacta/5551).

Salva citazione | Condividi | Citato da

I REPOSITORY: SCHEMI DI METADATI



METADATI: informazioni strutturate sui dati. Si riferiscono a informazioni che descrivono, spiegano, individuano o facilitano il recupero, l'uso o la gestione di una risorsa.

- Esistono **schemi** standard di metadati, **sia generici che specifici per disciplina**.
- Un **repository generalista**, che contiene tante risorse diverse, avrà dei metadati più generici.
- Un **repository disciplinare** può consentire dei metadati più dettagliati e specifici per la materia.

I REPOSITORY: SCHEMI DI METADATI

Basic informationrequired

Digital Object Identifier

10.5281/zenodo.6123290

Optional. Did your publisher already assign a DOI to your upload? If not, leave the field empty and we will register a new DOI for you. A DOI allows others to easily and unambiguously cite your upload. Please note that it is NOT possible to edit a Zenodo DOI once it has been registered by us, while it is always possible to edit a custom DOI.

Reserve DOI

Publication date

2022-03-15

Required. Format: YYYY-MM-DD. In case your upload was already published elsewhere, please use the date of first publication.

Title

What do we mean by "data" in the arts and humanities? Interview transcripts (Universi

Required.

Authors

Gualandi, BiancaUniversity of Bologna

Pareschi, LucaUniversity of Rome Tor Vergata

Peroni, SilvioUniversity of Bologna

+ Add another author

Description

B

I

S

x₂

x²

This dataset contains the anonymised transcripts of the interviews conducted between N department of Classical Philology and Italian Studies (FICLIT) at the University of Bologn analysis of the interviews, carried out using a grounded theory approach and the open so

Licenserequired

Access right

☒ Open Access

☐ Embargoed Access

☐ Restricted Access

☐ Closed Access

Required. Open access uploads have considerably higher visibility on Zenodo.

License

Creative Commons Zero v1.0 Universal

Required. Selected license applies to all of your files displayed on the top of the form. If you want to upload some of your files under different licenses, please do so in separate uploads. If you cannot find the license you're looking for, include a relevant LICENSE file in your record and choose one of the Other licenses available (Other (Open), Other (Attribution), etc.). The supported licenses in the list are harvested from opendefinition.org and spdx.org. If you think that a license is missing from the list, please contact us.

Fundingrecommended

Related/alternate identifiersrecommended

Specify identifiers of related publications and datasets. Supported identifiers include: DOI, Handle, ARK, PURL, ISSN, ISBN, PubMed ID, PubMed Central ID, ADS Bibliographic Code, arXiv, Life Science Identifiers (LSID), EAN-13, ISTC, URNs and URLs.

Related identifiers

10.5281/zenodo.6385132

has this upload as its source

Thesis

Optional. Resource type of the related identifier.

+ Add another related identifier

Metadata of Gualandi, Bianca, Pareschi, Luca, & Peroni, Silvio. (2022). What do we mean by "data" in the arts and humanities? Interview transcripts (University of Bologna, FICLIT) and qualitative data coding [Data set]. Zenodo. <https://doi.org/10.5281/zenodo.6123290>



SCHEMI DI METADATI

Per approfondire il tema ed esplorare standard di metadati disciplina-specifici, puoi utilizzare:

- **FAIRsharing** (<https://fairsharing.org/search?fairsharingRegistry=Standard>)
- **Research Data Alliance** (<https://rd-alliance.github.io/metadata-directory/standards/>)
- **DCC Metadata Standard Selector** (<https://www.dcc.ac.uk/guidance/standards/metadata>).

