

FAIR principles, Open Science & Data Stewardship



Giulia Pedonese
14 aprile 2025

Fonti

Questa presentazione è il risultato dell'adattamento delle seguenti fonti:

- van der Lek, I., Fišer, D., Frontini, F., & Pedonese, G. (2024, agosto 30). Introduzione ai Dati Linguistici: Standard e Archivi Digitali. Zenodo. <https://doi.org/10.5281/zenodo.13911935>
- van der Lek, I., Fišer, D., Samardzic, T., Simonovic, M., Assimakopoulos, S., Bernardini, S., Milicevic Petrovic, M., & Puskas, G. (2023). Integrating research infrastructures into teaching: Recommendations and best practices (Versione 2) <https://doi.org/10.5281/zenodo.8114407>. CC BY 4.0.
- CLARIN ERIC Official Website: <https://www.clarin.eu/> . CC BY 2.0
- Pasquale, V., & PASCHETTA, M. (2024, novembre 25). Un anno di Comunità Italiana Data Steward. GenOA Week 2024, Genova, Italy. Zenodo. <https://doi.org/10.5281/zenodo.14215870>
- Sito web Open-Science.it <https://open-science.it/>
- Pedonese, G. (2025, gennaio 17). Progettazione Didattica FAIR: la Metodologia di Skills4EOSC Applicata a CLARIN -IT. Zenodo. <https://doi.org/10.5281/zenodo.14698571>

Indice

- Perché applicare le buone pratiche della **Scienza Aperta**?
- A cosa servono i **principi FAIR**?
- Cos'è e come funziona un **repository**
- Come applicare una **licenza** d'uso
- **Citare correttamente** una risorsa: lo schema TASL
- Cos'è un **Data Management Plan** e perché è utile per la ricerca

Scienza Aperta: un nuovo paradigma

La Scienza Aperta (Open Science) rappresenta **un nuovo approccio al processo scientifico basato sul lavoro cooperativo e su nuove modalità di diffusione** della conoscenza attraverso l'utilizzo di tecnologie digitali e di nuovi strumenti di collaborazione ([European Commission, 2016](#))"

La scienza aperta comprende l'accesso libero agli articoli scientifici, l'accesso ai dati della ricerca pubblica e la ricerca collaborativa abilitata da strumenti e incentivi TIC ([OCSE, 2015](#))

