



# La figura del Data Steward nell'INFN

Francesca Marchegiani


Workshop sul Calcolo nell'INFN - Palau - 20-24 Maggio 2024

Licenza Creative Commons  
Attribuzione 4.0 Internazionale



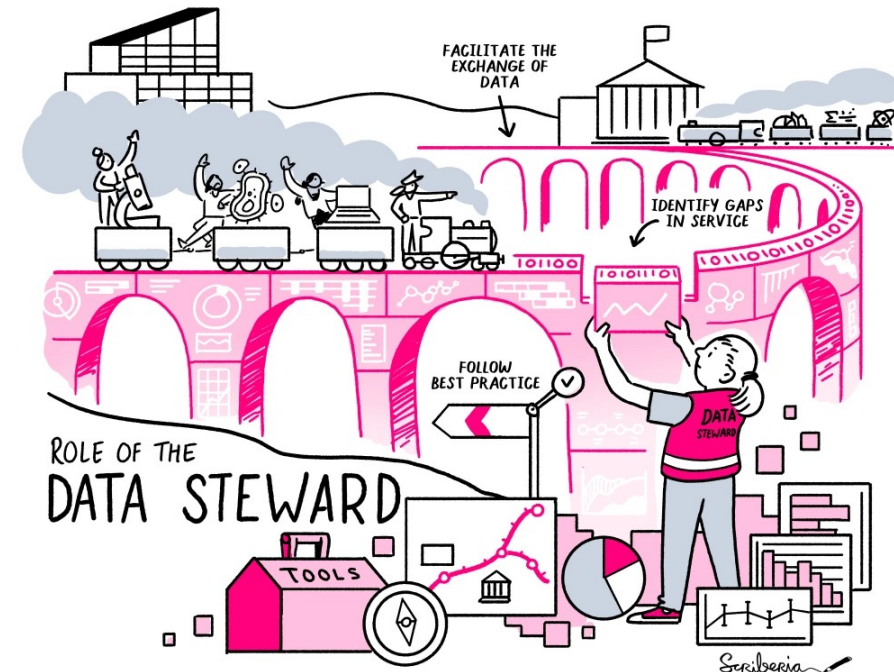


# Agenda

- 
- Introduzione
  - Importanza della gestione dei dati della ricerca
  - Gestione FAIR dei dati della ricerca
  - Chi è il data steward?
  - Competenze e compiti del data steward
  - Considerazioni dal Web
  - Formazione del data steward
  - Quadro europeo
  - Piano Nazionale Scienza Aperta (PNSA)
  - La figura del DS in Italia
  - Come si declina la figura del data steward nell'INFN

# Introduzione

Il data steward è una nuova **figura professionale**, con **conoscenze approfondite** in **ambiti** di ricerca **specifici**, preposta alla corretta ed efficace **gestione FAIR dei dati** della ricerca durante il loro **intero ciclo di vita** e con esperienza nelle tematiche dell'**Open Science**.



<https://openworking.files.wordpress.com/2022/04/data-stewards.jpg?w=1024>

# Importanza della gestione dei dati della ricerca

**Garantire** la **riproducibilità, trasparenza** e **integrità** della ricerca

**Garantire** la **validità** dei risultati scientifici (autentici, completi ed affidabili)

**Garantire** la **rintracciabilità** e il **riutilizzo futuro** dei dati

**Ottimizzare** l'uso delle **risorse** nel caso si duplichi una stessa ricerca

**Soddisfare** le richieste degli **enti finanziatori** (es. progetti Horizon) e le **normative** sulla protezione dei dati

**Concordare** la **proprietà** e la **condivisione** dei dati

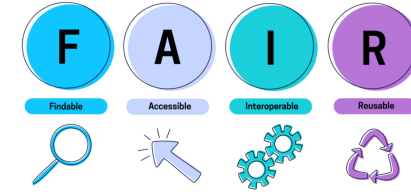
**Evitare** la **perdita di dati** (mancanza di documentazione adeguata per la loro interpretazione, obsolescenza dei formati e dei software che ne garantiscono la lettura, visualizzazione e analisi)

**Incentivare** la **collaborazione** tra ricercatori



<https://doi.org/10.54677/MNMH8546>

# Gestione FAIR dei dati della ricerca



**Principi FAIR<sup>1,2</sup>** (2016): insieme di linee guida (4 principi fondamentali e 15 principi accessori) per una corretta gestione dei dati della ricerca e che ne migliora la **rintracciabilità, accessibilità, interoperabilità e riutilizzo.**

Principi applicabili tanto agli esseri umani quanto alle macchine: capacità di trovare, accedere, combinare e riutilizzare dati attraverso l'uso di algoritmi.

