

# Gestione dei dati della ricerca

Tranquillo Vito

Per **gestione dei dati della ricerca** s'intende l'organizzazione, l'archiviazione, la conservazione e la condivisione dei dati raccolti e utilizzati in un progetto di ricerca. Implica la gestione quotidiana dei dati di ricerca durante la durata di un progetto di ricerca (ad esempio, utilizzando convenzioni di denominazione dei file coerenti). Implica anche decisioni su come i dati verranno conservati e condivisi dopo il completamento del progetto (ad esempio, il deposito dei dati in un repository per l'archiviazione e l'accesso a lungo termine).

Ci sono una serie di ragioni per cui la gestione dei dati di ricerca è importante:

- I dati, come gli articoli di giornale e i libri, sono un prodotto dell'attività di ricerca e quindi patrimonio dell'IZSLER
- I dati (soprattutto i dati digitali) sono fragili e si perdono facilmente.
- Ci sono crescenti requisiti di dati di ricerca imposti da finanziatori ed editori.
- La gestione dei dati di ricerca consente di risparmiare tempo e risorse a lungo termine.
- Una buona gestione aiuta a prevenire gli errori e aumenta la qualità delle analisi.
- I dati ben gestiti e accessibili consentono ad altri di convalidare e replicare i risultati.
- La gestione dei dati di ricerca facilita la condivisione dei dati di ricerca e, se condivisi, i dati possono portare a scoperte preziose da parte di altri al di fuori del team di ricerca originale.

## Cosa sono i dati della ricerca

I dati della ricerca sono definiti dal decreto legislativo n. 36/2006 come *documenti informatici, diversi dalle pubblicazioni scientifiche, raccolti o prodotti nel corso della ricerca scientifica e utilizzati come elementi di prova nel processo di ricerca, o comunemente accettati nella comunità di ricerca come necessari per convalidare le conclusioni e i risultati della ricerca*. Esempi di tale tipologia di dati sono forniti nella Direttiva (UE) 2019/1024 (cd. Direttiva Open Data): statistiche, risultati di esperimenti, misurazioni, osservazioni risultanti dall'indagine sul campo, risultati di indagini, immagini e registrazioni di interviste, oltre a metadati, specifiche e altri oggetti digitali. La Direttiva precisa anche che i dati della ricerca, come da definizione, sono diversi dagli articoli scientifici, in cui si riportano e si commentano le conclusioni della ricerca scientifica sottostante.

I dati della ricerca **DEVONO** essere resi disponibili gratuitamente per il riutilizzo per fini commerciali o non commerciali (dlgs36-2006/opendata/req/research/free)

I dati da considerare sono quelli che rappresentano il risultato di attività di ricerca finanziata con fondi pubblici e resi pubblici, anche attraverso l'archiviazione in una banca dati pubblica gestita a livello istituzionale o su base tematica, da ricercatori, organizzazioni che svolgono attività di ricerca e organizzazioni che finanziano la ricerca. Nel processo di identificazione dei dati della ricerca da rendere disponibili per il riutilizzo è necessario tenere conto della protezione dei dati personali, degli interessi commerciali, dei diritti di proprietà intellettuale<sup>1</sup> e dei diritti di proprietà industriale<sup>2</sup>. Il Decreto dispone che i dati della ricerca DEVONO rispettare i requisiti di reperibilità, accessibilità, interoperabilità e riutilizzabilità che rappresentano i 4 principi del framework FAIR (Findable - Accessible - Interoperable - Reusable)<sup>3</sup>. Di seguito le indicazioni per adempiere a tale disposizione<sup>4</sup>. Findable (Reperibile) - Il primo requisito da rispettare per garantire il riutilizzo dei dati è di rendere i dati reperibili da macchine ed essere umani. Per fare questo, dovrebbero essere resi disponibili i metadati attraverso una risorsa consultabile online e dovrebbe essere assegnato un identificatore persistente a dati e metadati.

Accessible (Accessibile) Deve essere possibile per essere umani e macchine accedere ai dati attraverso protocolli standard e aperti.