<https://open-science.it/cos-e-l-open-science>

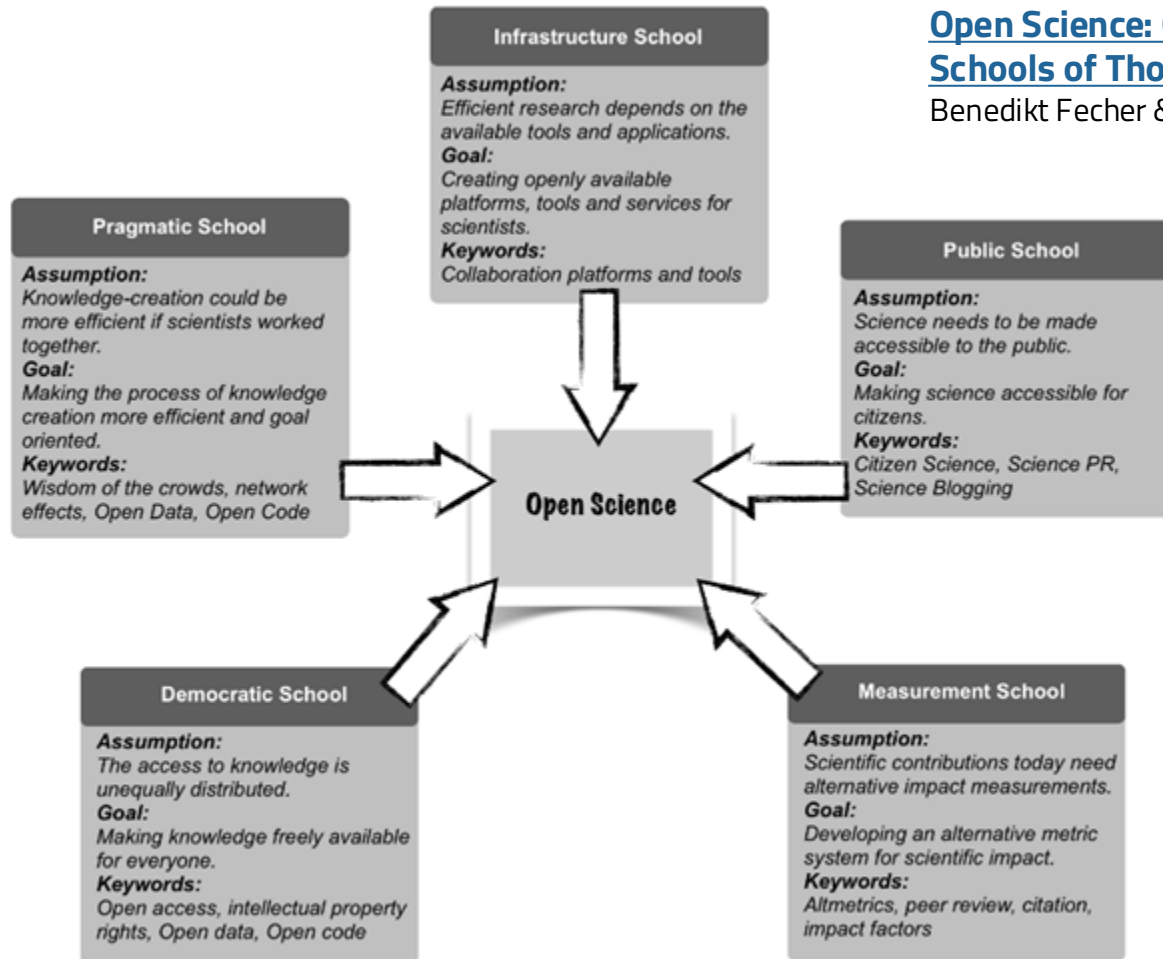
Research Data Lifecycle



1. Pianificazione
2. Raccolta
3. Elaborazione
4. Analisi
5. Archiviazione
6. Condivisione
7. Riutilizzo

Open Science: One Term, Five Schools of Thought

Benedikt Fecher & Sascha Friesike



Principi Guida FAIR

I principi FAIR per la gestione e la conservazione dei dati scientifici sono stati pubblicati nel 2016 con l'obiettivo principale di fornire linee guida per migliorare la gestione degli oggetti digitali nel Semantic Web. FAIR è l'acronimo inglese di:

- **Findability** - Trovabilità
- **Accessibility** - Accessibilità
- **Interoperability** - Interoperabilità
- **Riusability** - Riusabilità

Gestione dei dati FAIR

'FAIR Guiding Principles for scientific data management and stewardship', 2016

<https://www.go-fair.org/fair-principles/>

FAIR Principles

Compliance



Findability

Resource and its metadata are easy to find by both, humans and computer systems. Basic machine readable descriptive metadata allows the discovery of interesting data sets and services.

- ✓ F1. Resource is uploaded to a public repository.
- ✓ F2. Metadata are assigned a globally unique and persistent identifier.



Accessibility

Resource and metadata are stored for the long term such that they can be easily accessed and downloaded or locally used by humans and ideally also machines using standard communication protocols.

- ✓ A1. Resource is accessible for download or manipulation by humans and is ideally also machine readable.
- ✓ A2. Publications and data repositories have contingency plans to assure that metadata remain accessible, even when the resource or the repository are no longer available.



Interoperability

Metadata should be ready to be exchanged, interpreted and combined in a (semi)automated way with other data sets by humans as well as computer systems.

- ✓ I1. Resource is uploaded to a repository that is interoperable with other platforms.
- ✓ I2. Repository meta- data schema maps to or implements the CG Core metadata schema.
- ✓ I3. Metadata use standard vocabularies and/or ontologies.



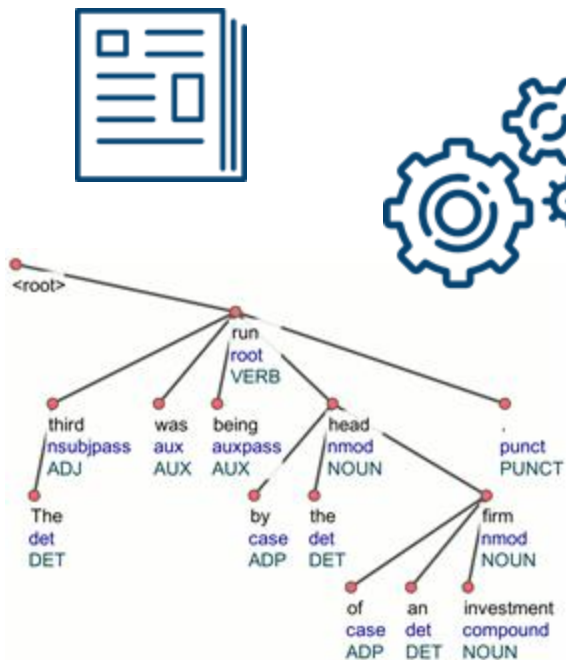
Reusability

Data and metadata are sufficiently well-described to allow data to be reused in future research, allowing for integration with other compatible data sources. Proper citation must be facilitated, and the conditions under which the data can be used should be clear to machines and humans.