## F<sub>indable</sub>

Dati rintracciabili sia per l'occhio umano che per le macchine in maniera univoca e certa.

- Identificativo persistente (PId)
- Metadati descrittivi comprensivi del PId
- Ricercabili online
- Metadati indicizzati

## I<sub>nteroperable</sub>

Dati strutturati in maniera da garantirne lo scambio ed il riutilizzo tra ricercatori e istituzioni di tutto il mondo.

- Formati largamente diffusi e standards
- Vocabolari controllati
- Schemi condivisi, ontologie, parole chiave
- Evitare formati e software proprietari

<https://doi.org/10.5281/zenodo.8383693>

## A<sub>ccessible</sub>

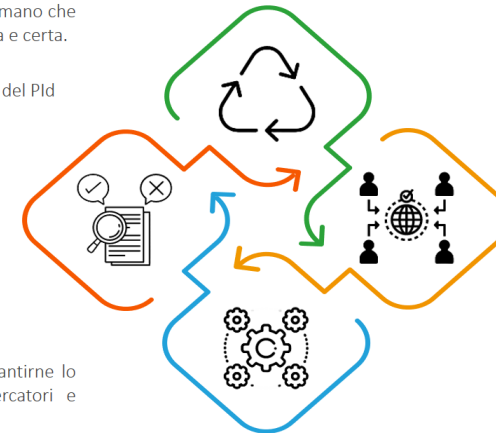
Dati recuperabili online attraverso protocolli standardizzati, reperibili e preservati in un orizzonte temporale a lungo termine.

- Interrogabili online con l'utilizzo di protocolli standardizzati
- Accesso limitato ai dati solo se necessario, accesso aperto ai metadati descrittivi: **As open as possible, as closed as necessary**
- Deposito in un trusted repository (es. Zenodo)

## R<sub>eusable</sub>

Dati corredati da una buona documentazione in modo da poter essere interpretati correttamente, replicati e/o combinati anche in contesti diversi.

- Readme files e documentazione
- Fonte e contesto di provenienza dei dati
- Strumenti necessari per riprodurre i risultati
- Licenze d'uso



<sup>1</sup> <https://www.go-fair.org/fair-principles>

<sup>2</sup> <https://doi.org/10.1038/sdata.2016.18>

# Chi è il data steward?

È una nuova **figura professionale**, con competenze consolidate nel proprio contesto disciplinare, e trasversali, che affianca e **supporta il ricercatore** nella **pianificazione** e **gestione responsabile ed etica** dei **dati della ricerca** durante tutto il loro **ciclo di vita**<sup>1</sup>.

Inserita in enti di ricerca ed università per rispondere alla crescente esigenza di garantire la riproducibilità della ricerca, la qualità dei dati, e l'adesione ai principi di Open Science nata nel contesto della **European Research Area** (ERA) e della **European Data Strategy**.

