

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

LEZIONE 3

Mario Marino

mario.marino6@unibo.it

Data Steward di Area Sociale, Research Division (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

Il dato e il suo
valore

Open Science

I principi FAIR

Research Data
Management

Data
Management
Plan

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 Stewards @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)

Esperienza in data management

FAIR principles e conoscenza di pratiche di OS





MARIO MARINO

Area Sociale



BIANCA GUALANDI

Area umanistica



GIULIA CALDONI

Area biomiedica



SARA COPPINI

Area tecnologica



RIPRENDIAMO LE FILA DEL DISCORSO...

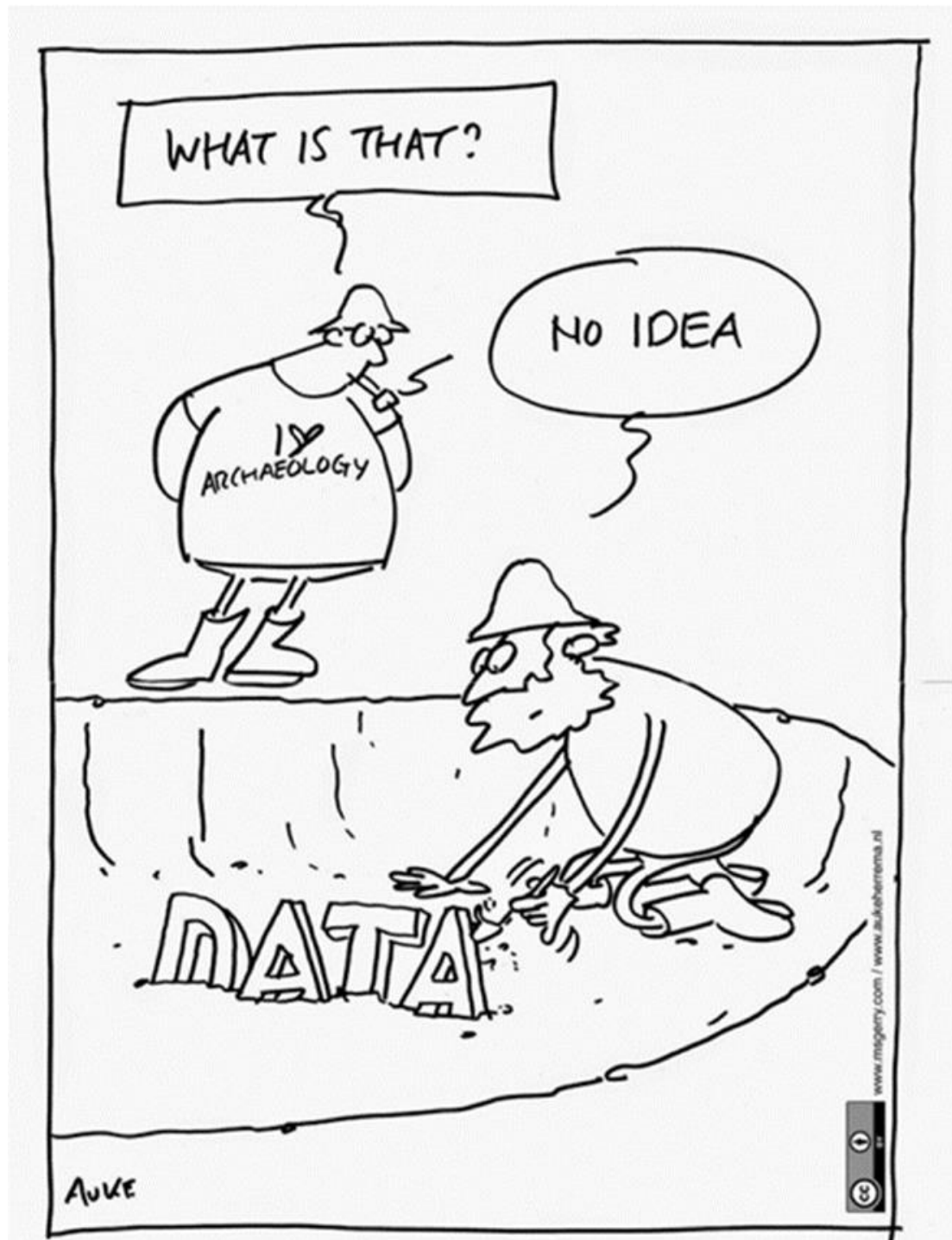


Image credit: <http://aukeherrema.nl> CC-BY

- 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...

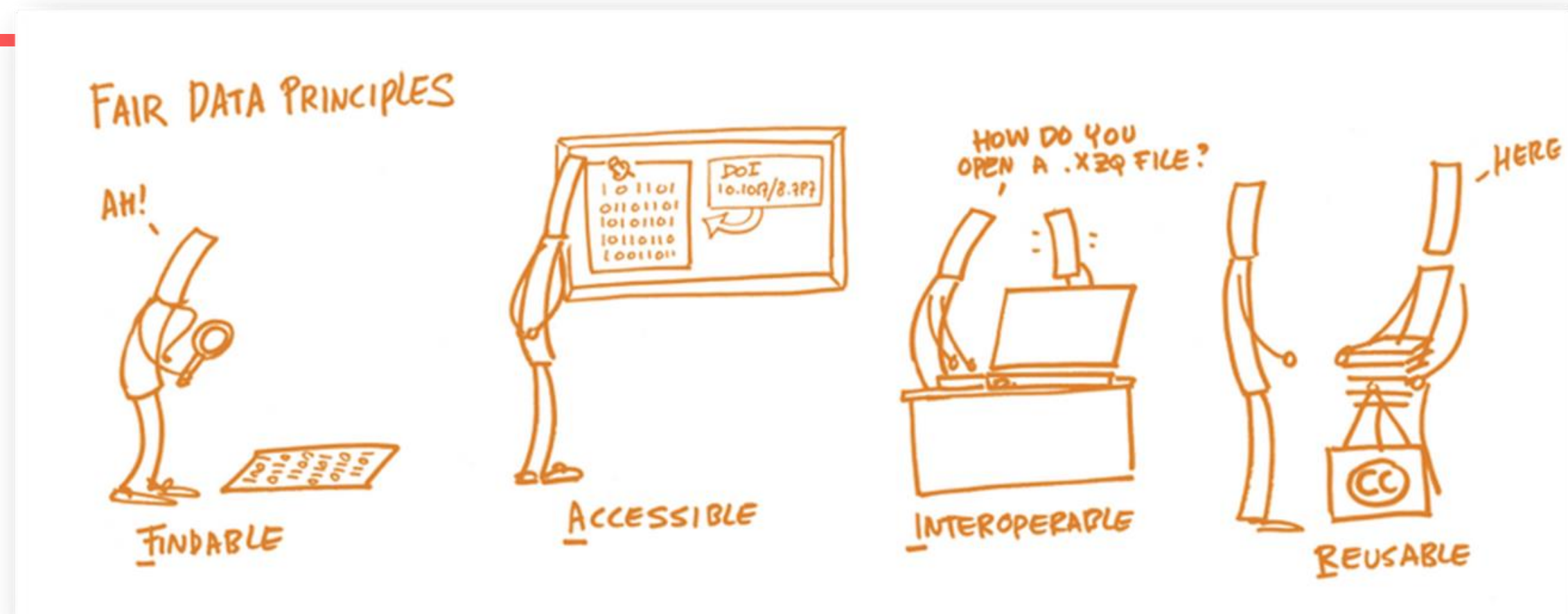


Image credit: <http://aukeherrema.nl> CC-BY

- 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 dati gestiti correttamente né tantomeno Open Data, "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...

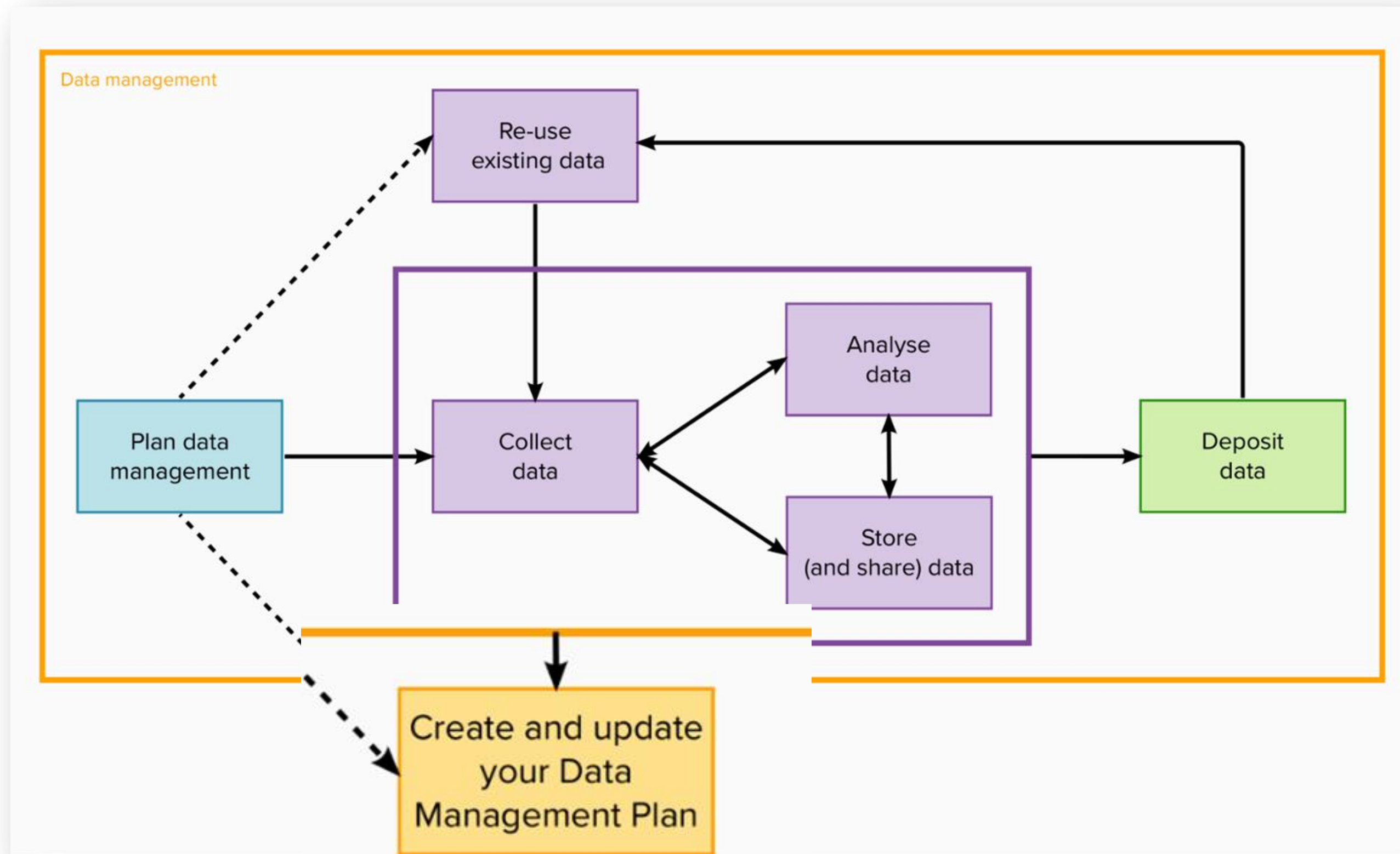


- Gestire attentamente i dati durante tutte le fasi del ciclo di vita del dato aiuta a rendere il processo della ricerca più efficiente.
- Le fasi attive della ricerca sono le più concitate: aver pianificato può aiutarti a portare avanti la ricerca in modo coerente.
- Non tutti i dati della ricerca sono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

DATA MANAGEMENT PLAN



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





COS'È UN DATA MANAGEMENT PLAN?

E' un documento strutturato che:

- Descrive i processi di acquisizione, analisi, gestione e conservazione dei dati durante un progetto di ricerca.
- Descrive chi è proprietario e responsabile dei dati.
- Descrive come i dati verranno preservati e condivisi durante e dopo il completamento di un progetto.



PERCHÈ REDIGERE UN DATA MANAGEMENT PLAN?

