

GLOSSARIO DataCite

<p>API (Application Programming Interface)</p>	<p>Un'API è un insieme di regole e protocolli che consente a software diversi di comunicare tra loro. Fornisce metodi standardizzati per richiedere e scambiare dati o funzionalità tra sistemi, applicazioni o servizi senza dover comprendere i dettagli interni del loro funzionamento. Le API sono fondamentali per l'integrazione di software, facilitando l'automazione, la scalabilità e lo sviluppo di nuove applicazioni su piattaforme esistenti.</p> <p style="text-align: right;"><i>Fonte: Oxford English Dictionary</i></p>
<p>Citation</p>	<p>Una citazione è un riferimento ai risultati di ricerca di supporto che sono alla base del contenuto pubblicato. Allo stesso modo in cui i ricercatori forniscono di routine un riferimento bibliografico ad altre risorse accademiche, dovrebbero essere inclusi anche i riferimenti ad altri risultati di ricerca come dati, software, strumenti e campioni. Questi tipi di risorse sono spesso condivisi, ma non sono spesso citati nello stesso modo degli articoli di riviste o di altre pubblicazioni. Ciò è importante perché facilita l'accesso, la trasparenza e la riproducibilità, il riutilizzo, il merito per i ricercatori e la visibilità per i repository che condividono i risultati della ricerca. I repository possono partecipare includendo informazioni sulle pubblicazioni che citano i dati nei metadati DOI .</p> <p style="text-align: right;"><i>Fonte: DataCite</i></p>
<p>Data</p>	<p>Un dato (digitale) è una rappresentazione reinterpretabile di informazioni in modo formalizzato adatto per la comunicazione, l'interpretazione o il trattamento. Esempi di dati includono una sequenza di bit, una tabella di numeri, i caratteri su una pagina, la registrazione dei suoni fatti da una persona che parla o un campione di roccia lunare.</p> <p style="text-align: right;"><i>Fonte: Open Archival Information System</i></p>
<p>Database</p>	<p>Un database è una raccolta organizzata di dati strutturati, memorizzati elettronicamente in modo sistematico. È progettato per consentire un accesso, una gestione, un aggiornamento e un recupero efficienti dei dati. I database sono spesso gestiti tramite un sistema di gestione di database (DBMS) e possono essere relazionali o non relazionali, a seconda della loro struttura. Sono utilizzati in vari campi per gestire grandi quantità di informazioni in modo ordinato e accessibile.</p> <p style="text-align: right;"><i>Fonte: Oxford English Dictionary</i></p>
<p>DataCite</p>	<p>DataCite è un'organizzazione no-profit, leader a livello mondiale, che fornisce identificatori persistenti (in particolare DOI) per i dati della ricerca e dei suoi risultati. L'adesione del CNR a DataCite come membro consente di assegnare DOI a tutti i risultati della ricerca rendendoli reperibili, citabili e connessi. I metadati associati vengono resi disponibili a tutta la comunità scientifica internazionale.</p> <p>DataCite NON è un Data repository.</p> <p style="text-align: right;"><i>Fonte: DataCite</i></p>

Dataset	<p>Un dataset è qualsiasi raccolta organizzata di dati in formato computazionale, definita da un tema o da una categoria che riflette ciò che viene misurato/osservato/monitorato. La presentazione dei dati nell'applicazione è resa possibile attraverso i metadata.</p> <p style="text-align: right;"><i>Fonte: CoreTrustSeal</i></p> <p style="text-align: right;"><i>Altre fonti:</i></p> <p style="text-align: right;">https://www.databricks.com/glossary/what-is-dataset https://dictionary.cambridge.org/dictionary/english/dataset https://www.treccani.it/enciclopedia/data-set_%28Dizionario-di-Economia-e-Finanza%29/ https://it.wikipedia.org/wiki/Dataset https://www.paris21.org/sites/default/files/1300.pdf https://dictionary.archivists.org/entry/data-set.html</p>
Data catalogue	<p>Un data catalogue è un inventario di risorse di dati all'interno di un'organizzazione. Fornisce metadata descrittivi che consentono di identificare, gestire e accedere ai dati disponibili. Lo scopo principale di un data catalogue è facilitare la scoperta dei dati e la comprensione delle loro caratteristiche, origine, utilizzo e disponibilità.</p> <p style="text-align: right;"><i>Fonte: Gartner Glossary</i></p>
Data dictionary	<p>Un data dictionary è un archivio formale di termini utilizzati per descrivere i dati.</p>
Data manager	<p>La figura del data manager rappresenta una funzione chiave, responsabile della gestione, organizzazione e supervisione dei dati nell'ambito di un'organizzazione o un progetto. Si occupa di assicurare che i dati siano raccolti, archiviati, protetti e utilizzati in modo efficiente e sicuro, conformemente alle normative e alle politiche di governance dei dati. Il ruolo include la gestione della qualità, dell'accesso e della protezione dei dati, spesso in contesti di ricerca, sanità o aziende che trattano grandi volumi di informazioni.</p> <p style="text-align: right;"><i>Fonte: European e-Competence Framework</i></p>
Data repository	<p>Un data repository è un archivio centralizzato e strutturato dove vengono memorizzati, gestiti e mantenuti dati digitali per consentire l'accesso e l'uso a lungo termine. Può contenere una varietà di formati di dati, come file di testo, immagini, video e dataset. I repository possono essere pubblici o privati e offrono strumenti per l'archiviazione, la conservazione e la condivisione dei dati in modo sicuro.</p> <p style="text-align: right;"><i>Fonte: OECD Glossary of Statistical Terms</i></p>
DOI	<p>Il DOI (<i>Digital Object identifier</i>) è un identificatore univoco e permanente assegnato a un oggetto digitale, come un articolo scientifico, un dataset, un rapporto o un documento. Rappresenta un collegamento persistente, rendendo accessibile l'oggetto su internet anche se il suo URL cambia nel tempo. Il DOI è gestito da un'organizzazione registrata ed è ampiamente utilizzato in ambito accademico e scientifico per garantire la tracciabilità e l'accesso duraturo alle risorse digitali. Il DOI non è un indicatore di qualità.</p> <p style="text-align: right;"><i>Fonte: International DOI Foundation</i></p>