**Figura ponte** tra ricercatori, infrastrutture ed organizzazioni di ricerca.



<sup>1</sup> <https://doi.org/10.5281/zenodo.3609515>

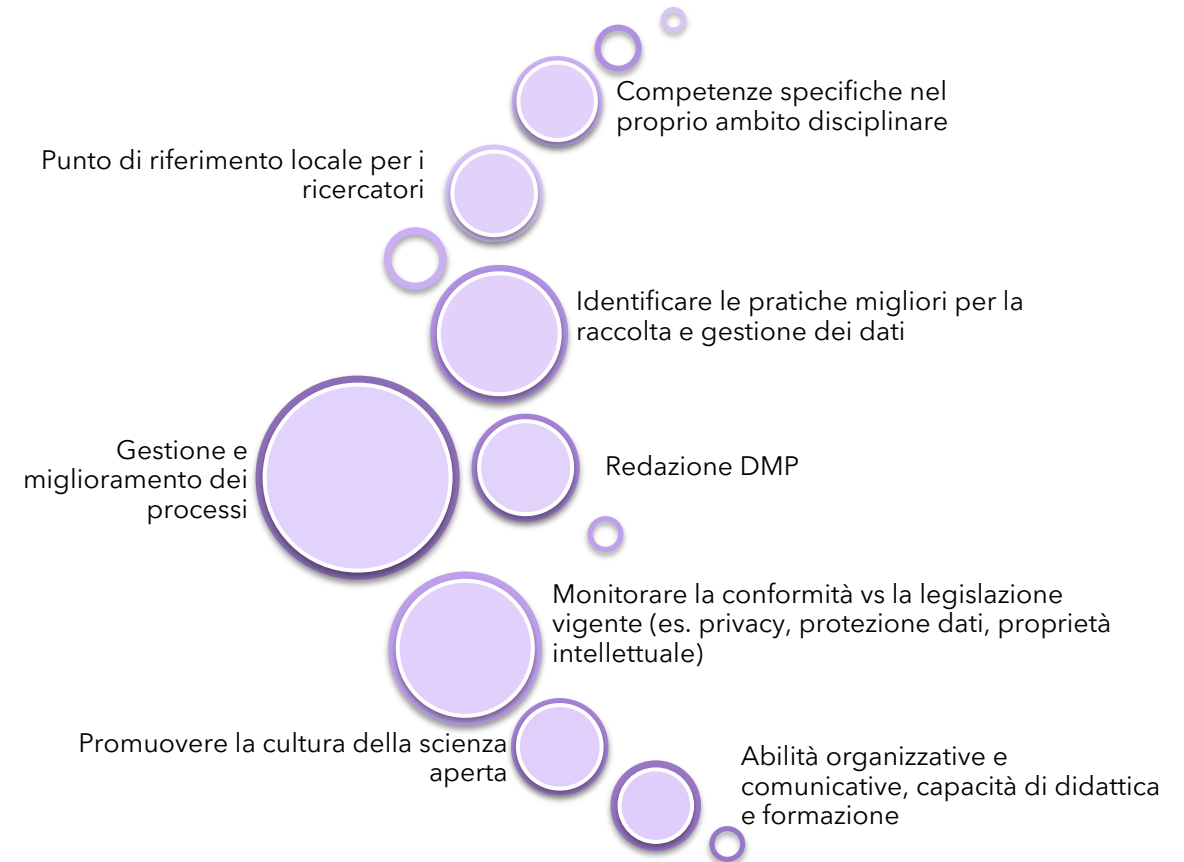
<sup>2</sup> <https://doi.org/10.5281/zenodo.10035214>

<sup>3</sup> <https://doi.org/10.5281/zenodo.4623713>

# Competenze e compiti del data steward

**Competenze trasversali:** disciplinari, informatiche, tecniche e legali.

**Compiti:** supportare i ricercatori nel compito di rendere FAIR i dati della ricerca, implementare i principi FAIR supportando al tempo stesso l'Open Science.



# Considerazioni dal web



André Ilharco | Institute of Public Policies - IPP | 27/06/2023

<https://www.prometheusnetwork.eu/blog/the-european-open-science-cloud-and-its-crucial-guardians-the-data-stewards/>

<https://it.linkedin.com/jobs/data-steward-offerte-di-lavoro>



## Data steward – an emerging and constantly changing role

Published: 2023-05-10

SND's first network meeting in 2023 brought together around 80 participants over two sunny days in Gothenburg. The theme of the meeting was the concept of data steward, which is an emerging role in Swedish higher education institutions.