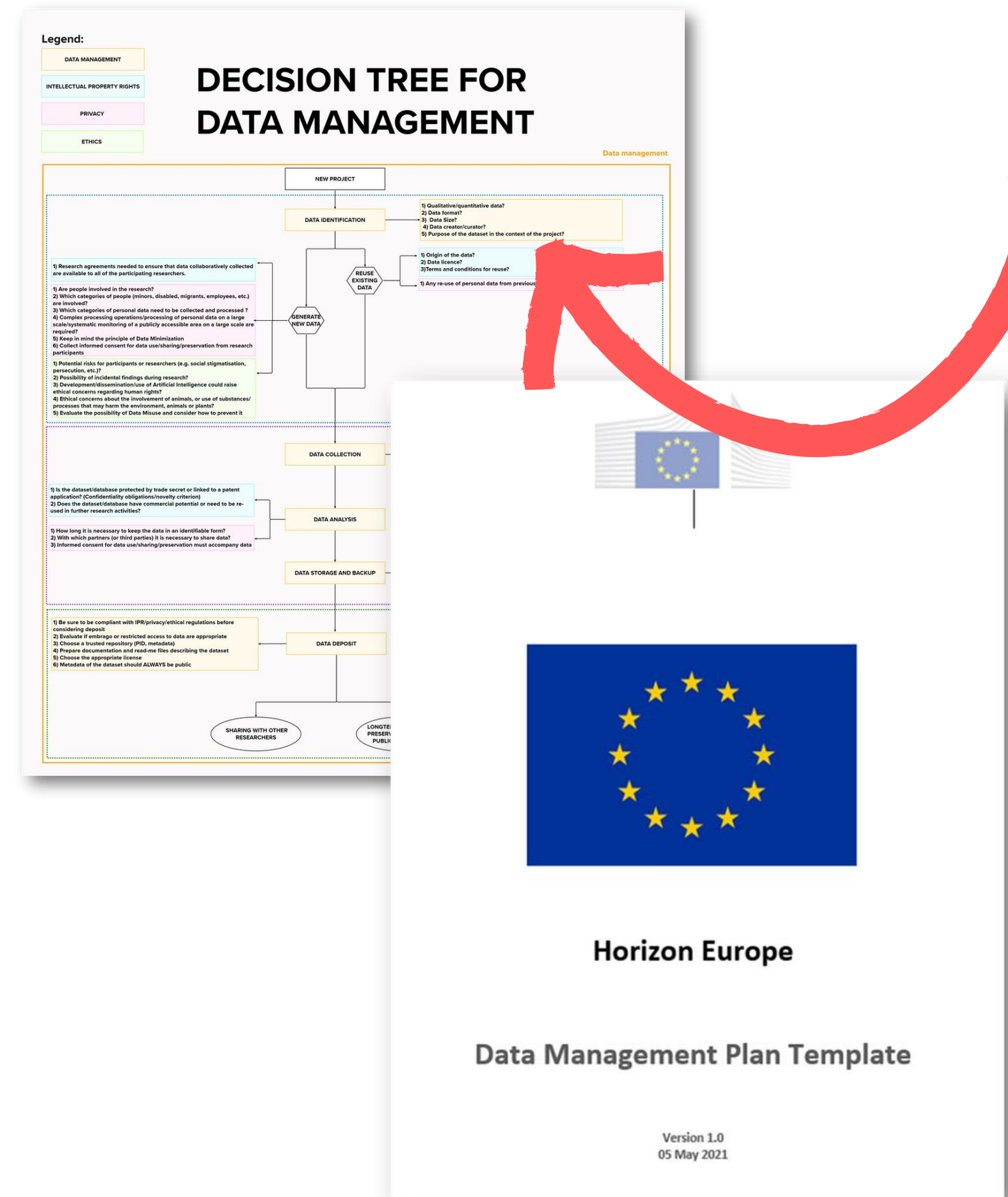
- Permette di identificare azioni chiave per assicurarsi che i dati della ricerca siano di qualità elevata, sicuri e sostenibili.
- Aiuta nel rendere FAIR i dati della ricerca e, dove possibile, disponibili apertamente.

DATA MANAGEMENT PLAN: DOCUMENTA LE RISPOSTE!

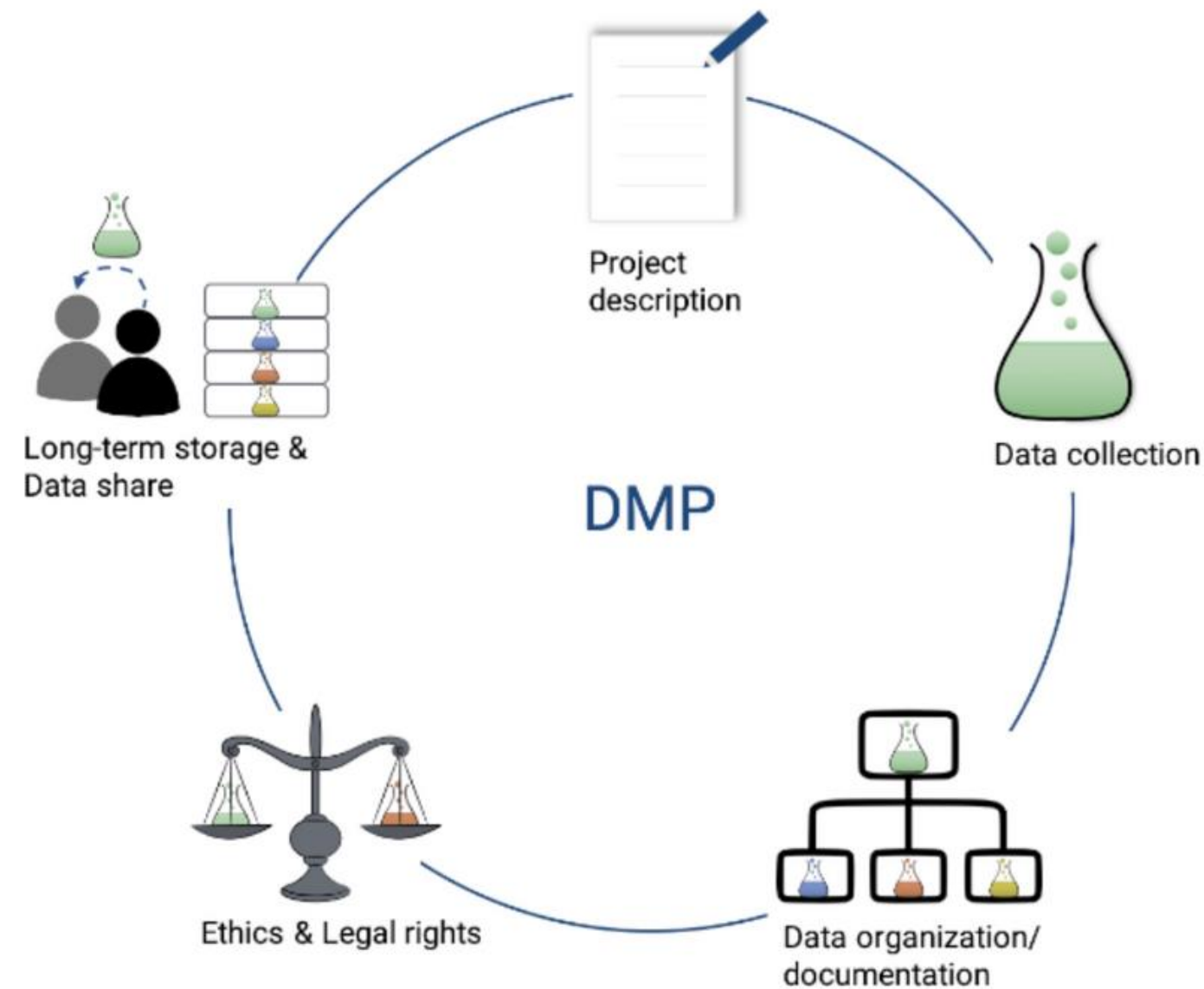
Il DMP viene redatto all'inizio del progetto e aggiornato nel corso del tempo.

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 consapevoli di come state adempiendo a tutti i vostri obblighi legali ed etici.



UN DATA MANAGEMENT PLAN DEVE INCLUDERE INFORMAZIONI SU:



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



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.
-  Migliorare la qualità delle tue applications.

CHI POTREBBE RICHIEDERE UN DATA MANAGEMENT PLAN?

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

https://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/data-management/data-management_en.htm

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



Horizon Europe

Data Management Plan Template

Version 1.0
05 May 2021

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.



Research Data Management Plans

Researchers **must create a data management plan (DMP)** at the time of their research proposal if any research data are to be collected or used. Plans should cover data types and volume, capture, storage, integrity, confidentiality, retention and destruction, sharing and deposit. A suggested minimum period of retention is three years from the end of the project, however requirements by funders and data providers, as well as disciplinary guidance should be observed.

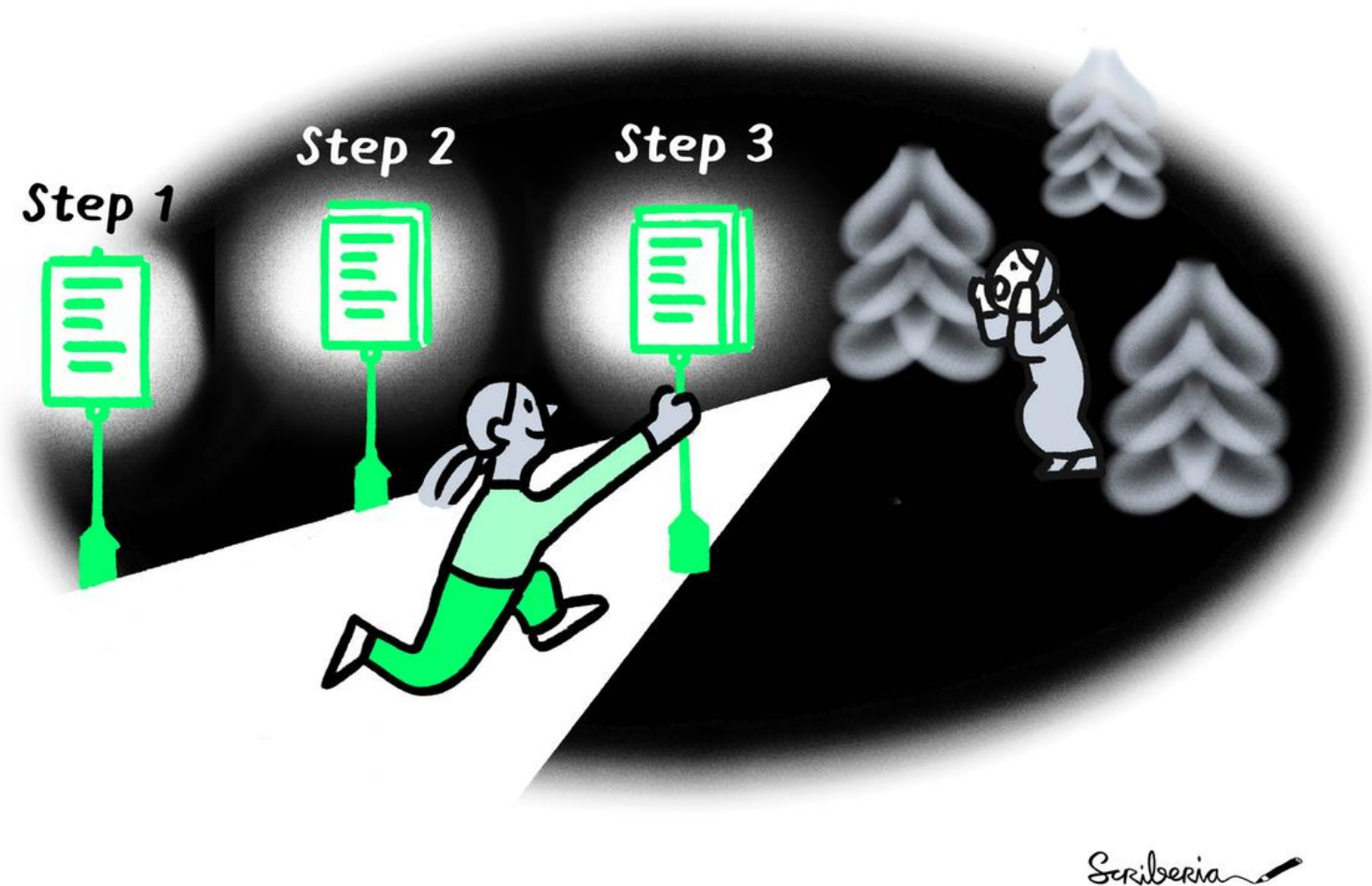


A further **general** guiding principle is that each and every staff member must act as a good steward in the management of relevant data.

1. Researchers

It is the responsibility of each **individual researcher** (or, in the case of a group of researchers, the research leader) to draw up a **Data Management Plan (DMP)** at the start of the research project and to follow up the agreements made in this plan. The DMP specifies the manner in which research data may be collected, stored, preserved, managed, documented and used. A sample DMP template is available on the Utrecht University intranet.

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.
- Non è necessario spiegare cosa significano "Open Access", "Open Science", "principi FAIR": piuttosto descrivi le azioni concrete che intraprenderai!

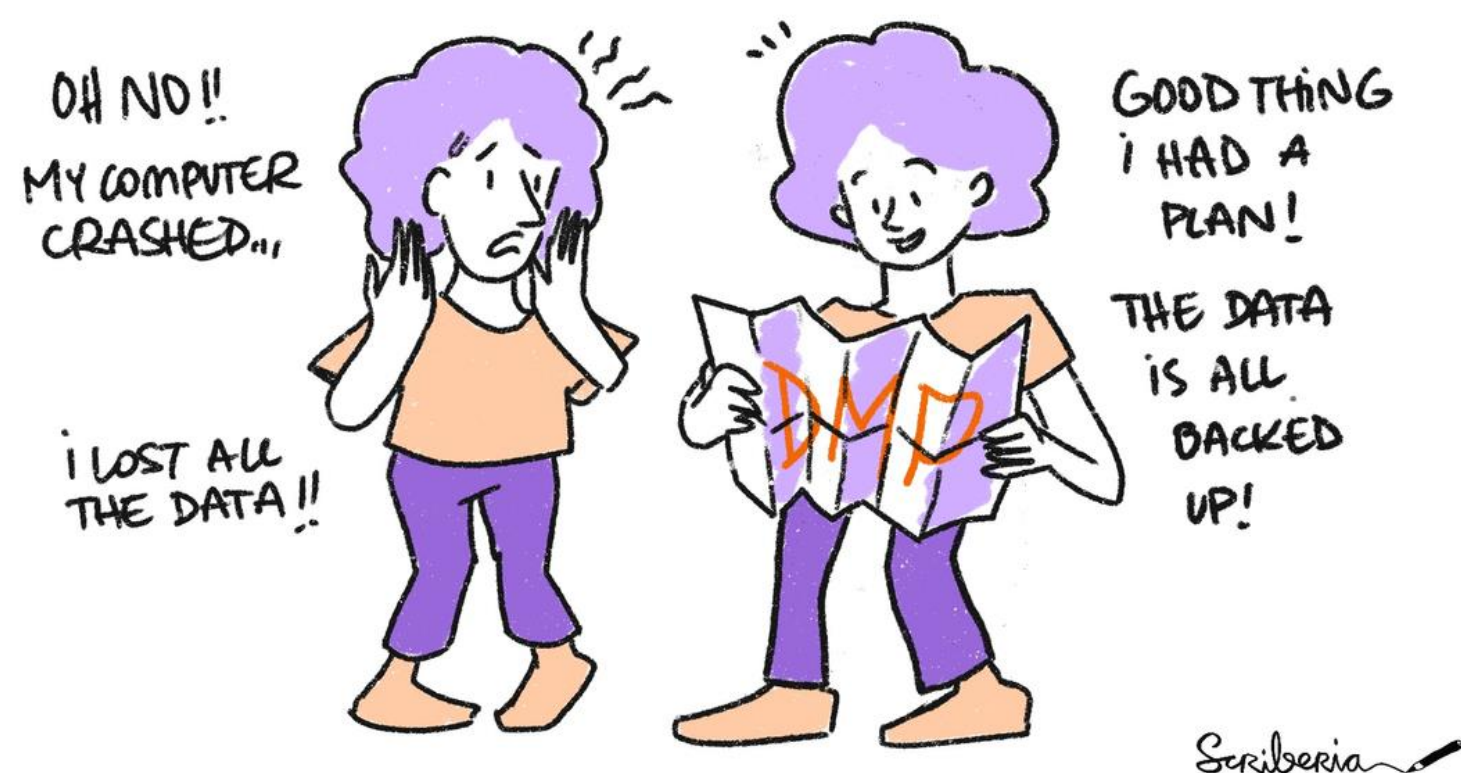
DATA MANAGEMENT PLAN: SUGGERIMENTI SULLA SCRITTURA

Non ci sono risposte sempre corrette per tutte le situazioni: **ogni progetto ha le sue specificità.**


Utilizza il DMP per provare di aver riflettuto su cosa fare e di aver **preso in considerazione l'approccio più ragionevole** per gestire correttamente la tua ricerca.