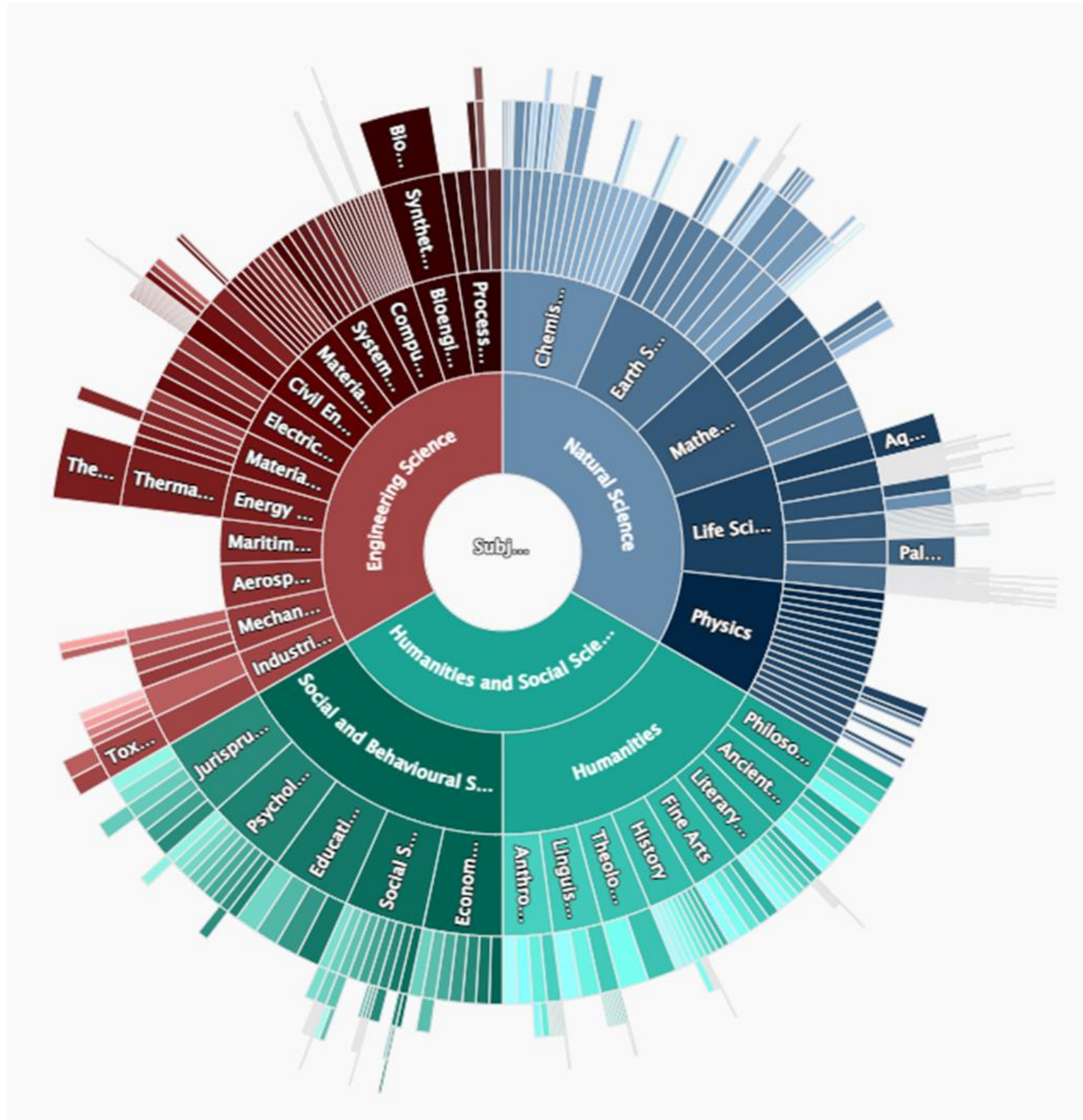


I REPOSITORY: INTEROPERABILITÀ

I repository forniscono anche un livello base di interoperabilità:

- supportano **standard per l'interoperabilità** e lo scambio di metadati (e.g. OAI-PMH)
- possono **raccomandare l'uso di ontologie o vocabolari specifici** (in particolare se sono repository disciplinari)

ONTOLOGIE E VOCABOLARI



- Ontologie e vocabolari per dati e metadati sono **best practices** di interoperabilità approvate dalla comunità
- Servono a **descrivere concetti e relazioni** all'interno di un dominio di conoscenza.
- Molti vocabolari e ontologie si possono trovare sul web: <https://fairsharing.org/browse/subject>

I REPOSITORY: CONDIZIONI DI ACCESSO E LICENZE








Una licenza **specifica ciò che gli altri possono fare con i dati** una volta che sono stati depositati.

- I dati sono **sempre coperti da copyright**, anche quando non includono una licenza!
- Se i dati non includono una licenza, **devi ottenere l'autorizzazione** dal creatore per copiare, condividere o pubblicare in altro modo questi dati
- Includere una licenza è sempre fortemente raccomandato. Esistono **strumenti online** per determinare quale licenza è più appropriata per il tuo set di dati, e.g.,
<https://creativecommons.org/choose/>,
<https://ufal.github.io/public-license-selector/>


LICENZE CREATIVE COMMONS


Permettono di rilasciare pubblicazioni e i dati con una licenza di copyright **chiara e *machine-readable***.