Interoperable (Interoperabile) - Dati e metadati devono poter essere combinati con altri dati e/o strumenti. Per questo, devono essere conformi a formati e standard riconosciuti.

Reusable (Riutilizzabile) - I dati devono essere ben documentati in modo che possano essere interpretati correttamente, replicati e/o combinati anche in contesti diversi. Ai dati, inoltre, bisogna assegnare una licenza chiara e accessibile in modo che si possa capire che tipo di riutilizzo è consentito

La necessità che i dati della ricerca seguano e siano conformi ai principi FAIR è ribadita anche dal Programma nazionale per la ricerca (PNR) 2021-2027<sup>5</sup>, che considera tale conformità come uno degli "strumenti" per l'attuazione della scienza aperta. Si legge: "I dati come tutti i

risultati della ricerca devono innanzitutto essere conformi ai criteri FAIR e cioè essere reperibili tempestivamente, accessibili su richiesta, interoperabili e comunque riusabili con le opportune regole, strumenti e risorse. I criteri FAIR sono la base per la fruizione e il riutilizzo di dati di diversa provenienza tematica e metodologica. Fare in modo che la produzione di dati FAIR diventi lo standard di riferimento per i risultati della ricerca finanziata con risorse pubbliche comporta un'innovazione sostanziale nella prassi scientifica sia a livello di produzione dei dati sia a livello della fruizione dei medesimi per elaborare nuove conoscenze. Si tratta di evoluzione delle prassi della ricerca (utilizzo di informazioni esistenti, sviluppo della multidisciplinarietà e della interdisciplinarietà, collaborazione formale e informale) validazione e valutazione dei risultati. Essa apre anche nuove prospettive di innovazione tecnologica (acquisizione automatica di dati FAIR ove possibile e sviluppo di servizi digitali per la cura, l'archiviazione, l'accesso, l'analisi e la fruizione di risorse di calcolo)". In tema di politiche di accesso aperto, il PNR include il "Piano nazionale per la scienza aperta" (non ancora adottato) che ha tra i suoi obiettivi, tra l'altro, l'accesso aperto, appunto, agli strumenti di produzione della ricerca, ai dati generati e alle pubblicazioni scientifiche nonché il sostegno all'accesso ai dati FAIR. Un ruolo essenziale nell'attuazione della strategia sulla scienza aperta e delle politiche di accesso aperto è svolto dalle infrastrutture di ricerca, oggetto del "Piano nazionale per le Infrastrutture di Ricerca"<sup>6</sup> (parte integrante, anch'esso, del PNR), a cui si rimanda per tutti gli aspetti che possono avere un impatto sull'applicazione del Decreto con riferimento alle banche dati pubbliche, gestite a livello istituzionale o su base tematica, necessarie per l'archiviazione dei dati della ricerca da rendere disponibili per il riutilizzo.

## **Policy dell'Istituto sulla gestione dei dati della ricerca**

### **Premessa**

Ai fini della presente policy si considera la definizione di "dati della ricerca" e di "afferenti all'IZSLER" così come da allegato 1 L'IZSLER riconosce l'importanza fondamentale dei dati prodotti durante l'attività di ricerca come un valido risultato scientifico. Pertanto riconosce la rilevanza della loro gestione per il mantenimento dei valori di qualità ed integrità della ricerca scientifica e si impegna ad applicare i più elevati standard per la loro raccolta, archiviazione e conservazione. L'IZSLER riconosce che dati della ricerca affidabili e facilmente reperibili sono alla base di ogni progetto di ricerca e sono altresì necessari per la verifica di attendibilità e correttezza della conduzione e dei risultati del progetto e per la sua riproducibilità L'IZSLER riconosce che i dati della ricerca sono un valore pubblico e costituiscono patrimonio dell'Istituto, nonché risorsa - anche a lungo termine - per la ricerca, la didattica ed il progresso della società...(impatto sociale—valore pubblico)

## **Ambiti di applicazione**

La presente “policy” si applica a tutti i progetti di ricerca dell’IZSLER limitatamente alle parti di cui essa è responsabile attraverso i propri ricercatori che sono tenuti ad osservarla. Nel caso la ricerca sia stata finanziata da parti terze ed esistano accordi specifici relativi al controllo dei dati, al loro accesso e conservazione, gli accordi prevalgono sulla presente policy sempre nel rispetto della normativa vigente.