Giustifica le tue scelte e decisioni.

Non c'è una ricetta unica alla compilazione, quindi **non copiare** il lavoro degli altri!



DOVE TROVARE DEI DATA MANAGEMENT PLAN 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\(CC BY 4.0\)](#)
[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, it explains the way research data are handled, organized, licensed and made openly available to the public, and how they will be preserved after the project is completed. This DMP also provides motivations when versions or parts of the project research data cannot be openly shared on account of third-party copyright issues, confidentiality or personal data protection requirements, or when open dissemination could jeopardize the project achievements. This DMP reflects the current state of the art of PERCEIVE project (M6). The details and the final number of the project data sets may vary during the course of research. The variations will be recorded in updated versions of this DMP.


 **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.

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	



Upload Communities

 [giulia.caldoni2@unibo.it](#)

All versions

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 projec

Uploaded on December 18, 2020


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

D9.1 Data Management Plan

 Buresi Pascal;  Schoots Fleke;  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

View

 **argos**

[Home](#) [Public DMPs](#) [Public Dataset Desc.](#)


Published DMPs

Sort by: Published

DMP



Published: 29 March 2023

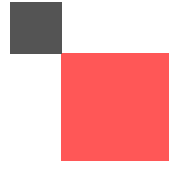
DIAMOND - Machine-Actionable Data Management Plan

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

Contained Datasets: (1)

DIAMOND D3.1 - Initial development plan: Surveys

 Export  All DMP Versions



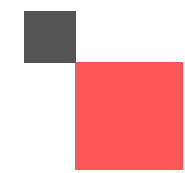
QUALCHE PUNTO CHIAVE



- Il Data Management Plan è un documento strutturato che ti permette di documentare tutte 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.
- E' un documento che può variare notevolmente da un progetto all'altro perchè ogni progetto ha le sue specificità. Non esiste una ricetta unica e giusta per compilarlo!



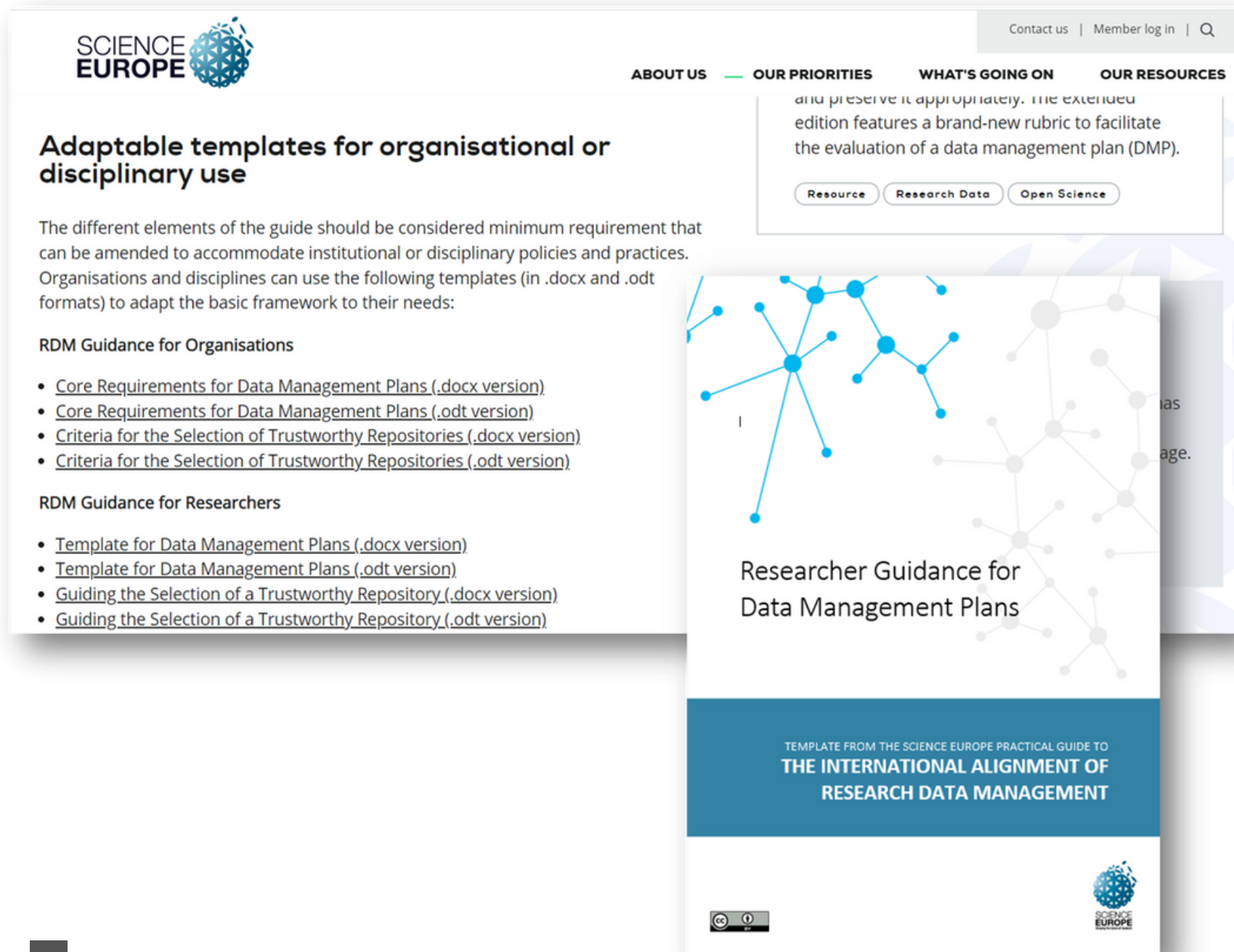




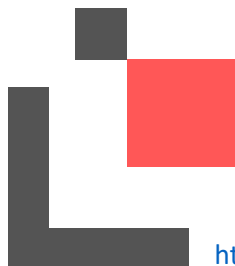
SCRIVERE UN DATA MANAGEMENT PLAN



DATA MANAGEMENT PLAN: I TEMPLATE

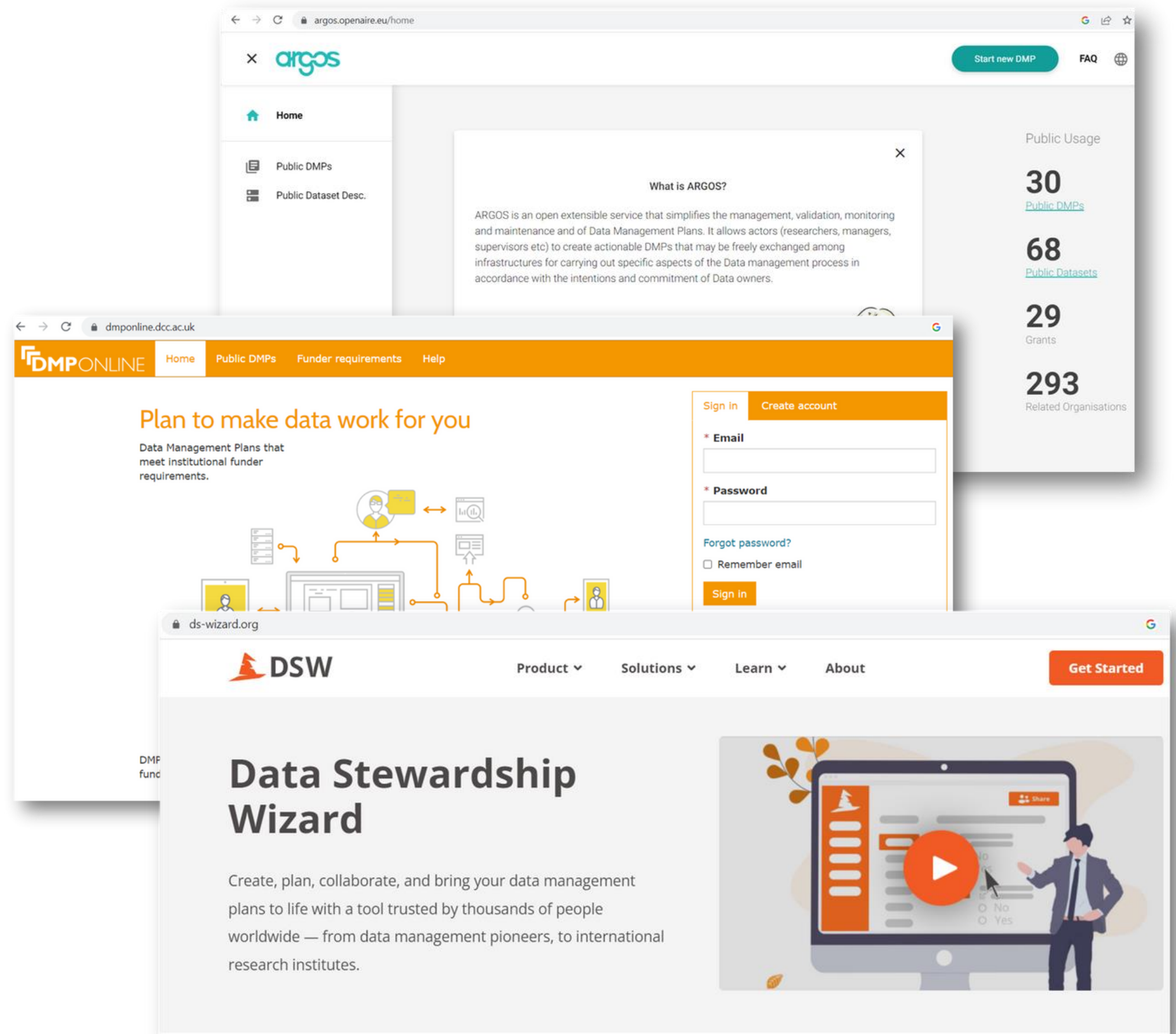


- Un template di Data Management Plan ti guida nella compilazione attraverso domande specifiche a cui rispondere.
- Il template di Science Europe raccoglie le informazioni minime da inserire.
- Horizon Europe ha basato il suo template su quello di Science Europe.