- È possibile scegliere diverse condizioni di riutilizzo, in diverse combinazioni, dalla più aperta (CC 0) alla più chiusa (CC BY-NC-ND)
- Permettono agli altri di capire se e come riutilizzare i tuoi materiali senza dover affrontare inutili ostacoli.


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	✓	✓	✗	✗	✗


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 REPOSITORY: CONDIZIONI DI ACCESSO E LICENZE

Dati personali? Devi seguire il GDPR e proteggere la privacy

Dati sensibili? Devi seguire il GDPR, potresti dover coinvolgere comitati etici

Riutilizzi dati protetti da **copyright**?

Vuoi **brevettare**?

Vuoi pubblicare (magari una monografia) e **temi che le idee possano essere rubate?**

Possibili strategie:

- Anonimizzazione
- **Accesso controllato**
- **Comitato di accesso**

Controlla le licenze! Il riutilizzo può essere limitato, come l'accessibilità di nuovi risultati

Imponi un **embargo**

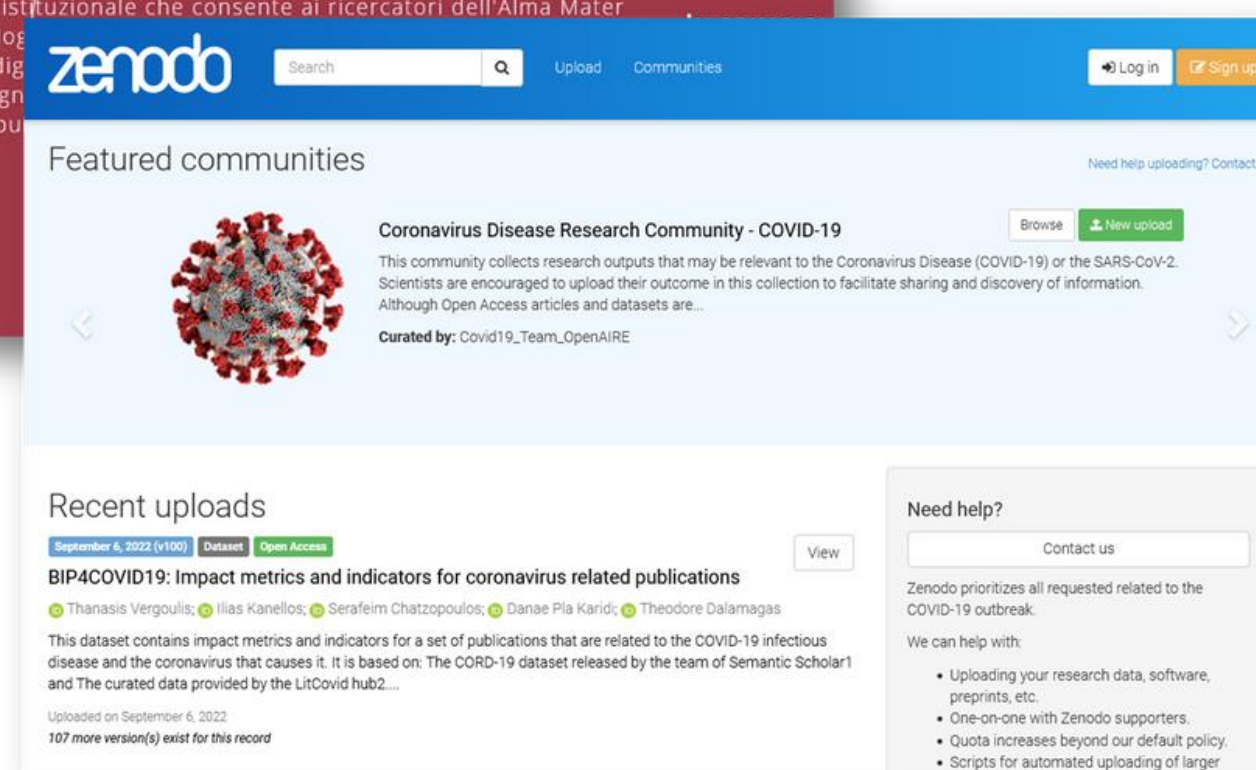
COME SCEGLIERE DOVE DEPOSITARE I DATI?