<https://snd.se/en/nyheter/data-steward-en-framvaxande-roll-i-standing-forandring>

## Data Steward, the fastest growing job

23 February 2024

According to LinkedIn, the profession of Data Steward is gaining popularity, ranking second in its list of the twenty fastest-growing professions of 2024. This trend comes as no surprise in the scientific community, where the importance of FAIR and open data for open science is increasingly recognised. Initiatives, projects, and communities are springing up rapidly, and universities, colleges, and knowledge institutes are increasingly appointing Data Stewards. We are also witnessing a steady rise in Digital Competence Centres.

<https://dans.knaw.nl/en/news/data-steward-the-fastest-growing-job/>

**ETH** zürich

Blog: Attention, Prof: You need a data steward for your team.

<https://ghe.ethz.ch/ghe-blog-news/2024/02/blog-attention-prof-you-need-a-data-steward-for-your-team.html>

<https://www.garr.it/it/news-e-eventi/2150-sondaggio-sui-data-steward-in-italia-una-mappatura-delle-figure-a-supporto-della-gestione-dei-dati-della-ricerca>



# Formazione di un data steward

**Necessari investimenti** per la **formazione** di queste nuove figure professionali che devono essere introdotte all'interno di enti di ricerca ed università con dei programmi specifici <sup>1</sup>.

All'**estero** esistono percorsi focalizzati alla formazione della figura del DS

- Diploma universitario in Scientific Data Management (Université de Montpellier - Francia)

In **Italia** i percorsi che più si avvicinano alla formazione dei DS sono di recente attuazione

- Corso Universitario di Aggiornamento Professionale (CUAP) per Data Steward (UniTO - 2024)

**Formazione continua** ed **integrata** nel tempo che necessariamente deve essere **modulare**, unita alla partecipazione a corsi, seminari, webinar e reti nazionali ed internazionali (es. Conferenze EUDAT, Open Science Café, Simposi EOSC, RDA).

Il progetto europeo **Skills4EOSC** (coordinato dal GARR) ha sviluppato metodologie per armonizzare i profili curriculari, su cui saranno costruiti i percorsi formativi, e risorse didattiche per la formazione di professionisti che sostengono la gestione dei dati di ricerca e promuovono la scienza aperta.

<sup>1</sup> <https://doi.org/10.5281/zenodo.8101903>

# La figura del DS nel contesto internazionale

In alcuni paesi europei (Austria, Danimarca, Olanda<sup>1,2,3</sup>, UK) la **figura del data steward** è stata **già introdotta all'interno dei dipartimenti di ricerca e delle facoltà** e segue un percorso di carriera all'interno del **supporto alla ricerca** quindi separato dalle carriere scientifiche.

Possibilità di accedere a **fondi statali** a tempo determinato per la fase di «start-up»

**Sostenibilità a lungo termine** è lasciata alle organizzazioni di ricerca che devono prevedere l'assunzione a tempo indeterminato di queste figure.

A livello europeo esiste una task force all'interno della EOSC (**Data stewardship, curriculum and career paths EOSC-A Task Force** <sup>4,5</sup>) per la definizione della figura del Data Steward e delle loro attività principali.



1 <https://doi.org/10.5281/zenodo.4623713>  
2 <https://doi.org/10.5281/zenodo.4320504>  
3 <https://doi.org/10.5281/zenodo.3609516>  
4 <https://doi.org/10.5281/zenodo.10573892>  
5 <https://doi.org/10.5281/zenodo.10925394>