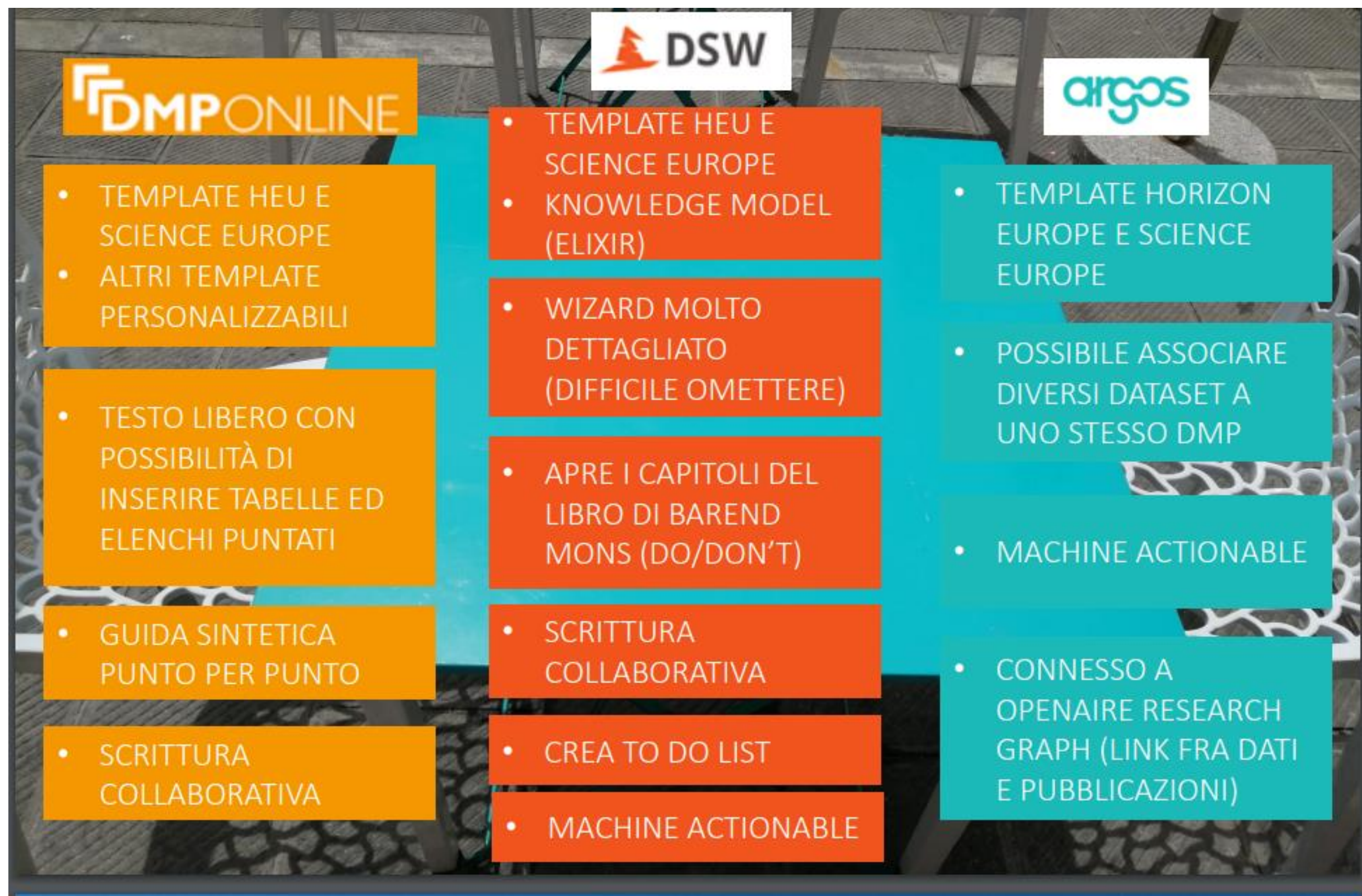


DATA MANAGEMENT PLAN: GLI STRUMENTI ONLINE

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



DATA MANAGEMENT PLAN: GLI STRUMENTI ONLINE





QUALCHE CONSIGLIO PRATICO PRIMA DI INIZIARE



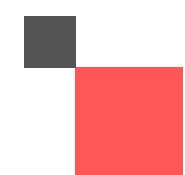
Identifica correttamente i dati della ricerca. Ricorda che il fondamento di un buon RDM è conoscere quali sono i tuoi dati.



Fai una lista completa dei materiali di ricerca e descrivi dettagliatamente aspetti tecnici e provenienza. Per farlo, segui i suggerimenti dati dal template che stai usando.



Interpreta correttamente le domande che ti vengono poste. Alcuni termini hanno significati specifici e spesso non intuitivi nel contesto del RDM.



LE SEZIONI DEL DATA MANAGEMENT PLAN





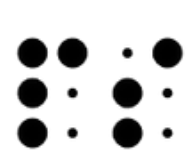
LE SEZIONI DI UN DATA MANAGEMENT PLAN

Seguiamo il template di
Science Europe

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

DESCRIZIONE DEI DATI E DELLA LORO RACCOLTA/DEL RIUSO DI DATI GIÀ ESISTENTI

Cosa deve essere descritto in un Data Management Plan?



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.

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 collected or produced.• State any constraints on re-use of existing data.• Explain how data provenance will be documented.• Briefly state the reasons if the re-use of any existing data 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> spreadsheets, textual (documents), image, audio, video, etc.• Give details on the data format: the way in which data is stored, often reflected by the filename extension (e.g. doc, txt, or pdf).• Justify the use of certain formats. For <u>example</u> on staff expertise within the host organisation, or on staff expertise within the host organisation, or on staff expertise within the research community, or on the software that will be used.• Give preference to open and standard formats and long-term reuse of data (several repositories 'preferred formats').• Give details on the volumes (they can be expressed in bytes), and/or in numbers of objects, etc.

DESCRIZIONE DEI DATI E DELLA LORO RACCOLTA/DEL RIUSO DI DATI GIÀ ESISTENTI:QUALCHE SUGGERIMENTO

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

- Tipi di dati
- Dati appena raccolti o riutilizzati?
- Come vengono creati i dati
- Nome breve e descrizione dei dati
- Formati di file dei dati
- Volume dei dati (stima)

DESCRIZIONE DEI DATI E DELLA LORO RACCOLTA/DEL RIUSO DI DATI GIÀ ESISTENTI: QUALCHE SUGGERIMENTO

- Se ancora non hai materialmente i dati descrivi quali dati hai intenzione di trattare e come.
- Delinea fin dalle prime fasi la struttura dei dataset.
- Il DMP di fine progetto dovrà essere comprensivo di una lista di tutti i dataset utili alla riproducibilità della 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/>
- British Film Institute: <https://www.bfi.org.uk/industry-data-insights>

GENERALISTI

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



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 DI DATI

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).



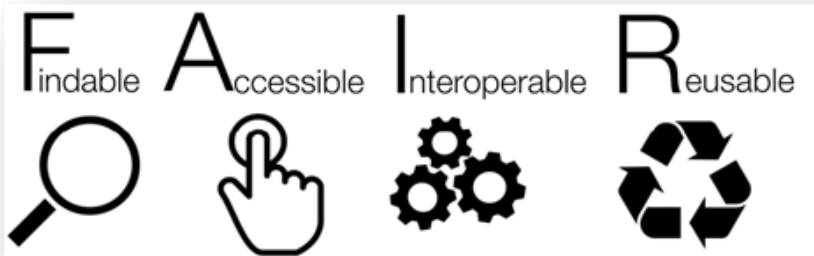
DIMENSIONI

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 (fate attenzione ai costi di archiviazione per grandi volumi di dati che potreste dover prendere in considerazione).

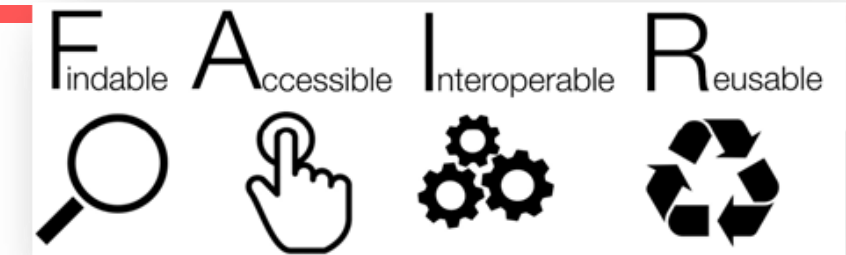
PERCHÈ SCEGLIERE DEI FORMATI "OPEN"?

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
Quantitative tabular data with extensive metadata. A dataset with variable labels, code labels, and defined missing values, in addition to the matrix of data.	Proprietary formats of statistical packages e.g. SPSS (.sav), Stata (.dta), .sas7bdat. Delimited text and command ('setup') file (SPSS, Stata, SAS, etc.) containing metadata information. Some structured text or mark-up file containing metadata information, e.g. DDI XML file.	 SPSS portable format (.por). MS Access (.mdb/.accdb).
Quantitative tabular data with minimal metadata. A matrix of data with or without column headings or variable names, but no other metadata or labeling.	Comma-separated values (CSV) file (.csv). Tab-delimited file (.tab). Including delimited text of given character set with SQL data definition statements where appropriate.	Delimited text of given character set – only characters not present in the data may be used as delimiters (.txt). Widely-used formats: MS Excel (.xls/.xlsx), MS Access (.mdb/.accdb), OpenDocument Spreadsheet (.ods).

PERCHÈ SCEGLIERE DEI FORMATI "OPEN"?



- I dati devono essere combinabili e utilizzabili con altri dati o strumenti.
- Il formato dei dati deve essere aperto e interoperabile da vari strumenti, compresi altri database.



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



UN ESEMPIO DI DESCRIZIONE DATI NEL DMP

Table 1
Summary of data format

Type of data	Formats used during data processing	Formats for sharing reuse and preservation
Numerical or textual tabular data	Microsoft Excel (.xls/.xlsx)	Comma-separated values (.csv) <i>Microsoft Excel (.xls/.xlsx) as an exception when format conversion to .csv was not possible because of the presence of special characters, as happened with dataset n°1 and 4.</i>
Qualitative textual data	Microsoft Word (.doc/.docx)	Rich Text Format (.rtf) or text (.txt)
Audio data	mp3 format (.mp3)	<i>Audio recordings were deleted and only the processed transcripts were shared and preserved.</i>
Topic modeling data	Mallet format (.mallet)	Comma-separated values (.csv)
Simulation model data	Text model format (.mdl)	<i>The mathematical model was saved using standard differential equations symbols in .csv and .txt files because the simulation model was developed using a proprietary software. Simulated values were saved as numerical data, as specified above.</i>
Statistical data	STATA format (.dta)	Comma-separated values (.csv), Stata format (.dta)

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)"
DOI 10.6092/unibo/amsacta/6244

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.

- 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.

- 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.

- 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

DOCUMENTAZIONE E QUALITÀ DEI DATI



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

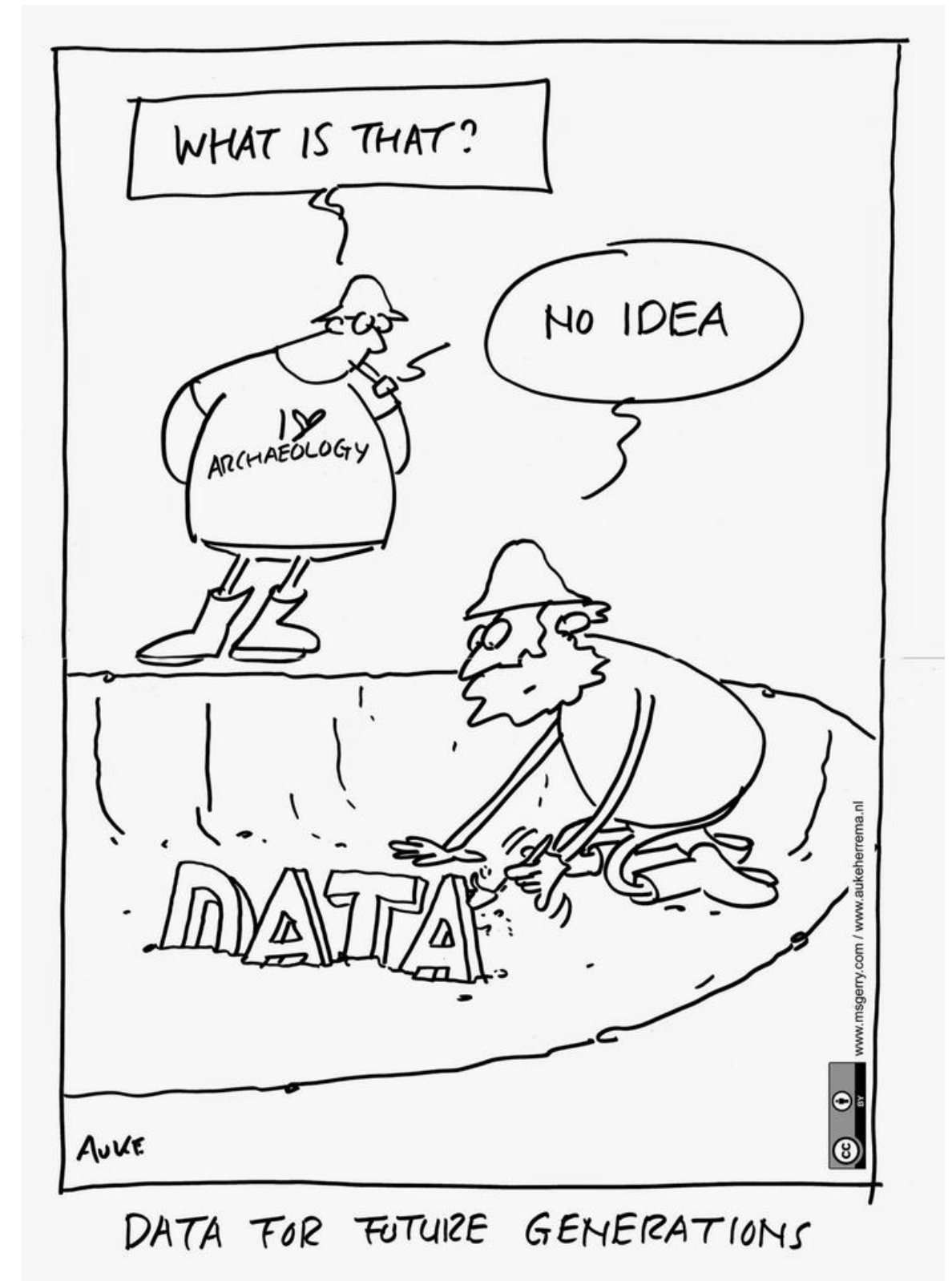


DOCUMENTAZIONE: informazioni che rendono i dati più comprensibili agli altri, permette di capirli e interpretarli anche molto dopo la raccolta.

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.

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.





METADATI

- Informazioni che aggiungono struttura ai dati per renderli machine-readable.
- Strutturali (su un oggetto in sé) o descrittivi (sul contenuto di un oggetto).
- È possibile utilizzare vocabolari controllati.
- 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.
- Human-readable.
- Molti ricercatori accompagnano i dati con un README file, che spiega come sono stati raccolti i dati, cosa significano i metadati ecc.
- Può anche essere "in-file", come i commenti nel codice.

COS'È UN READ ME FILE?