1. Utilizza re3data per trovare un **repository disciplinare** OPPURE



2. Utilizza il **repository della tua istituzione** OPPURE



3. Utilizza un **repository generalista** come Zenodo



Repository details

Zenodo

General Institutions Terms Standards

Name of repository	Zenodo
Additional name(s)	Research. Shared
Repository URL	https://zenodo.org/
Subject(s)	Humanities and Social Sciences Life Sciences Natural Sciences Engineering Sciences
Description	ZENODO builds and operates a simple and innovative service that enables researchers, scientists, EU projects and institutions to share and showcase multidisciplinary research results (data and publications) that are not part of the existing institutional or subject-based repositories of the research communities. ZENODO enables researchers, scientists, EU projects and institutions to: easily share the long tail of small research results in a wide variety of formats including text, spreadsheets, audio, video, and images across all fields of science. display their research results and get credited by making the research results citable and integrate them into existing reporting lines to funding agencies like the European Commission. easily access and reuse shared research results.
Contact	info@zenodo.org https://zenodo.org/contact
Content type(s)	Standard office documents Networkbased data Images Structured graphics Audiovisual data Scientific and statistical data formats Raw data Plain text Structured text Archived data other Source code

Repository details

Zenodo

General Institutions Terms Standards

Persistent identifier system(s)	DOI
Name of the repository software	other
Data citation guideline	https://about.zenodo.org/
Author Identifier system(s)	ORCID
Enhanced Publication	unknown
Quality management	yes
Application programming interfaces (2)	
API type	OAI-PMH
API type	REST
Metadata standards (2)	
Metadata standard name	DataCite Metadata Schema
Metadata standard scheme	DCC
Metadata standard name	Dublin Core
Metadata standard scheme	DCC

Repository details

AMS Acta

General Institutions Terms Standards

Name of repository

AMS Acta

Additional name(s)

Institutional Research Repository
 Alma Mater Studiorum Acta
 AMSActa

Repository URL

<https://amsacta.unibo.it/>

Subject(s)

Humanities and Social Sciences Life Sciences Natural Sciences Engineering Sciences

Description

AMS Acta is the institutional open access repository which enables the researchers of the Alma Mater Studiorum - University of Bologna to share, preserve and showcase their scientific results making them easily accessible, citable and reusable. The repository collects and disseminates scientific publications, research data and preprints. AMS Acta is the archiving infrastructure for the University of Bologna's researchers participating in the European programme H2020 that requires open access and the mandatory deposit of all funded peer-reviewed publications in a repository. AMS Acta collects both publications and research data so it is also a suitable archiving solution for the funded projects participating in the H2020 Open Research Data Pilot.

Contact

almadi@unibo.it
 eprints@cib.unibo.it

Content type(s)

Structured text Scientific and statistical data formats Standard office documents

Repository details

AMS Acta

General Institutions Terms Standards

Persistent Identifier system(s)

DOI

Name of the repository
 software

EPrints

Versioning

yes

Author identifier system(s)

ORCID

Enhanced Publication

yes

Quality management

unknown

Application programming interfaces (1)

API type

OAI-PMH

Metadata standards (1)

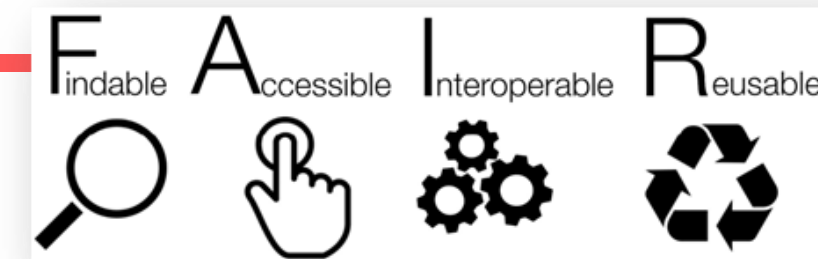
Metadata standard name

Dublin Core

Metadata standard scheme

DCC

IL RUOLO DEL REPOSITORY NEL GESTIRE I DATI IN MODO FAIR



Assegnano ai dataset identificatori persistenti e un set di metadati

Conservano dati e i metadati e li rendono persistenti nel tempo e rintracciabili in rete. I dati stessi possono essere apertamente accessibili, o accessibili attraverso un sistema di autenticazione e autorizzazione, se la loro natura ne impedisce l'apertura.



I possibili riutilizzi dei dati vengono dichiarati tramite l'utilizzo di licenze

CONDIVISIONE E PRESERVAZIONE A LUNGO TERMINE: QUALCHE SUGGERIMENTO

- Deposita i dati necessari a **verificare indipendentemente** la tua ricerca. Se hai scartato determinati dati perché non particolarmente rilevanti spiega il processo di decisione nel tuo DMP
- Crea una tabella nella quale indicare il **repository che intendi utilizzare**, quale schema di metadati supporta, che licenze consente di attribuire e quali PID fornisce.