# Piano Nazionale Scienza Aperta (PNSA)

Approvazione nel 2022

*Dati della ricerca: "As open as possible, as closed as necessary"*

PIANO NAZIONALE PER LA SCIENZA APERTA (2021 – 2027)	
Assi di intervento	Piano di intervento
1. Pubblicazioni scientifiche	<ul style="list-style-type: none"> <li>• Accesso aperto alle pubblicazioni</li> <li>• Forme non commerciali di pubblicazione</li> <li>• Quadro normativo in materia di diritto d'autore</li> <li>• Sistema di monitoraggio</li> <li>• Risorse formative aperte</li> </ul>
2. Dati della ricerca	<ul style="list-style-type: none"> <li>• FAIRification nel sistema ricerca</li> <li>• Integrazione in EOSC</li> <li>• Produzione collaborativa di dati</li> <li>• Formazione delle figure tecniche</li> </ul>
3. Valutazione della ricerca	<ul style="list-style-type: none"> <li>• Processi e criteri di valutazione</li> <li>• Collaborazione tra istituzioni e tra ricercatori</li> <li>• Pubblicare in accesso aperto</li> <li>• Revisione paritaria aperta</li> <li>• Infrastruttura nazionale</li> </ul>
4. Scienza aperta, comunità scientifica e partecipazione europea	<ul style="list-style-type: none"> <li>• Percorso organico verso la scienza aperta</li> <li>• Attività di coordinamento a livello europeo</li> </ul>
5. Apertura dei dati della ricerca su SARS-COV-2 e Covid-19	<ul style="list-style-type: none"> <li>• Portale nazionale per dati FAIR e testi su COVID19</li> <li>• Modelli di dati aperti sulla salute pubblica</li> </ul>

Tab. 1 - Struttura e obiettivi del piano

...non solo pubblicazioni scientifiche ma anche dati....

Nell'**asse di intervento n. 2 (Dati della ricerca)**, tra i piani di intervento, c'è una sezione dedicata alla "**Formazione delle figure tecniche**".

# La figura del DS in Italia (1 / 3)



ALMA MATER STUDIORUM  
UNIVERSITÀ DI BOLOGNA

## Esperienza @ UNIBO <sup>1</sup>

- 2022 - progetto **Data Steward@UNIBO** per reclutare queste figure professionali con il compito di accompagnare ricercatrici e ricercatori nella gestione dei dati della ricerca, e garantire trasparenza e accessibilità ai risultati del lavoro scientifico.
- Ciascun DS offre supporto su un **dominio disciplinare specifico** ma ha al tempo stesso competenze trasversali nella gestione FAIR dei dati e nelle tematiche dell'Open Science
- Il **team** costituisce il **punto di contatto unico** per ricevere indicazioni sulla gestione dei dati della ricerca
- **Servizio centralizzato presso l'Area della Ricerca** (e non delocalizzate all'interno dei diversi dipartimenti come implementato nella realtà dell'università olandese TU Delft)
- **Interazione continua con le altre strutture** ed aree dell'amministrazione coinvolte (gruppo privacy e DPO, knowledge transfer office, sistema bibliotecario di ateneo, biblioteca digitale) nella gestione della ricerca

<sup>1</sup> <https://doi.org/10.6092/unibo%2Famsacta.7244>

# La figura del DS in Italia (2/3)



## Esperienza @ INGV <sup>1</sup>

### Registro Dati

- Supporto alla compilazione dei metadati
- Valutazione delle schede di metadati
- Sviluppo e gestione infrastruttura informatica

### Data Management Plans

- Supporto alla redazione di DMP

### Supporto all'Amministrazione

- Revisione di documenti che includono attività relative ai dati (es. contratti, accordi)
- Rendicontazione da parte dei Direttori di Sezione
- Compilazione Piano Triennale di Attività (PTA)
- Piano Integrato di Attività e Organizzazione (PIAO)

### Open Science

- Seminari presso le sezioni
- Partecipazione a convegni
- Partecipazione a ICDI
- Partecipazione a GdL OS CoPER

<sup>1</sup> <https://agenda.infn.it/event/37219/contributions/208114/attachments/109195/155058/2023%2008%2028%20Data%20Steward%20-%20L'esperienza%20INGV.pdf>

# La figura del DS in Italia (3/3)

**Nascita** della **comunità italiana dei data steward** <sup>1</sup> a dicembre 2023.

**Competence centre per Open Science, FAIR e EOSC (CC-ICDI)** <sup>2</sup>

- Creare una rete di esperti, iniziative e infrastrutture di ricerca con competenze funzionali a supporto della comunità nazionale per la scienza aperta, i principi FAIR e per la partecipazione italiana alla European Open Science Cloud (EOSC)

**Portale italiano sulla scienza aperta** <sup>3</sup>: contiene informazioni generali, materiale formativo, eventi, novità e aggiornamenti di carattere nazionale e internazionale, materiale sia introduttivo sia risorse utili per problematiche specifiche, analisi e informazioni avanzate.