[INSERT PROJECT LOGO]

funded from the European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement N. [INSERT G.A. NUMBER]

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

[INSERT PROJECT LOGO]

funded from the European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement N. [INSERT G.A. NUMBER]

"PROJECTxyz_WP1_T1-2_FocusGroups_20181108_v01.csv(/.ods)"

1 README file

"README_PROJECTxyz_WP1_T1-2_FocusGroups_20181108_v01.rtf"

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]

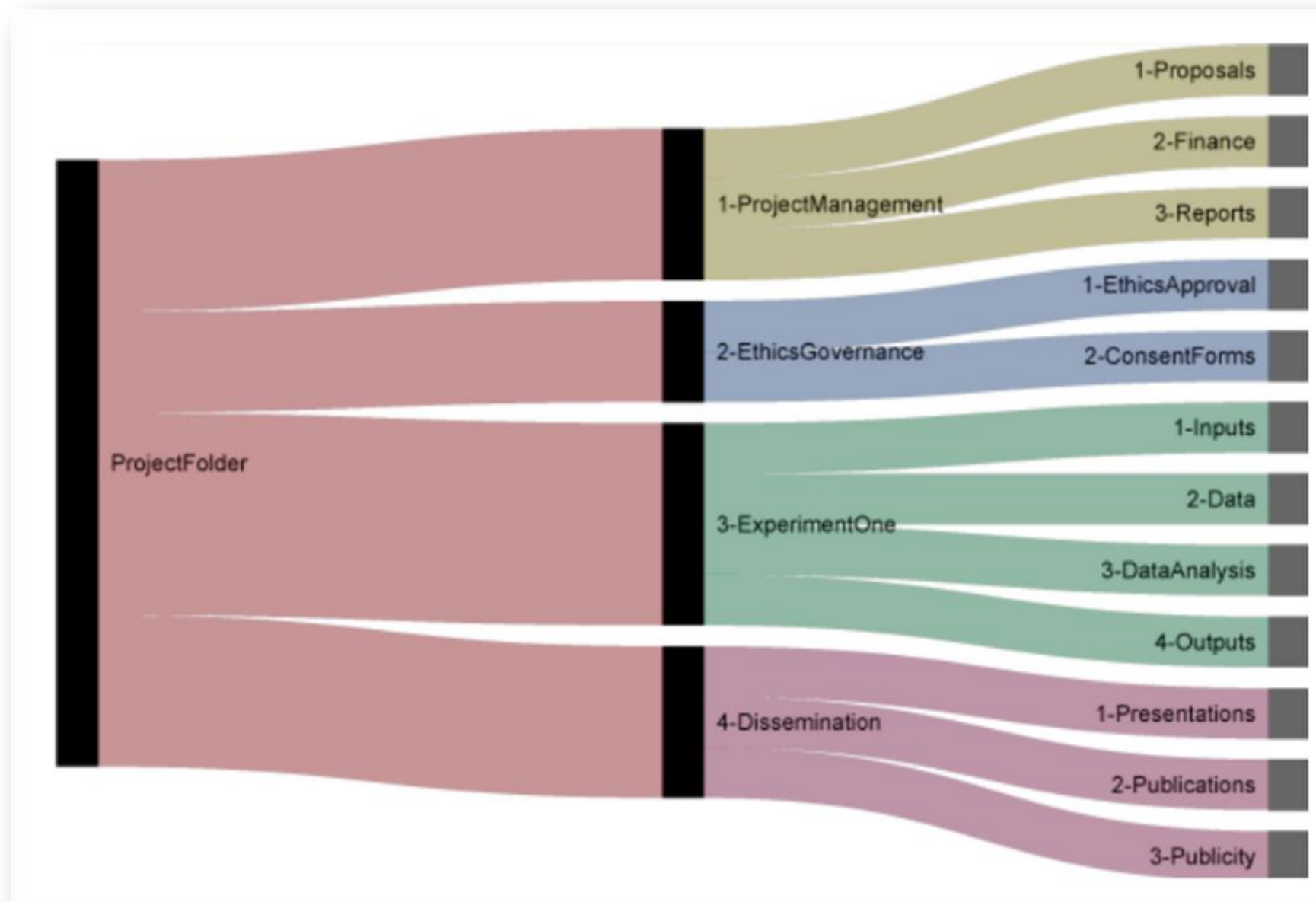
Codebook of variables

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

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

CURARE LA RACCOLTA DEI DATI: ORGANIZZAZIONE DELLE CARTELLE

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



http://nikola.me/folder_structure.html

È possibile strutturare le cartelle:

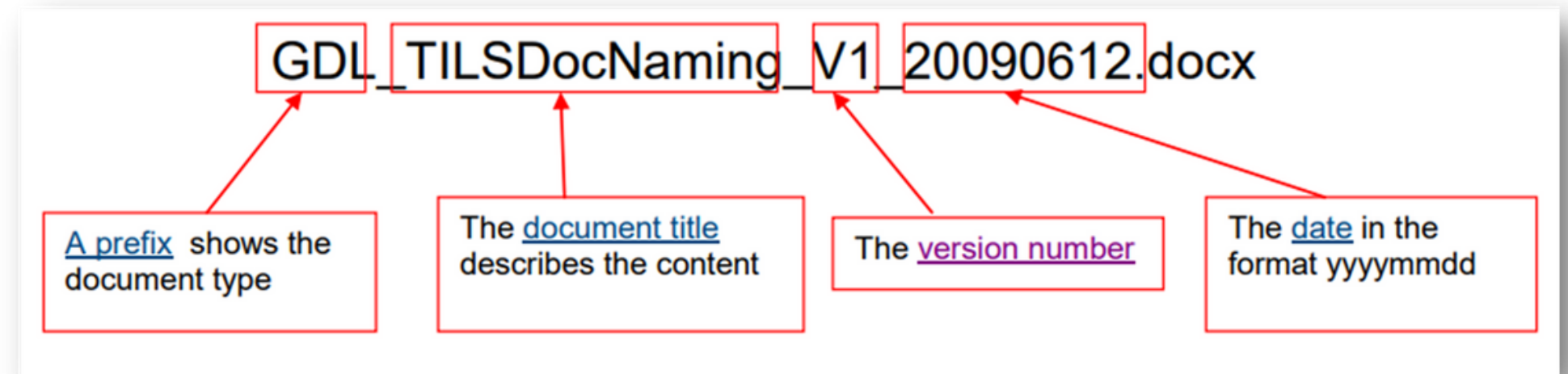
- in base alla persona che ha generato i dati/cartella;
- cronologicamente (mese, anno, sessioni);
- per progetto (come nell'esempio seguente);
- in base al metodo/attrezzatura di analisi;
- al tipo di dati.

CURARE LA RACCOLTA DEI DATI: NOMENCLATURA DI FILE E CARTELLE

E' essenziale per garantirne l'accessibilità e l'interoperabilità.

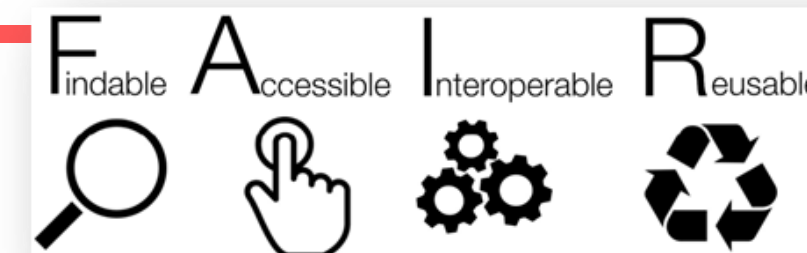
I nomi dei file devono essere leggibili sia dalla macchina che dall'uomo.

- Mantenere un nome breve ma descrittivo
- Evitare caratteri speciali o spazi
- Usare maiuscole o trattini bassi
- Utilizzare una formattazione coerente della data, ad esempio ISO 8601: AAAA-MM-GG
- Includere un numero di versione, se applicabile



https://www.data.cam.ac.uk/files/gdl_tilsdocnaming_v1_20090612.pdf

PERCHÈ LA DOCUMENTAZIONE È FONDAMENTALE?



I dataset devono essere accompagnati da metadati e parole chiave significative.



I metadati devono utilizzare un linguaggio standardizzato e condiviso a livello internazionale dai diversi servizi di indicizzazione.



I dati devono essere descritti e documentati nel miglior modo possibile, per garantirne la qualità e per poter essere replicati e combinati in contesti diversi.



UN ESEMPIO DI INTENZIONE DI DOCUMENTAZIONE NEL DMP

For each deposited data set, all relevant documentation explaining data collection procedures and analysis (such as codebooks, methodologies, etc.) are made available along with the data, in order to guarantee intelligibility, reproducibility and the validation of the project findings. Moreover, the deposited documentation specifies the tools and software recommended to reproduce and reuse the data, when necessary. (See Tab.3 for examples of tools and software enabling reuse of the dataset)

STORAGE E BACKUP DURANTE IL PROCESSO DI 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

E' 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?

- 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?

- 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 legal, ethical, or security reasons, 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 E BACKUP DURANTE IL PROCESSO DI RICERCA: QUALCHE SUGGERIMENTO

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



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 su Internet
- 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?



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 locale con backup automatici;
- Backup automatici di tutto il contenuto del laptop/PC (Windows: Cronologia file; Mac: Time Machine);
- Tenere conto delle diverse versioni dei file.



UN ESEMPIO DI SCELTA DI STORAGE NEL DMP

Data shared among Partners **did not** contain sensitive data because they **were** be anonymized, with the only exception of data relative to interviewees who specifically asked **not** to be anonymized, such as public stakeholders.

At each institution, research data **were** stored in computers, laptops, intranets or hard-drives accessible through institutional password periodically modified according to national law provisions for data security and protected by regularly updated antiviruses. None of the project data **were** left inadvertently available.

All the research materials stored in computers **were** subject to regular backup in order to safeguard them from accidental losses. For example, thanks to a partnership with Google® offering unlimited storage space to the University of Portsmouth staff, PBS team performed **a** secure backup of its data on Google Drive® on a daily base for the whole length of the project. All the data **were** password protected. If mobile devices **were** used to store data files (e.g. backup files), they **were** kept in a safe place accessible only to the researchers involved or **were** encrypted with *ad-hoc* software.

A cloud storage solution **was** adopted for data sharing among research teams. In this case, as well, regular backup of the data **was** performed to ensure data recovery. In addition all Partners **were** asked to keep local updated copies of all their files.

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)"
DOI 10.6092/unibo/amsacta/6244



QUALCHE PUNTO CHIAVE



- Per compilare correttamente il DMP segui un template e interpreta correttamente le domande che ti vengono poste.
- 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.





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.

Legend:

- DATA MANAGEMENT
- INTELLECTUAL PROPERTY RIGHTS
- PRIVACY
- ETHICS

DECISION TREE FOR DATA MANAGEMENT

Data management

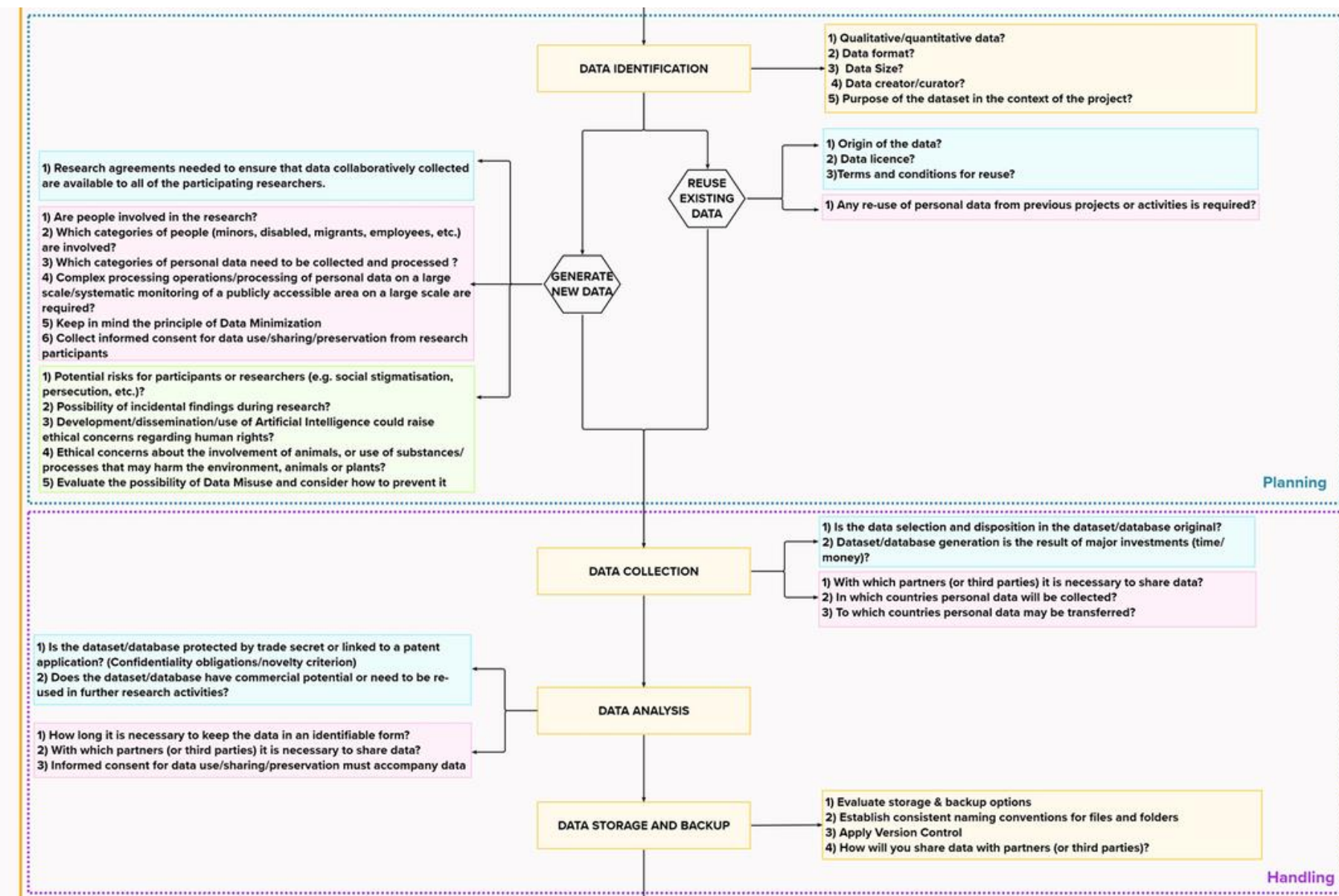
4 Legal and ethical requirements, codes of conduct	
4a If personal data are processed, how will compliance with legislation on personal data and on security be ensured?	<ul style="list-style-type: none">• Ensure that when dealing with personal data protection laws (for example GDPR) are complied with:<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.