- ✓ R1. Metadata are released with a clear and accessible usage license.
- ✓ R2. Metadata about data and datasets are richly described with a plurality of accurate and relevant attributes.

Un esempio di gestione dei dati:

le risorse linguistiche in CLARIN



CLARIN

Common Language Resources and
Technology Infrastructure

Cosa sono le Risorse Linguistiche?

Per Risorse Linguistiche si intende **l'insieme di dati e metadati relativi al linguaggio in forma leggibile dalla macchina**. Esse vengono utilizzate per costruire, migliorare o valutare algoritmi o sistemi di linguaggio naturale (linguaggio umano) e vocale e, sempre più spesso, per l'apprendimento automatico. Sono inoltre utilizzati come risorse fondamentali ad esempio per gli studi linguistici e di traduzione, per l'editoria elettronica, per le transazioni internazionali, per gli specialisti di aree tematiche e per gli utenti finali

Fonte: [ELRA/ELDA](#)



Dove trovare le Risorse Linguistiche? Gli Archivi Digitali

“Definiamo un repository - o deposito o archivio digitale - come un servizio gestito da organizzazioni di ricerca, in cui i materiali di ricerca sono conservati, gestiti e resi accessibili”. Fonte: [DataCite](#)

Per visualizzare i diversi tipi di risorse linguistiche e comprendere il modo in cui vengono descritte (il modello dei metadati), si può consultare la pagina [ELRC-SHARE repository](#)



<https://ilc4clarin.ilc.cnr.it/>

Esempio di un Corpus in ILC4CLARIN

The screenshot shows the ILC4CLARIN Repository interface for the item 'Travel Writings on Italy'. The page includes a header with the repository name and a 'View item' link. Below the title, there is a yellow banner with instructions on how to cite the item and a 'CITE' button. A red circle with the number 1 points to this banner. The main content area lists various metadata fields: Authors (Sprugnoli, Rachelle), Item identifier (http://hdl.handle.net/20.500.11752/OPEN.976), Project URL (https://files.google.com/view/travelwritingsonitaly/home), Demo URL (https://drive.google.com/file/d/1B-TXND7JAARVYVRYWYTLONZ7N6UCU/view?usp=sharing), Referenced by (https://doi.org/10.26434/chemrxiv-2018-07-11-111111), Date issued (2018), Type (corpus, text), Size (50 files), Language(s) (English), Description (A collection of travel writings - non-fictional narratives (reports, diaries, letters) and guidebooks - about Italy written by English native authors and published between the country unification and the beginning of the 20th century. Texts are available in both TEI-XML and TEI-XML format. Texts have been collected from Project Gutenberg. TEI was automatically generated by GutenTag v0.1.4 and then manually checked and integrated.), Publisher (Fondazione Bruno Kessler - Trento), Subject(s) (Travel writings, Cultural texts), and Collection(s) (ILC4CLARIN - OPEN Data & Tools). A red circle with the number 2 points to the Authors field. A red circle with the number 3 points to the Referenced by field. A red circle with the number 4 points to the Description field. A red circle with the number 5 points to the Publisher field. A red circle with the number 6 points to the 'Files in this item' section, which shows a download button for the entire corpus (10.57 MB) and a button for downloading individual files. The page also includes social media sharing icons and a 'Show full item record' link.

1. Informazioni sulla citazione tramite handle
2. Campi di metadati che descrivono il corpus
3. Il corpus è citato in una rivista
4. Il corpus è descritto, ad esempio i testi sono disponibili in formato UTF-8 e TEI-XML
5. Informazioni sull'editore
6. Istruzioni per il download

CLARIN Resource Families

Corpora

- Computer-Mediated Communication Corpora
- Corpora of Academic Texts
- Historical Corpora
- L2 Learner Corpora
- Legal Corpora
- Literary Corpora
- Manually Annotated Corpora
- Multimodal Corpora
- Newspaper Corpora
- Oral History Corpora
- Parallel Corpora
- Parliamentary Corpora
- Reference Corpora
- Sign Language Resources
- Spoken Corpora