Image credit: <http://aukeherrema.nl> CC-BY

UN ESEMPIO PRATICO

DMP identifier	Licensing and restrictions	Quality assurance processes	Repository for long-term preservation
Workshop Annotation	CC-BY 4.0	-	Zenodo
UM	CC-BY 4.0	Internal peer review. Unit tests and integration tests for continuous development.	-
Community	GPL 3.0	Unit and system tests	Zenodo
Recomm	CC-BY 4.0	User testing	Zenodo
LDH	CC-BY 4.0	Unit and system tests	Zenodo
LDHReg	Apache 2.0	Unit testing, Code review, Issue tracking, User testing	Zenodo
IGmuseum	CC-BY-NC 4.0	RDF data validation	Zenodo
IGclassifier	CC-BY 4.0	RDF data validation	Zenodo
IUcitizen	CC-BY 4.0	Issue tracking, User testing	Zenodo, AMSActa
Reasoner	MIT	Unit testing, Code review, Issue tracking, User testing	Zenodo
Ontologies	Apache 2.0	Unit testing, Code reviewing, Issue tracking	Zenodo
IMMAdata	CC-BY 4.0	Ontology consistency validation, Usability	Zenodo, AMSActa
IMMAPilot	CC-BY-NC 4.0	RDF data validation	Zenodo
MuOpinion	Apache 2.0	Unit testing, Pair programming, User tests	Zenodo
HaifaPilot	CC-BY 4.0		Zenodo
Demog	CC-BY 4.0		Zenodo
Ethnog	CC-BY-NC 4.0		Zenodo
PortableMu	CC-BY-NC 4.0		Zenodo
VRpopup	CC-BY-NC 4.0		Zenodo
FTMdata	CC-BY 4.0	Data quality is ensured by the application used for collecting data (IUcitizen).	Zenodo

Table 26 Reusability of datasets and software

	Reuse or development of metadata standards	Persistent identifiers
	Zenodo Metadata set, CLAUDIA Metadata set	Zenodo (DOI)
	-	-
	Zenodo Metadata set	Zenodo (DOI)
	Zenodo Metadata set	Zenodo (DOI)
	Zenodo Metadata set	Zenodo (DOI)
ctive	Zenodo Metadata set	Zenodo (DOI)
	Zenodo Metadata set	Zenodo (DOI)
	Zenodo Metadata set	Zenodo (DOI)
	Zenodo Metadata set, AMSActa Metadata set	Zenodo (DOI), AMSActa (DOI)
ctive	Zenodo Metadata set	Zenodo (DOI)
oc	Zenodo Metadata set	Zenodo (DOI)
	Zenodo Metadata set, AMSActa Metadata set	Zenodo (DOI), AMSActa (DOI)



LE SEZIONI DI UN DATA MANAGEMENT PLAN

seguendo il template di
Science Europe

1. Descrizione dei dati e della loro raccolta/del riutilizzo di dati già esistenti
2. Documentazione e qualità dei dati
3. Storage e backup durante il processo di ricerca
4. Richieste etiche e legali, codici di condotta
5. Condivisione e preservazione a lungo termine dei dati
- 6. Responsabilità e costi della gestione dei dati**

6. RESPONSABILITÀ E COSTI

Nell'ambito dei progetti di ricerca che siano portati avanti da più ricercatori è opportuno **definire anticipatamente i ruoli e le responsabilità** sia in termine di gestione dei dati sia in termini di compilazione e di gestione del DMP.

La gestione dei dati della ricerca che può essere **dispendiosa** sia in termini di tempo che di altri costi, per esempio:

- **Raccolta** dei dati, e.g., trascrizione da file audio
- **Descrizione** e documentazione
- **Archiviazione** (se i dati sono di grandi dimensioni)
- **Conservazione**, e.g. conversione di dati/file.