RESPONSABILITÀ: COMPLIANCE ETICA

Riconoscere questioni e criticità di rilievo etico in relazione a obiettivi, metodologia e impatti del progetto di ricerca;

Prevedere (e dimostrare) il rispetto dei principi etici e delle normative a fronte delle criticità etiche rilevate.

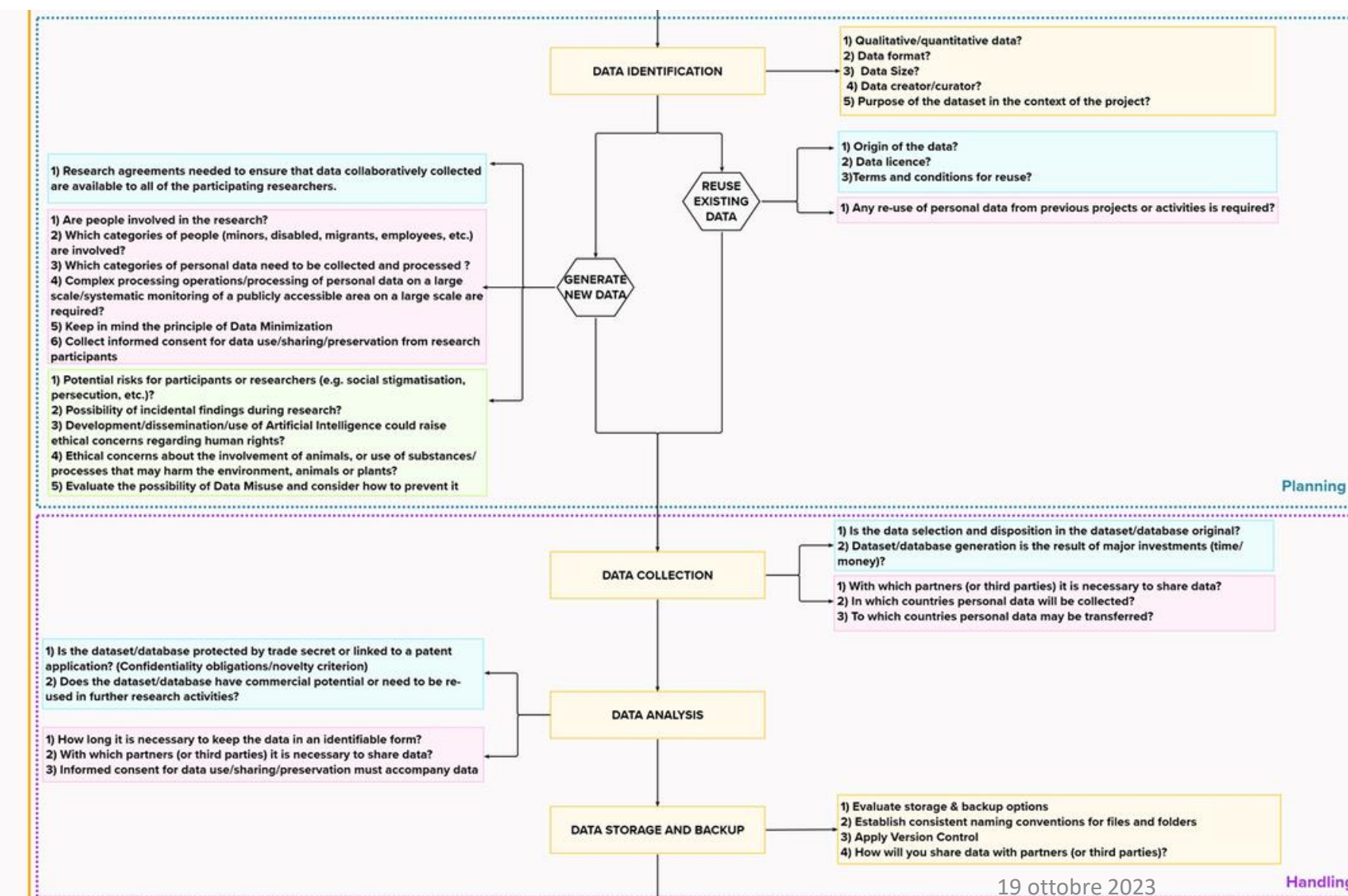
- RICERCA CON ESSERI UMANI (o con cellule staminali embrionali umane o embrioni umani o cellule o tessuti umani)
- PRIVACY
- POTENZIALI RISCHI per partecipanti o ricercatori
- INCIDENTALI FINDINGS (possibilità di scoperte inattese)
- INTELLIGENZA ARTIFICIALE (sviluppo, diffusione, uso)
- Coinvolgimento di ANIMALI
- Uso di sostanze o processi che possono DANNEGGIARE L'AMBIENTE.



RESPONSABILITÀ: TUTELA DEI DATI PERSONALI

La tutela dei dati personali è un tema etico e deontologico: riguarda il modo corretto di svolgere un'attività, la "research integrity".

- 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



RESPONSABILITÀ: TUTELA DEI DATI PERSONALI



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

TUTELA DEI DATI PERSONALI: QUALCHE SUGGERIMENTO

- Pianifica e ottieni un consenso informato all'inizio della ricerca
- Cifra i documenti che contengono i dati personali
- Consenti l'accesso a questi dati solo a persone autorizzate
- Pseudonimizzali subito per renderli più sicuri
- Appena possibile anonimizza i dati, operazione che renderà in un futuro depositabili i tuoi dati
- Raccogli dati personali solo se sono strettamente necessario agli scopi della ricerca

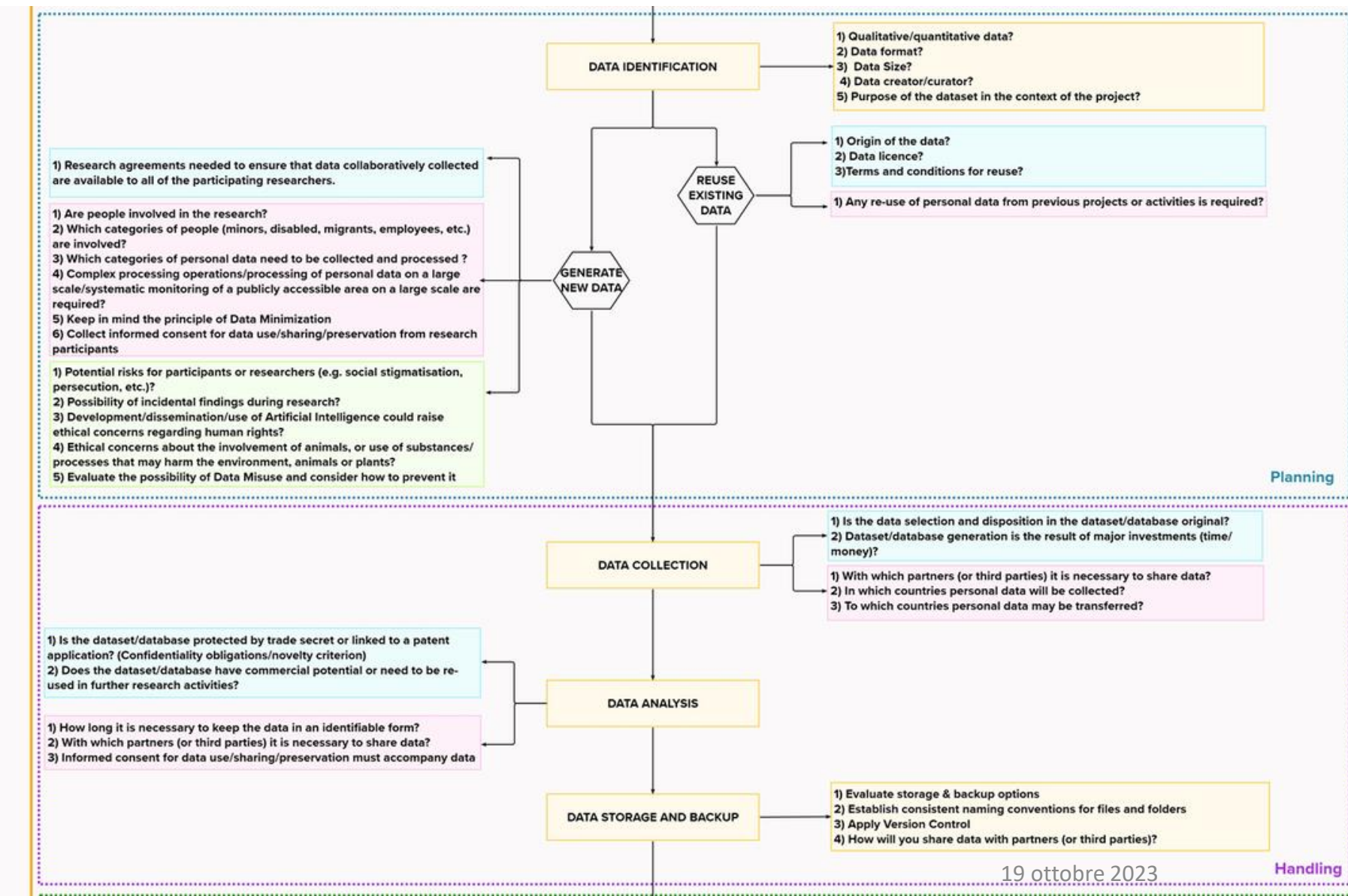
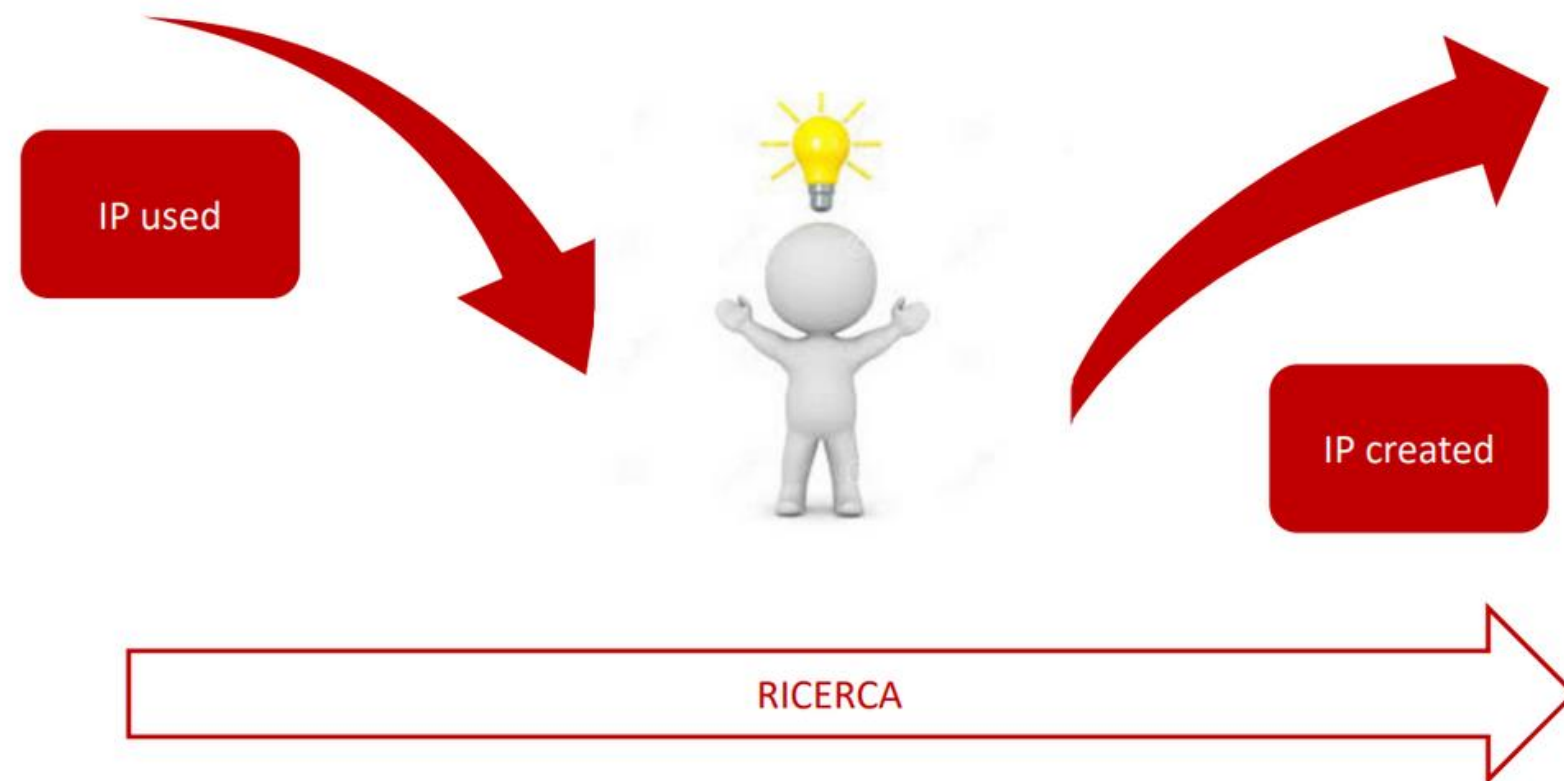


LE STRATEGIE SCELTE DEVONO ESSERE RIPORTATE NEL DMP!

RESPONSABILITÀ: TUTELA DELLA PROPRIETÀ INTELLETTUALE

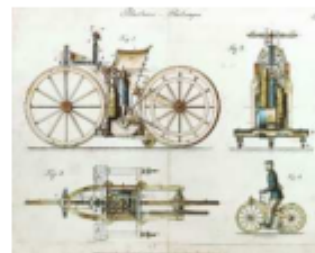
Insieme di diritti legali volti ad assicurare la tutela delle creazioni della mente umana in campo scientifico, industriale e artistico.