Lexical Resources

- Language Models
- Lexica
- Dictionaries
- Conceptual Resources
- Glossaries
- Wordlists

Tools

- Corpus Query Tools
- Normalisation
- Named Entity Recognition
- Part-of-Speech Tagging and Lemmatisation
- Tools for Sentiment Analysis

Principi FAIR in CLARIN

- **FINDABLE** attraverso punti di accesso come il Virtual Language Observatory e la Federated Content Search
- **ACCESSIBLE** attraverso i centri CLARIN che forniscono repository, metadati e un sistema di login federato
- **INTEROPERABLE** grazie al Language Switchboard, all'uso di vocabolari ([Concept Registry](#)) e a modi chiari e ben documentati di collegare i metadati
- **REUSABLE** grazie a licenze, formati e standard, alla curatela attiva dei metadati e ai controlli sui collegamenti

<https://www.clarin.eu/fair>

I Metadati

Lo scopo dei metadati per una risorsa è consentire la catalogazione e la scoperta attraverso mezzi standardizzati per riportare:

- **Chi** ha creato la risorsa
- **Cosa** è il contenuto della risorsa
- **Quando** è stata creata la risorsa
- **Dove** si trova la risorsa
- **Perché** è stata creata la risorsa

I metadati **devono essere in formati standard** (cioè compatibili e condivisi con altre risorse) e leggibili anche dalle macchine per essere indicizzabili, cercabili, confrontabili e conservabili

Identificatori Persistenti (PID)

Un identificatore persistente (PID) è un tipo di metadato. Il suo scopo è quello di **etichettare in modo univoco un oggetto digitale e garantire che questa etichetta non cambi nel tempo**

Esistono diversi tipi di PID, come DOI (digital object identifier), Handle (HNDL), ARK (Archival Resources Key)

ESEMPLI:

- Questo è un DOI: [10.5281/zenodo.8305539](https://doi.org/10.5281/zenodo.8305539)
- Questo è un Handle: hdl.handle.net/20.500.11752/ILC-567

Identificatori Persistenti (PID)

Utilizzando un PID si può **arrivare a una pagina accessibile con l'oggetto digitale elencato e i suoi metadati**. L'accesso effettivo all'oggetto digitale da questa pagina può essere limitato. I PID sono **leggibili dalla macchina** e aiutano a distinguere tra:

- Materiali diversi
- Versioni diverse dello stesso materiale, per esempio:
<https://zenodo.org/records/13816183>

The screenshot shows the Zenodo upload interface. At the top is a blue header with the Zenodo logo, a search bar, and links for 'Communities' and 'My dashboard'. Below the header is a grey bar with the text 'Select the community where you want to submit your record.' and a 'Select a community' button. The main content area has a blue bar labeled 'Files' with a right arrow. Below this is a storage status bar showing 'Storage available: 0 out of 100 files, 0 bytes out of 50.00 GB'. The central area is a large grey box with the text 'Drag and drop files' and '- or -' followed by an 'Upload files' button. Below this is a yellow warning box stating: 'File addition, removal or modification are not allowed after you have published your upload.' The 'Basic information' section is a blue bar with a right arrow. It contains a 'Digital Object Identifier' section with the question 'Do you already have a DOI for this upload?' and two radio buttons: 'Yes' and 'No'. The 'Yes' option is selected. Below this is a text input field containing '10.5281/zenodo.14671900' and a small 'X' icon. A note below the input field states: 'Reserve a DOI by pressing the button (so it can be included in files prior to upload). The DOI is registered when your upload is published.' The 'Resource type' section has a dropdown menu with a right arrow. The 'Title' section has a text input field.

Diritti di Proprietà Intellettuale

La proprietà intellettuale si riferisce alla creazione di attività intellettuali. I **diritti di proprietà intellettuale** (DPI) proteggono gli interessi dei creatori e dei proprietari fornendo loro diritti sulla loro creazione.

Ai fini della creazione di dati della ricerca, vengono utilizzati il diritto d'autore e i diritti connessi, rami principali dei DPI. Essi sono **definiti per proteggere le creazioni letterarie e artistiche, le esecuzioni e i fonogrammi, definendo i diritti degli autori, dei proprietari, degli esecutori, dei produttori e delle emittenti.**

Nota: Qualsiasi opera creata dall'utente è di default il detentore del diritto d'autore per l'utente stesso o per la sua istituzione/datore di lavoro.