## **Trattamento dei dati della ricerca**

Nel rispetto della vigente normativa in materia di protezione dei dati personali e di proprietà intellettuale, nonché delle disposizioni contenute nello Statuto e nei regolamenti dell’Università e fatti salvi gli specifici accordi per il finanziamento della ricerca stipulati con terze parti,

i dati della ricerca, una volta pubblicati, sono archiviati e resi liberamente disponibili all’uso per finalità di ricerca scientifica o storica, o di pubblico interesse

I dati della ricerca, una volta pubblicati, devono essere archiviati nell’archivio digitale dell’IZSLER (AIR) o in un archivio digitale che rispetti gli standard internazionali appartenente ad istituzioni non commerciali accademiche o di ricerca, enti statali o intergovernativi

I dati della ricerca collegati alle pubblicazioni alle quali si riferiscono devono essere archiviati in modo corretto, completo, affidabile, rispettandone l’integrità. Devono inoltre essere accessibili, identificabili, tracciabili, interoperabili e, laddove possibile, disponibili per usi successivi (principi FAIR).

Non esistendo sui dati diritti di terze parti e non sussistendo usi vietati per legge, i dati della ricerca possono essere associati ad una licenza per il libero utilizzo (ad esempio Creative Commons), garantendo la tracciabilità degli usi e il credito verso la fonte originaria.

La durata minima del periodo di archiviazione è di 10 anni dal rilascio al pubblico dei risultati. Trascorso tale periodo i dati possono essere cancellati, a meno che permangano le finalità di cui in premessa ovvero rilevanti interessi di ricerca.

Nel caso i dati della ricerca vengano distrutti o cancellati (per il termine del periodo di archiviazione o per ragioni etiche o giuridiche), tale azione viene intrapresa dopo aver tenuto conto di tutti gli aspetti etici e legali. La eventuale cancellazione o distruzione di dati della ricerca deve essere tracciabile e la relativa documentazione accessibile. Allo stesso tempo si devono tenere in considerazione gli interessi di eventuali terze parti finanziatrici della ricerca e di altri portatori di interessi, così come aspetti di confidenzialità e sicurezza.

## **Responsabilità, doveri, diritti**

La responsabilità della raccolta, gestione e manutenzione dei dati è condivisa tra i ricercatori dell'IZSLER e l'IZSLER stesso, nei limiti dei rispettivi ambiti di competenza e nel rispetto del Codice etico.

E' responsabilità dei ricercatori IZSLER la gestione dei dati della ricerca e dei datasets, ovvero:

- Raccolta, documentazione, archiviazione, accesso, uso e conservazione (o distruzione) dei dati della ricerca, compresa la definizione di protocolli e responsabilità nel gruppo di ricerca che vanno incluse in un Data Management Plan (DMP) 1 compilato, laddove previsto da contratti con finanziatori della ricerca o altre entità legali, già nella fase iniziale del progetto
- Elaborazione e aggiornamento del DMP (si rinvia qui alle linee guida sui DMP e ai modelli di DMP) e definizione dei diritti di utilizzo dei dati dopo il termine del progetto;
- Realizzazione di copie di sicurezza (backup) secondo le linee guida e armonizzazione con i requisiti istituzionali e legali o contrattuali relativi ai dati della ricerca e alla gestione dei record che li descrivono.
- In generale, tutti i dati rilevanti -positivi e negativi- per la ricerca stessa, devono essere archiviati e conservati con i medesimi criteri di cui sopra su adeguati supporti informatici, anche al di fuori di specifici DMPs, in condizioni di accesso riservato al/i responsabile/i del progetto/i. E' compito del responsabile del progetto di ricerca definire quali tra i dati debbano essere conservati, tenendo conto dei doveri risultanti da contratti con parti terze.

E' responsabilità dell'IZSLER nominare personale di supporto competente e fornire risorse, strumenti e infrastrutture adeguate in termini di software e hardware atti alla conservazione dei dati e conformi a questa policy.