Quando rileva l'IP



RESPONSABILITÀ: TUTELA DELLA PROPRIETÀ INTELLETTUALE

BREVETTI: per un'invenzione sono concessi da un governo e 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

RESPONSABILITÀ: TUTELA DELLA PROPRIETÀ INTELLETTUALE

DIRITTO D'AUTORE: è una 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.



Opere
dell'ingegno

Copyright



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

TUTELA DELLA PROPRIETÀ INTELLETTUALE: QUALCHE SUGGERIMENTO

- Valuta fin dall'inizio se i dati che utilizzi possono essere utilizzati commercialmente.
- Valuta se possono servire per un brevetto.
- Controlla le condizioni di riutilizzo di dati già esistenti, per esempio la licenza .
- Studia quale è la repository migliore per depositare e preservare i tuoi dati anche in forma chiusa.
- Scegli una licenza opportuna per i tuoi dati.



LE STRATEGIE SCELTE DEVONO ESSERE RIPORTATE NEL DMP!



UN ESEMPIO DI GESTIONE DELLE QUESTIONI DI PRIVACY, IPR E ETICA NEL DMP

Research in PERCEIVE involved questionnaires, interviews and surveys with adults participants and focus groups with key informants (selected practitioners and experts of the EU Cohesion Policy). All aspects of collection of data involving personal data (in particular in WP1, WP2, WP3, WP4, WP6) were covered by the Ethics Requirements document (WP9).

All personal data collected within the PERCEIVE project from questionnaires, interviews, surveys and focus groups were carefully protected in compliance with relevant national data protection legislation of the EU member states, with the Regulation (EU) 2016/679 (GDPR)²⁶, and with the procedures defined by the European Code of Conduct for Research Integrity.

As a general principle, personal data resulting from the focus groups, interviews, observation and questionnaires were separated from the research results, and were handled by different members of the research team. In regards to the respondents in the survey, they were selected at random and their name and address were not recorded. The data were stored in a way not to allow the identification of the subject, adopting measures for pseudonymisation (i.e. names replaced by initials or pseudonyms); results of questionnaires and interviews were transmitted or made available to the other project partners as anonymous data.

In principle, research did not involve personal sensitive data. Notwithstanding, in certain circumstances, data collected during interviews and focus groups were potentially sensitive (e.g.

participants might disclose political opinions). Only personal characteristics that are strictly necessary, for theoretical reasons and to the benefit of the research, were collected; such characteristics were used to compare participants and to shape their views and opinions.

Files containing questionnaire data for statistical analysis, transcripts of interviews and focus groups, transcripts of field observations, photos, minutes, videos, action diaries, etc.) were stored in computers, laptops, intranets or hard-drives of the research institutions accessible through institutional password modified periodically (every 3 months in case of storage of sensitive data), and protected by regularly updated antiviruses. Files containing “sensitive” data were stored encrypted. Password-protected and encrypted files were accessible only to authorized members of the research teams receiving preliminarily specific information and training on the procedures for data collection, storage etc. None of the project data were left inadvertently available by being left on desks or in unlocked rooms. All the research materials stored in computers were subjected to back up regularly (according to each institutions’ regulations) in order to safeguard them from accidental losses.

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 di dati
- 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.

CONDIVISIONE E PRESERVAZIONE A LUNGO TERMINE: QUALI DATI?



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



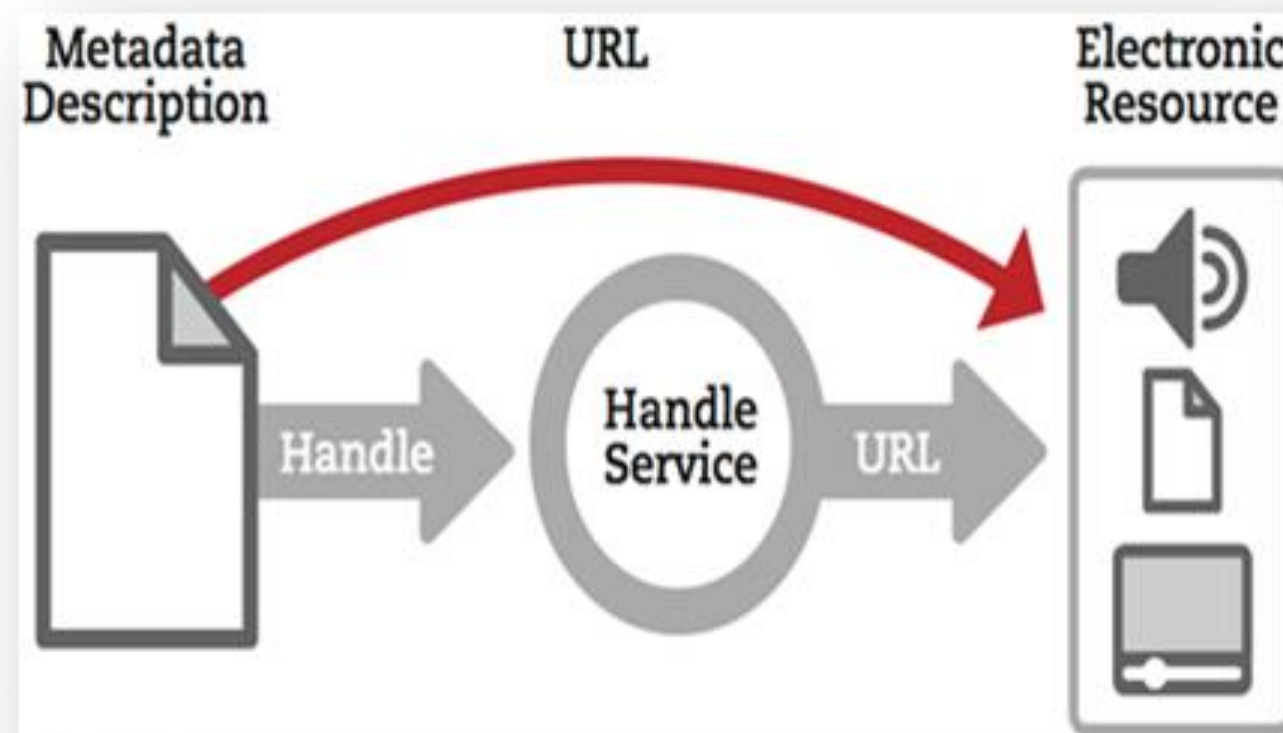
Tieni i dati che potrebbero esserti utili per progetti di ricerca futuri e quelli che puoi raccogliere in una sola occasione o che hai ottenuto dopo notevoli sforzi personali.



In alcuni casi i dati non possono essere tenuti alla fine della ricerca per accordi precedentemente presi con parti terze.

LE CARATTERISTICHE DI UN REPOSITORY: IDENTIFICATORI PERSISTENTI (PID)

- Un identificatore persistente (PID) è un riferimento duraturo a una risorsa.
- Lo scopo principale del PID è fornire le informazioni necessarie per identificare, verificare e localizzare la risorsa in modo affidabile.
- Un PID può essere collegato a un insieme di metadati che descrivono un elemento piuttosto che all'elemento stesso.



Publication date:
November 24, 2017

DOI:
DOI: [10.5281/zenodo.1065991](https://doi.org/10.5281/zenodo.1065991)

Keyword(s):
FAIR, FAIRness, checklist, research data, Findable, Accessible, Interoperable, Reusable, PID, repository, DOI, metadata, licence, data sharing, research data management,

Grants:
European Commission
• EUDAT2020 - EUDAT2020 (654065)

License (for files):
[CC BY](https://creativecommons.org/licenses/by/4.0/) Creative Commons Attribution 4.0

LE CARATTERISTICHE DI UN REPOSITORY: IDENTIFICATORI PRESTISTENTI (PID)

Un identificatore persistente (PID) è un **referimento duraturo a una risorsa**.

La risorsa può essere:

- una **persona** (ricercatori, autori, collaboratori: e.g. ORCID, ISNI)
- un'**organizzazione** (e.g. ROR)
- un **oggetto** (pubblicazioni, dati, software: e.g. DOI, Handle, ARK, URN)



Molti repository assegneranno un PID quando un oggetto viene depositato.

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
Intellectual Property Storage Preservation

License (for files):
[Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/)

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

Versions

Version 1 Oct 12, 2022
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.](#)

LE CARATTERISTICHE DI UN REPOSITORY: SCHEMI DI METADATI



METADATI: informazioni strutturate sui dati.

Si riferiscono a informazioni strutturate 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.

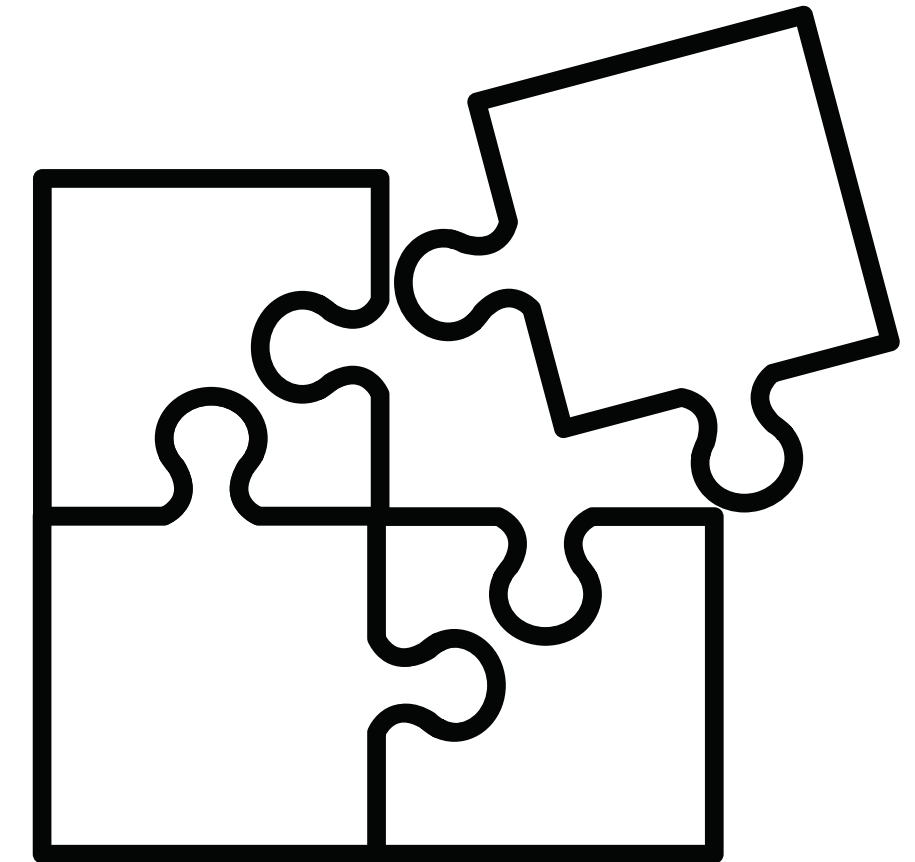
Per trovare standard di metadati disciplina-specifici, puoi utilizzare:

- 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>).

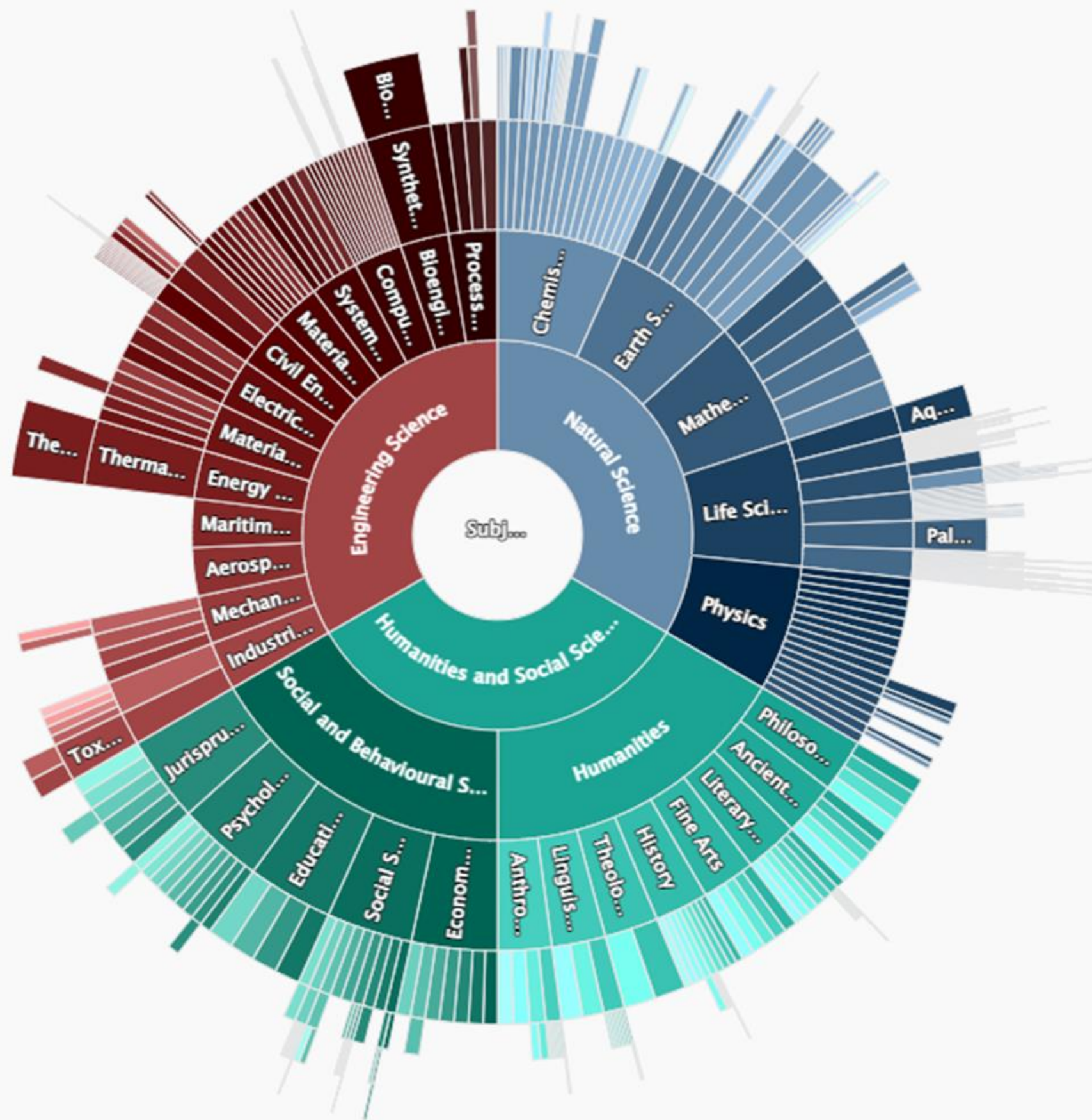
LE CARATTERISTICHE DI UN REPOSITORY: INTEROPERABILITÀ

Dati interoperabili possono essere **combinati** con altri insiemi di dati da parte di esseri umani o computer in maniera duratura nel tempo e con formati leggibili dalle macchine.