Copyright

Solo il titolare del copyright è autorizzato a distribuire ciò che viene creato. Se si desidera trasferire questo diritto anche ad altre persone, è possibile farlo tramite una licenza.

In assenza di eccezioni o limitazioni, **è possibile riutilizzare un'opera esistente se questa è concessa in licenza all'utente o se è concessa al pubblico tramite una licenza pubblica come le licenze Creative Commons (CC) o le licenze Free-Libre / Open Source Software (FLOSS).**







Come Applicare una Licenza CC?

Il contenuto di queste slide è riadattato dalle seguenti fonti:







- CC FAQ website distribuito con licenza CC BY 4.0
- Creative Commons license Wikipage (16 gen 2025) su Wikipedia distribuito con licenza Creative Commons Attribution-ShareAlike 4.0
- CC License Compatibility Chart

Come Applicare una Licenza CC?

Una licenza permette di **definire i diritti e gli obblighi relativi all'uso del vostro lavoro**. Una delle migliori fonti di informazione per la concessione di licenze per risorse digitali riutilizzabili è Creative Commons. Ci sono molte opzioni di licenza CC disponibili che si basano sui quattro diritti fondamentali:

Icon	Right	Description
	Attribution (BY)	Licensees may copy, distribute, display, perform and make derivative works and remixes based on it only if they give the author or licensor the credits (attribution) in the manner specified by these. Since version 2.0, all Creative Commons licenses require attribution to the creator and include the BY element.
	Share-alike (SA)	Licensees may distribute derivative works only under a license identical to ("not more restrictive than") the license that governs the original work. (See also copyleft .) Without share-alike, derivative works might be sublicensed with compatible but more restrictive license clauses, e.g. CC BY to CC BY-NC.)
	Non-commercial (NC)	Licensees may copy, distribute, display, perform the work and make derivative works and remixes based on it only for non-commercial purposes.
	No derivative works (ND)	Licensees may copy, distribute, display and perform only verbatim copies of the work, not derivative works and remixes based on it. Since version 4.0, derivative works are allowed but must not be shared.

Come Applicare una Licenza CC?

License name ⇅	Abbreviation ⇅	Icon ⇅	Attribution required ⇅	Allows remix culture ⇅	Allows commercial use ⇅	Allows Free Cultural Works ⇅
Attribution	BY		Yes	Yes	Yes	Yes
Attribution-ShareAlike	BY-SA		Yes	Yes	Yes	Yes
Attribution-NonCommercial	BY-NC		Yes	Yes	No	No
Attribution-NonCommercial-ShareAlike	BY-NC-SA		Yes	Yes	No	No
Attribution-NoDerivatives	BY-ND		Yes	No	Yes	No
Attribution-NonCommercial-NoDerivatives	BY-NC-ND		Yes	No	No	No

[Creative Commons license Wikipage](#) (2025, Jan 14) su Wikipedia con licenza [Creative Commons Attribution-ShareAlike License 4.0](#) / immagine tagliata

Scegliere una Licenza

Lo strumento CC License Chooser **aiuta gli autori a condividere il proprio lavoro in modo standardizzato**, fornendo licenze di copyright che consentono la condivisione e il riutilizzo dell'opera creativa alle condizioni scelte.

Per applicare una licenza **è sufficiente indicare quale licenza CC si intende applicare alla propria opera. È quindi necessario verificare i termini** della licenza scelta e **rispettare i requisiti** in essa contenuti.

Scegliere una Licenza

Si raccomanda di includere un link all'atto di licenza CC pertinente (ad esempio, <https://creativecommons.org/licenses/by/4.0>) oppure il full text della licenza (nel file README).

Esempio «Questo lavoro ha licenza secondo [Creative Commons Attribution 4.0 International License](#).»

Compatibilità di Licenze

Se la mia risorsa è una **combinazione di materiale preesistente** tratto da diverse fonti e riutilizzato così com'è, basta citare correttamente ciascuna delle parti dichiarandone la provenienza e la rispettiva licenza.

Se si effettuano adattamenti di materiale precedentemente rilasciato con licenza CC, **la licenza CC originale si applica sempre al materiale che si sta adattando, anche una volta adattato.** La licenza che si può scegliere per la propria rielaborazione (chiamata "licenza dell'adattatore") dipende da quale licenza si applica al materiale originale.

- **NOTA:** Una modifica raggiunge il livello di un adattamento ai sensi della legge sul diritto d'autore quando l'opera modificata si basa sull'opera precedente ma manifesta una creatività nuova sufficiente per essere tutelata dal diritto d'autore.