<sup>1</sup> <https://doi.org/10.5281/zenodo.10035214>

<sup>2</sup> <https://doi.org/10.5281/zenodo.5071055>

<sup>3</sup> [www.open-science.it](http://www.open-science.it)

# @ INFN



- 2021- Costituito GLOS (**gruppo di lavoro per l'Open Science**)

- L'INFN aderisce agli assi di intervento codificati dal **Piano Nazionale per la Scienza Aperta**



PIANO NAZIONALE PER LA  
SCIENZA APERTA



- Con la Delibera n. 16717 del 21/07/2023 è stato approvato il **"Disciplinare per l'accesso aperto ai prodotti della ricerca dell'Istituto Nazionale di Fisica Nucleare"**.



- In corso la migrazione dell'**archivio istituzionale OAR** da INFN-CT a INFN-CNAF/S.N. (presentazione "Open Science per INFN, INFN per Open Science" a cura di Stefano Dal Pra INFN-CNAF)

- Creato un **gruppo di lavoro** per la **stesura dei DMP** dei laboratori nazionali



- **Percorso di formazione** attivato con UNIBO

<https://web.infn.it/openscience/>

# @ INFN

- Partecipazione alle attività OS in Italia e in EU
  - Co-coordinamento del gruppo di lavoro Open Science della CoPER
  - Membro della Italian Computing and Data Infrastructure (ICDI)
  - Partecipazione ai progetti promossi da EOSC
- Co-coordinamento di progetti PNRR multidisciplinari (ICSC/Terabit/...)
- Utilizzo di molteplici piattaforme open-source (Indico, Pandora, Bluemeet, Newdle,...)
- Pubblicazioni scientifiche: Coordinatore italiano di Scoap3, oa2020, scienceeurope, coalition-s
- Valutazione della ricerca: firmato accordo europeo sulla riforma della valutazione della ricerca e partecipazione alla coalizione CoARA.
- Realizzazione progetto CovidStat
- Partecipazione allo sviluppo di ESCAPE Virtual Research Environment (VRE) ed utilizzo nell'ambito di esperimenti di materia oscura





# ... messaggi fondamentali ...

- Data steward: nuova figura professionale
- Competenze disciplinari e trasversali
- Supporto al ricercatore negli aspetti tecnico-scientifici ed amministrativi della gestione dei dati della ricerca
- Formazione continua vista la prospettiva dinamica dell'OS
- Lavoro di squadra perchè c'è continua interazione sia coi ricercatori che con altri servizi dell'ente

Grazie per  
l'attenzione

Francesca Marchegiani (INFN-LNGS)

[francesca.marchegiani@lngs.infn.it](mailto:francesca.marchegiani@lngs.infn.it)

A nome del gruppo:

Luciano Gaido (INFN-TO) - [luciano.gaido@to.infn.it](mailto:luciano.gaido@to.infn.it)

Stefano Bianco (INFN-LNF) - [stefano.bianco@lnf.infn.it](mailto:stefano.bianco@lnf.infn.it)

Daniele Bonacorsi (INFN-BO) - [daniele.bonacorsi@bo.infn.it](mailto:daniele.bonacorsi@bo.infn.it)

Concezio Bozzi (INFN-FE) - [bozzi@fe.infn.it](mailto:bozzi@fe.infn.it)

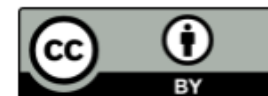
Stefano Dal Pra (INFN-CNAF) - [stefano.dalpra@cnafe.infn.it](mailto:stefano.dalpra@cnafe.infn.it)

Sandra Parlati (INFN-LNGS) - [sandra.parlati@lngs.infn.it](mailto:sandra.parlati@lngs.infn.it)

Irene Piergentili (INFN-LNF) - [irene.piergentili@lnf.infn.it](mailto:irene.piergentili@lnf.infn.it)

Lorenzo Rinaldi (UNIBO) - [lorenzo.rinaldi@unibo.it](mailto:lorenzo.rinaldi@unibo.it)

Licenza Creative Commons  
Attribuzione 4.0 Internazionale



# Brainstorming – idee e suggerimenti