- I repository attendibili devono supportare gli **standard per l'interoperabilità**: e.g. OAI-PMH.
- I repository (in particolare quelli specifici del dominio) possono raccomandare l'uso di **ontologie o vocabolari specifici**.



LE CARATTERISTICHE DI UN REPOSITORY: INTEROPERABILITÀ



- Standard e vocabolari per dati e metadati sono **best practices** di interoperabilità approvate dalla comunità.
- I vocabolari e le ontologie servono a **descrivere concetti e relazioni** all'interno di un dominio di conoscenza.

Ci sono molti vocabolari e ontologie che si possono trovare sul web:

<https://fairsharing.org/browse/subject>



LE CARATTERISTICHE DI UN REPOSITORY: LICENZE

La licenza di copyright 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 le informazioni sulla licenza non sono incluse con i dati!
- Se i dati non includono una licenza, devi ottenere l'autorizzazione dal proprietario per copiare, condividere o pubblicare in altro modo questi dati.

Esistono strumenti online per determinare quale licenza è più appropriata per il tuo set di dati:











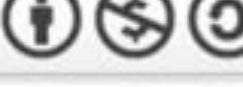
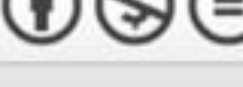
<https://creativecommons.org/choose/>


<https://ufal.github.io/public-license-selector/>


LE CARATTERISTICHE DI UN REPOSITORY: LICENZE


Rilascia le tue pubblicazioni e i tuoi dati con una licenza di copyright **chiara e permissiva** in modo da:


- Capire come controllare i tuoi materiali e puoi impostare le condizioni per il loro riutilizzo;
- Permettere agli altri di trovare, leggere e 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.

LE CARATTERISTICHE DI UN REPOSITORY: CONDIZIONI DI ACCESSO

Dati **personali**? Deve seguire il GDPR e proteggere la privacy

Dati **sensibili**? Deve seguire il GDPR, potrebbe voler coinvolgere comitati etici

Riutilizzare i dati sotto **copyright**?

Vuoi **brevettare**? Deve imporre un embargo

Vuoi pubblicare e temi che le idee possano essere rubate? Può imporre l'embargo e/o pubblicare i dati sottostanti insieme alla pubblicazione

Possibili strategie

- Anonimizzazione
- Accesso controllato
- Comitato di accesso

Controlla le licenze!

Il riutilizzo può essere limitato, come l'accessibilità di nuovi risultati

Possibili strategie

- Deposito dati con embargo
- Accesso controllato

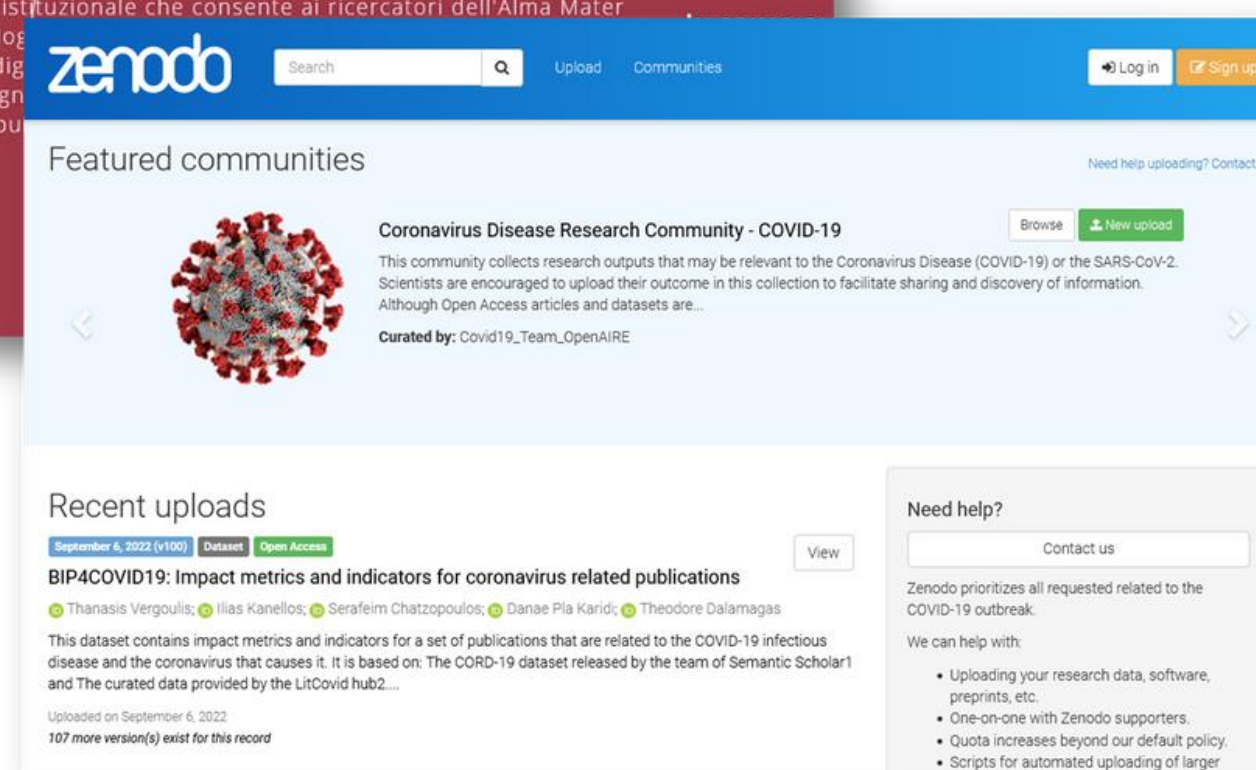
COME SCEGLIERE DOVE DEPOSITARE I DATI?



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	<p>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.</p>
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 almadl@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 EPrints
software

Versioning yes

Author identifier system(s) ORCID

Enhanced Publication yes

Quality management unknown

Application programming interfaces (1)

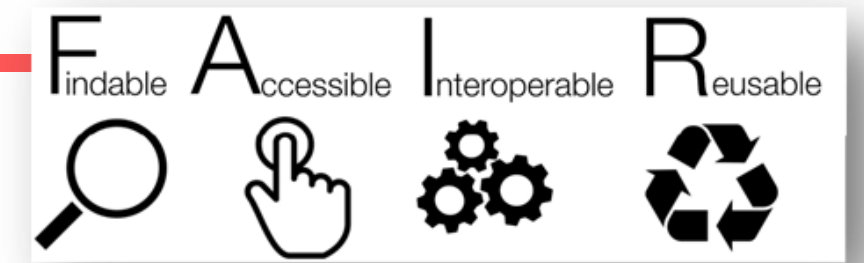
API type OAI-PMH


Metadata standards (1)

Metadata standard name Dublin Core

Metadata standard scheme DCC

IL RUOLO DEL REPOSITORY NEL GESTIRE I DATI IN MODO FAIR



F_{indable}  Ai dataset devono essere assegnati identificatori persistenti (ad esempio, DOI, Handle, URN).

A_{ccessible} 

- I dati e i metadati devono essere conservati in archivi o depositi che li rendano persistenti nel tempo e rintracciabili in rete.
- I dati possono essere apertamente accessibili (default) o accessibili attraverso un sistema di autenticazione e autorizzazione, se la natura dei dati ne impedisce l'apertura.

R_{eusable} 

Il riutilizzo dei (meta)dati deve essere dichiarato sotto una o più licenze aperte chiare e accessibili (ad esempio, Creative Commons, GPL per il software e altre).

CONDIVISIONE E PRESERVAZIONE A LUNGO TERMINE DEI DATI: QUALCHE SUGGERIMENTO

- Inserisci i dati necessari a verificare indipendentemente la tua ricerca.
- Se hai scartato determinati dati perchè non particolarmente rilevanti puoi dedicargli una sezione a parte in cui descrivi il processo di decisione.
- Crea una tabella nella quale indicare i repository che intendi utilizzare, quale schema di metadati supporta, che licenze consente di attribuire e quali PID fornisce.



UN ESEMPIO DI SCELTA DEL REPOSITORY NEL DMP

Data underlying articles stemming from the project and published after the project end will be deposited in the chosen repository at the latest at the time of the editorial acceptance of the manuscript to permit crosslinking between the publication and the data. Public access to these data will be given at the time of publication.

Table 2
Summary of repositories

Repository name	URL	Type	Partner
AMS Acta	https://amsacta.unibo.it/	Institutional	UNIBO
Quality of Government (QoG)	http://qog.pol.gu.se/data	Institutional	UGOT
Dipòsit Digital de la Universitat de Barcelona Dades	http://diposit.ub.edu/dspace/handle/2445/56364	Institutional	UB
Zenodo	https://zenodo.org/	Multi-disciplinary	IEA IAFE- NRI PBS WU UGOT UNIBO

The table shows the repositories chosen by the partners for both dissemination and long term preservation.

D'Alberto, Riccardo ; Raggi, Meri (2022) CONSOLE Project - Deliverable 7.6 - "Data Management Plan". University of Bologna, p. 32. DOI 10.6092/unibo/amsacta/7076 <<https://doi.org/10.6092/unibo%2Famsacta%2F7076>>.

Table 2
Summary of repositories

Repository name	URL	Type	Partner
AMS Acta	https://amsacta.unibo.it/	Institutional	UNIBO
Quality of Government (QoG)	http://qog.pol.gu.se/data	Institutional	UGOT
Dipòsit Digital de la Universitat de Barcelona Dades	http://diposit.ub.edu/dspace/handle/2445/56364	Institutional	UB
Zenodo	https://zenodo.org/	Multi-disciplinary	IEA IAFE- NRI PBS WU UGOT UNIBO

The table shows the repositories chosen by the partners for both dissemination and long term preservation.

Table 3
Features of the chosen repositories

Repository name	Permanent ID	OpenAIRE compatibility?	Indexed in re3data ¹³ catalogue?
AMS Acta	DOI	OpenAIRE 3.0 (OA, funding) OpenAIRE Data (funded, referenced datasets)	Yes, https://www.re3data.org/repository/r3d100012604
Quality of Government (QoG)	-	-	Yes, https://www.re3data.org/repository/r3d100012231
Dipòsit Digital de la Universitat de Barcelona Dades	Handle	OpenAIRE 3.0 (OA, funding)	Yes, https://www.re3data.org/repository/r3d100012304
Zenodo	DOI	OpenAIRE 3.0 (OA, funding) OpenAIRE Basic (DRIVER OA)	Yes, https://www.re3data.org/repository/r3d100010468

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)"
DOI 10.6092/unibo/amsacta/6244

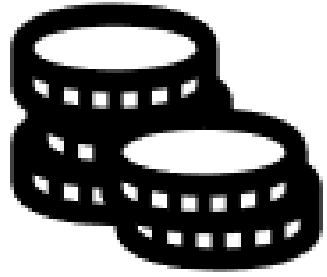
RESPONSABILITÀ E COSTI DELLA GESTIONE DEI DATI

La gestione dei dati della ricerca che può essere dispendiosa per il ricercatore, sia in termini di tempo che di denaro.

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.

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.

COSTI E RESPONSABILITÀ



- Raccolta dei dati: e.g. acquisizione dataset esterni, formattazione, organizzazione, trascrizione.
- Descrizione dei dati, i metadati e la documentazione: richiedono molto tempo, soprattutto se effettuati in una fase successiva del progetto.
- Archiviazione: e.g. raccogliendo dati di grandi dimensioni.
- Accesso ai dati e la sicurezza: e.g. accesso remoto tramite VPN.
- Conservazione: e.g. convertire i dati/file in formati aperti.
- Riutilizzo dei dati: e.g. anonimizzazione, diritti d'autore, condivisione dei dati.

COSTI DELLA GESTIONE DATI: QUALCHE SUGGERIMENTO

The cost of data management

How to calculate costs?

How to use this costing tool?

Estimating costs RDM tool

What will it cost to manage and share my data?

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 .	2 hrs to 2 days, depending on the complexity of your project
		Check if there is a department within your organization to support data management planning.	
1. Data Collection	Acquiring External datasets	Your library may be able to help you acquire a license to a crucial database	<u>Example:</u> A faculty licence on a database for macro-economic analysis: €18.000/y
		Do you plan to use existing data, and is the data available at a commercial partner? In research data repositories, data can be available at no or low costs	

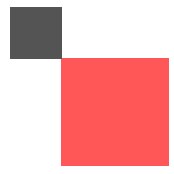
How to use this costing tool?

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.
- La gestione dei dati della ricerca che può essere dispendiosa per il ricercatore, sia in termini di tempo che di denaro.







GRAZIE