Adapter's license chart		Adapter's license						
		BY	BY-NC	BY-NC-ND	BY-NC-SA	BY-ND	BY-SA	PD
Status of original work	PD							
	BY							
	BY-NC							
	BY-NC-ND							
	BY-NC-SA							
	BY-ND							
	BY-SA							

Compatibilità di Licenze

Nota: **CC sconsiglia di utilizzare una licenza se il riquadro corrispondente è giallo**, anche se ciò è tecnicamente consentito dai termini della licenza. Se lo fate, dovrete fare attenzione a segnalare che l'adattamento coinvolge più diritti d'autore con termini diversi, in modo che gli utenti a valle siano consapevoli dei loro obblighi di rispettare le licenze di tutti i titolari dei diritti.

Le caselle grigio scuro indicano le licenze che non possono essere utilizzate come licenza dell'adattatore.

Attribuzione

Il diritto di attribuzione è un diritto morale degli autori che protegge il rapporto personale tra l'autore e l'opera creata, anche se il creatore non possiede il copyright. **Il riconoscimento dei materiali riutilizzati attraverso l'attribuzione è sempre necessario** (anche se non è un requisito della licenza, come nel caso delle licenze pubblico dominio). L'attribuzione ideale segue l'approccio TASL

Titolo - qual è il nome dell'opera

Autore - chi permette di utilizzare l'opera (nome e link)

Fonte (Source)- dove si può trovare l'opera (link aggiunto al titolo)

Licenza - come si può usare l'opera (nome e link alla licenza)

Come Attribuire Correttamente?

Se il vostro lavoro è una modifica o un adattamento di un'altra opera, indicatelo e **fornite l'attribuzione al creatore dell'opera originale. Dovreste anche includere un link all'opera che avete modificato e indicare quale licenza si applica a quell'opera.**

Per le opere create da altri e che si stanno incorporando nel proprio lavoro, si deve rendere facile per gli altri sapere chi ha creato quali parti dell'opera, utilizzando l'approccio TASL.

Vd. anche [Recommended practices for attribution](#)

Come Attribuire Correttamente?

Se si riutilizza una risorsa, si raccomanda di fornire **l'attribuzione all'inizio della presentazione o del testo**. Lo stesso vale per le parti riutilizzate di una risorsa (solo testo o combinazione di testo e immagini).

L'attribuzione delle **immagini e delle tabelle riutilizzate** va posta **sotto il contenuto**.

<https://h2iosc-training-platform.ilc4clarin.ilc.cnr.it/>

Attribuzione

Questo corso traduce in italiano e aggiorna i materiali di:

van der Lek, Iulianna; Fišer, Darja. (2023). *Introduction to Language Data: Standards and Repositories*. In [UPSKILLS Learning Content](https://upskillsproject.eu/project/standards_repositories/): https://upskillsproject.eu/project/standards_repositories/, CC BY 4.0.

L'adattamento si è svolto nell'ambito del progetto [Humanities and cultural Heritage Italian Open Science Cloud](#) Work Package 8 "Training, Capacity Building, Engagement", a cura del personale CNR-ILC dedicato all'Attività 8.2 "Teach CLARIN, Teach with CLARIN".

Progetto H2IOSC - Humanities and cultural Heritage Italian Open Science Cloud finanziato dall'Unione Europea NextGenerationEU - PNRR M4C2 - Codice progetto IR0000029 - CUP B63C22000730005.



Immagine tagliata da: van der Lek, Iulianna, Darja Fišer, Francesca Frontini, e Giulia Pedonese. «Introduzione Ai Dati Linguistici: Standard E Archivi Digitali». 31 agosto 2024. <https://doi.org/10.5281/zenodo.13911935>, con licenza [CC BY 4.0](#).

Come Attribuire Correttamente?

Per riassumere:

- Fare attenzione alle modalità indicate dagli autori e dare loro la precedenza
- Controllare nel file README o la presenza di un file .cff
- Eventualmente adattare lo stile di citazione alla bibliografia che si sta componendo
- In mancanza di indicazioni, seguire lo schema TASL

Come Attribuire Correttamente?

Per Far sì che gli altri possano citare il nostro lavoro correttamente, consigliamo di stabilire una citazione standard e di comunicarla esplicitamente (in fondo alla presentazione, nell'Introduzione, nella Guida al Riuso, nel file .cff per il deposito ecc.).