6 Data management responsibilities and resources

- | | |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 6a Who (for example role, position, and institution) will be responsible for data management (<u>i.e.</u> the data steward)? | <ul style="list-style-type: none">• Outline the roles and responsibilities for data management/stewardship activities for example data capture, metadata production, data quality, storage and backup, data archiving, and data sharing. Name responsible individual(s) where possible.• For collaborative projects, explain the co-ordination of data management responsibilities across partners.• Indicate who is responsible for implementing the DMP, and for ensuring it is reviewed and, if necessary, revised• Consider regular updates of the DMP. |
| 6b 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)? | <ul style="list-style-type: none">• Explain how the necessary resources (for example time) to prepare the data for sharing/preservation (data curation) have been costed in. Carefully consider and justify any resources needed to deliver the data. These may include storage costs, hardware, staff time, costs of preparing data for deposit, and repository charges.• Indicate whether additional resources will be needed to prepare data for deposit or to meet any charges from data repositories. If yes, explain how much is needed and how such costs will be covered. |

TUTELA DELLA PROPRIETÀ INTELLETTUALE: QUALCHE SUGGERIMENTO

The cost of data management

How to calculate costs?

How to use this costing tool?

Estimating costs RDM tool

Estimating costs RDM tool

DMP PHASE	ACTIVITY	COMMENTS AND SUGGESTIONS	COSTS
Preparing	Make a Data Management Plan	Make a DMP before you start creating data; make decisions about managing your data. You can find the template for H2020 DMPs here . Check if there is a department within your organization to support data management planning.	2 hrs to 2 days, depending on the complexity of your project
1. Data Collection	Acquiring External datasets Do you plan to use existing data, and is the data available at a commercial partner?	Your library may be able to help you acquire a license to a crucial database In research data repositories, data can be available at no or low costs	<u>Example:</u> A faculty licence on a database for macro-economic analysis: €18.000/y

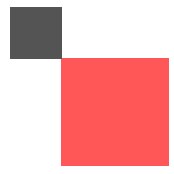
How to use this costing tool?

Step 1: Check the data management activities in the table and tick those that may apply to your proposed research.

Step 2: For each selected activity, estimate the additional time and/or other resources needed and cost this, e.g. people’s time or physical resources needed such as hardware or software. Find out which resources, e.g. for data storage and backup, are available to you from your institution. Consider whether you need a dedicated data manager.

Step 3: Add these data management costs to your research application. Coordinate resourcing and costing with your institution, research office and institutional IT services.

Step 4: Plan the data management activities in advance to avoid them competing with the need to focus on research excellence.



QUALCHE PUNTO CHIAVE



- Le responsabilità etiche e legali nella ricerca influenzano le modalità di raccolta, gestione, conservazione e condivisione dei dati e possono essere di diversa natura.
- Non tutti i dati che utilizzi nella tua ricerca sono necessari per la sua comprensione, verifica e riproducibilità.
- Un repository per i dati non fa tutto il lavoro per rendere i tuoi dati il più FAIR possibile. Tuttavia, offre un'ottima struttura per ottenere le basi giuste.
- Descrivi tutti i passaggi che abbiamo visto nel tuo DMP.



CONSIGLI DI APPROFONDIMENTO

Open Science:

Open Science: A Practical Guide for Early-Career Researchers (1.0)

<https://doi.org/10.5281/zenodo.7716153>

Open Science Toolkit (UNESCO) <https://www.unesco.org/en/open-science/toolkit>

Portale Italiano Open-Science <https://open-science.it>

Webinar Open Science Cafè <https://www.icdi.it/it/attivita/competence-centre/open-science-cafe>

Gestione FAIR dei Dati di Ricerca:

Open AIRE guidelines <https://guidelines.openaire.eu/en/latest/>

The Turing Way handbook <https://the-turing-way.netlify.app/index.html>

The hands-on guide to research data management for KU Leuven researchers, students, and research support

staff in the humanities and social sciences. <https://doi.org/10.5281/zenodo.4596323>

LEGO® Metadata for Reproducibility game pack <http://eprints.gla.ac.uk/196477/>

Reprojuice, the reproducibility game <https://seriousgaming.tudelft.nl/game/reprojuice>

CONSIGLI DI APPROFONDIMENTO

Gestione FAIR dei Dati di Ricerca – strumenti e infrastrutture per discipline umanistiche:

DARIAH Pathfinder to Data Management Best Practices in the Humanities.

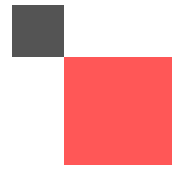
<https://campus.dariah.eu/resource/posts/dariah-pathfinder-to-data-management-best-practices-in-the-humanities>

Wuttke, U. Introduction to Humanities Research Data Management.

<https://doi.org/10.5281/zenodo.1491250>

Giglia, E. FAIR data in the Humanities. <https://doi.org/10.5281/zenodo.5510388>

The road to FAIR. <https://roadtofair.hypotheses.org/266>



GRAZIE