-Fornire assistenza per la progettazione e compilazione dei DMPs

- Sviluppare servizi e meccanismi per la registrazione dei progetti, per consentire il deposito, l'archiviazione e la conservazione dei dati della ricerca e mantenerne l'accessibilità durante e al termine del progetto di ricerca. Far sì che gli afferenti all'IZSLER possano esercitare le loro responsabilità e ottemperare a obblighi previsti dai contratti con i finanziatori della ricerca o altre entità legali

## **Validità**

La presente "policy" entra in vigore il xxxxx e verrà aggiornata ogni due anni

La "policy" è stata approvata dal CdA in data xxxx

## **Allegato 1**

### **Definizione di dati della ricerca**

I dati della ricerca sono record fattuali (dati numerici, simboli, testi, immagini o suoni) utilizzati come fonti primarie della ricerca scientifica necessari per convalidare i risultati della ricerca.

A supporto della definizione di dati della ricerca si allegano tre modelli considerati nella definizione corrente:

- a) LERU Roadmap for Research Data (LERU Research Data Working Group, Advice Paper No. 14 – December 2014): “Research data, from the point of view of the institution with a responsibility for managing the data, includes: All data which is created by researchers in the course of their work, and for which the institution has a curational responsibility for at least as long as the code and relevant archives/record keeping acts require, and third-party data which have originated within the institution or come from elsewhere.”
- b) The Australian Griffith University: “Research data are factual records, which may take the form of numbers, symbols, text, images or sounds, which are used as primary sources for research, which are commonly accepted in the research community as necessary to validate research findings.”
- c) The University of Minnesota: “Research data are data in any format or medium that relate to or support research, scholarship, or artistic activity. They can be classified as:
  - Raw or primary data: information recorded as notes, images, video footage, paper surveys, computer files, etc.
  - Processed data: analyses, descriptions, and conclusions prepared as reports or papers
  - Published data: information distributed to people beyond those involved in data acquisition and administration

### **Definizione dei ricercatori e personale di supporto alla ricerca dell'IZSLER ai quali si applica la presente policy.**

Personale della ricerca (Ds e D), Dirigenti veterinari, Dirigenti Sanitari, Dirigenti chimici, personale amministrativo di supporto alla ricerca

## **Data management Plan (DMP)**

Il DMP è un documento che progetta l'intero ciclo di vita dei dati della ricerca e può essere continuamente aggiornato. Il DMP assicura che i dati della ricerca siano accessibili, tracciabili, disponibili, autentici, citabili e che rispondano a requisiti legali definiti con chiarezza e a misure appropriate di sicurezza in relazione al loro riutilizzo. Nella loro forma ideale i DMP saranno in un formato machine actionable.

Documenti importanti per la valutazione della policy: <https://www.force11.org/group/fairgroup/fairprinciples>  
Elenca i principi a cui le istituzioni devono ispirarsi <http://learn-rdm.eu/en/research-data-management-toolkit-now-available/> raccoglie la documentazione su cui ci si è basati per la elaborazione della policy (linee guida del progetto europeo guidato da LERU) [https://ec.europa.eu/research/openscience/pdf/eosc\\_declaration.pdf#view=fit&pagemode=none](https://ec.europa.eu/research/openscience/pdf/eosc_declaration.pdf#view=fit&pagemode=none)  
documento che andrà in discussione alla riunione dei rettori della LERU per essere sottoscritto (raccoglie indicazioni e suggerimenti di LEARN e di FORCE11)

## **Riferimenti**

### **Linee Guida Open Data**

<file:///C:/Users/vito.tranquillo/Downloads/lg-opendata-docs.pdf> Le presenti Linee Guida, emesse ai sensi dell'art. 12 del decreto legislativo n. 36/2006 e s.m.i., dell'art. 71 del decreto legislativo n.82/2005 (CAD) e dell'art. 7 della Deliberazione n. 160/2018 recante "Regolamento per l'adozione di linee guida per l'attuazione del Codice dell'Amministrazione Digitale", definiscono le regole tecniche relative all'apertura dei dati e al riutilizzo dell'informazione del settore pubblico.