Modalità di citazione

Esempio:

È possibile adattare e condividere il contenuto sotto la licenza [CC-BY 4.0](#) a condizione che:

- si dia il giusto credito agli autori utilizzando la **modalità di citazione standard**:

van der Lek, Iulianna, Darja Fišer, Francesca Frontini, e Giulia Pedonese. «Introduzione Ai Dati Linguistici: Standard E Archivi Digitali». 31 agosto 2024. <https://doi.org/10.5281/zenodo.13911935>.

- si includa il link alla licenza.
- si indichi se sono state apportate modifiche al contenuto.

Immagine tagliata da: van der Lek, Iulianna, Darja Fišer, Francesca Frontini, e Giulia Pedonese. «Introduzione Ai Dati Linguistici: Standard E Archivi Digitali». 31 agosto 2024. <https://doi.org/10.5281/zenodo.13911935>, con licenza [CC BY 4.0](#).

Cos'è un Data Management Plan?

Un Piano di Gestione dei Dati

Un **Data Management Plan (DMP)** è un documento formale che descrive come verranno gestiti i dati durante e dopo un progetto di ricerca.

Il suo scopo principale è garantire che i dati siano ben organizzati, accessibili, utilizzabili e conservati nel tempo, in linea con i principi FAIR (Findable, Accessible, Interoperable, Reusable) e con le buone pratiche della scienza aperta.

Elementi Chiave

Un DMP solitamente include informazioni su:

- **Tipologia di dati**

- Quali dati verranno raccolti, generati o riutilizzati?
- In quale formato saranno archiviati?

- **Standard e metadati**

- Quali standard verranno seguiti per garantire interoperabilità e riutilizzo?
- Come verranno documentati i dati?

Elementi Chiave

- **Conservazione e sicurezza**

- Dove e come verranno archiviati i dati durante il progetto?
- Quali misure di sicurezza verranno adottate per proteggerli?

- **Condivisione e accesso**

- Quali dati saranno resi pubblicamente accessibili?
- Ci saranno restrizioni o requisiti di licenza specifici?

Elementi Chiave

- **Responsabilità e gestione a lungo termine**
 - Chi sarà responsabile della gestione dei dati?
 - Dove e per quanto tempo verranno conservati dopo la fine del progetto?

Perché è importante avere un DMP?

- Aiuta i ricercatori a **organizzare e gestire** i dati in modo efficiente.
- Facilita la **condivisione e il riutilizzo** da parte di altri studiosi.
- È spesso un **requisito per ottenere finanziamenti** da enti pubblici e privati.
- **Riduce il rischio di perdita** o corruzione dei dati nel tempo.

Alcuni enti di finanziamento, come la **Commissione Europea** e la **National Science Foundation (NSF)**, richiedono un DMP come parte integrante delle proposte di progetto.

Plan and follow your data

Create machine actionable DMPs.

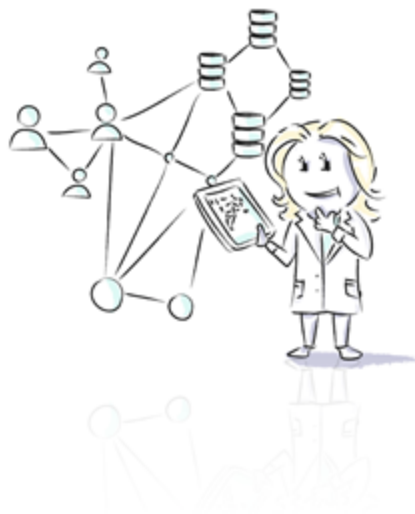
Configure to best fit your discipline.

Link to EOSC components out of the box.

Share easily in your repository.

Bring your Data Management Plans closer to where data are generated, analysed and stored.

Start your DMP



<https://argos.openaire.eu/splash/>

Cos'è un Data Steward?

«Il data steward è un professionista dei dati della ricerca: nel suo ruolo supporta i ricercatori nella gestione attenta e responsabile dei dati e dei prodotti della ricerca durante il loro ciclo di vita, ossia dalla pianificazione alla condivisione con il pubblico nell'ottica di realizzare una scienza sempre più aperta e trasparente».

https://open-science.it/article?rpk=313614&prs_sel=p_datasteward&tpc_sel=t_gestione_dei_dati_della_ricerca

Il lancio della Comunità



Circa 1 anno fa a **Roma** si è svolto il primo incontro in presenza della Comunità.

Alla presenza di **50 professionisti** dei dati della ricerca, provenienti da diverse regioni e istituzioni italiane, sono stati delineati gli **obiettivi fondamentali** di questo gruppo.



Roma, 7 novembre 2023

Comunità Italiana dei Data Steward (CIDS)

Pasquale, V., & PASCHETTA, M. (2024, novembre 25). Un anno di Comunità Italiana Data Steward. GenOA Week 2024, Genova, Italy. Zenodo. <https://doi.org/10.5281/zenodo.14215870>

Comunità italiana dei Data Steward

Cosa è e come partecipare

Perché una Comunità di Data Steward?

Siamo una Comunità nata dal basso con il supporto del Centro di Competenza ICDI e del progetto europeo [Skills4EOSC](#), con l'esigenza di collaborare per

- aiutarci a vicenda nel costruire nuove competenze spesso non direttamente accessibili nei percorsi di formazione tradizionale
- collaborare per favorire la convergenza su soluzioni condivise alla gestione dei dati della ricerca
- aumentare la consapevolezza e la conoscenza del ruolo del Data Steward all'interno e all'esterno delle Istituzioni di ricerca.

Cosa fa la Comunità?

La parola d'ordine è condivisione!

La costruzione della Comunità è un lavoro in divenire, proprio perché lo facciamo in modo partecipativo. Comuniciamo attraverso una mailing list (attualmente più di 100 iscritti) per annunciare e promuovere eventi, confrontarci con i colleghi e attivare collaborazioni.

Ogni circa 2 mesi ci incontriamo online per scambiarci idee e condividere soluzioni. Recentemente, è stato costituito un gruppo di lavoro per la redazione di un *manifesto* della Comunità, che contenga non solo informazioni su chi siamo e cosa vorremmo fare insieme, ma anche sulle modalità di partecipazione.

PASCHETTA, M., CALDONI, G., Pasquale, V., Solieri, A., Maltese, V., Gualandi, B., Tarallo, A., Spadi, A., Pedonese, G., Marchegiani, F., Osmenaj, E., Pavone, G., Sacchi, S., Ferretti, R., Ildiko Kis, N., & Coppini, S. (2025). **Manifesto della Comunità Italiana Data Steward (CIDS)**. Zenodo.

<https://doi.org/10.5281/zenodo.15129936>

https://open-science.it/article?rpk=313614&prs_sel=p_datasteward&tpc_sel=t_gestione_dei_dati_della_ricerca

Sei anche tu un Data Steward?

1. Ti trovi spesso a parlare di dati?
2. Sei la "persona dei dati" all'interno del tuo gruppo di ricerca, team, dipartimento o servizio?
3. Sei la persona che sviluppa/implementa i processi, i database e i flussi di lavoro dei dati all'interno del tuo gruppo di lavoro?
4. Dedichi una parte significativa del tuo tempo alla gestione, estrapolazione, pulizia, controllo qualità di dati e metadati?
5. Amministri o fornisci supporto ad infrastrutture o servizi per la gestione dei dati?
6. Hai supportato la scrittura, l'implementazione, la revisione di un piano di gestione dei dati (data management plan, DMP)?
7. Fornisci supporto alla formazione e/o all'implementazione della corretta gestione dei dati di ricerca come servizio centrale o a un gruppo di ricerca?
8. Sei incaricato della curatela, della conservazione o della FAIRificazione dei dati?
9. Stai lavorando allo sviluppo o all'implementazione di una policy per la gestione dei dati di ricerca?

Checklist tradotta e riadattata da Sonrai – Irish Data Stewardship Network (<https://datastewards.ie/>).

Come partecipare



Scrivi a **data-stewards@lists.icdi.it**



Inserisci alcune informazioni utili:

Qual è la tua formazione?

Per quale istituzione lavori?

Qual è il tuo ruolo lavorativo?

Cosa vorresti trovare nella Comunità?

Cosa vorresti condividere con la Comunità?

Comunità Italiana dei Data Steward (CIDS)

Quiz



<https://www.menti.com/al6k9a9usnsk>

Per approfondire

Le tematiche trattate in questa lezione sono approfondite in questo corso online, che potete trovare registrandovi sul sito di H2IOSC Training Environment.

Qui tutte le informazioni sulla piattaforma e su come accedervi:

<https://www.h2iosc.cnr.it/training-infrastructure/>

Iniziative della comunità CLARIN

CLARIN Newsflash

aggiornamento mensile sulle attività dei consorzi nazionali

CLARIN Café

webinar informali e interattivi per discutere su temi di interesse



Grazie per l'attenzione!

Contatti:

<https://forum.clarin.eu/>
formazione@clarin-it.it