<p>FAIR (principi lato macchina)</p>	<p>FAIR (Findable Accessible Interoperable Reusable)</p> <p>Findable - Qualsiasi risorsa scientifica deve essere rintracciabile affinché possa circolare, massimizzando così l’impatto della ricerca. A questo scopo sono necessari metadati accurati e processabili dalle macchine, che rendano possibile rilevare in automatico dataset o altri prodotti della ricerca.</p> <p>Accessible - Una volta trovata, si potrebbe voler fruire una determinata risorsa. Tuttavia - poiché non è sempre possibile rendere pubblicamente accessibile qualsiasi contenuto - per ciascun prodotto della ricerca va indicato quali sono le modalità per consultarlo, per esempio se serve identificarsi o inoltrare la richiesta di accesso. Idealmente tutto ciò dovrebbe essere processabile in automatico dalla macchina.</p> <p>Interoperable - I principi che operano per l’interoperabilità cercano di rispondere ai requisiti, sempre più necessari, di integrare varie fonti di dati, interagire con applicazioni o flussi di lavoro per l’analisi, l’archiviazione, o qualsiasi altro tipo di elaborazione.</p> <p>Reusable - Il riuso è un aspetto molto importante nella pratica della scienza aperta. Per esempio uno stesso set di dati si può riusare per replicare un esperimento - stessa domanda e stesso metodo - per cercare nuove risposte o altri metodi per rispondere alla stessa domanda. O si possono usare per rispondere a domande di ricerca completamente diverse. Affinché i dati siano riusabili, i dati devono essere ben descritti, con metadati curati.</p> <p style="text-align: right;">Fonte: OPEN SCIENCE</p> <p style="text-align: right;">Altre info ed esempi: https://sites.google.com/unimib.it/research-data-mgmt-sc-parma/fair-research-data?authuser=0</p>
<p>Infrastructure</p>	<p>L'infrastruttura si riferisce al framework di base fondamentale di sistemi e reti che sottostanno a una struttura, processo o organizzazione. I PID (<i>Persistent Identifier</i>) svolgono un ruolo cruciale nell'infrastruttura di ricerca che supporta lo sviluppo, la distribuzione e l'uso di ricerca e informazioni. La ricerca si basa sull'infrastruttura sottostante che include piattaforme e servizi digitali come quelli sviluppati da DataCite.</p> <p style="text-align: right;">Fonte: DataCite</p>
<p>Landing page</p>	<p>Tutti i DOI si risolvono, tramite il server di gestione globale, in una landing page. La landing page è una pagina html, visualizzabile sul web, che può essere creata per permettere all’utente di accedere a dei contenuti. Nel caso di DataCite è un modo per l'utente di accedere al contenuto sottostante (ad esempio, dove può scaricare un set di dati) e contiene metadati sulla risorsa. È responsabilità dell'entità che crea il DOI fornire tale landing page. Una landing page/URL deve essere registrata per creare un DOI pubblico e reperibile.</p> <p style="text-align: right;">Fonte: DataCite</p>

<p>Metadati</p>	<p>I metadati sono dati che descrivono altri dati, fornendo informazioni aggiuntive per identificarli, organizzarli, e facilitarne l'uso. Possono includere dettagli come l'autore, la data di creazione, il formato del file, la provenienza, e le specifiche tecniche di un dataset o di un documento. I metadati sono fondamentali per migliorare la scoperta, la gestione e l'interoperabilità dei dati in vari contesti, inclusi archivi digitali, database e data catalogue.</p> <p style="text-align: right;">Fonte: National Information Standards Organization</p>
<p>Open science</p>	<p>L'Open science (scienza aperta) si riferisce a un processo di ricerca scientifica trasparente, collaborativo e accessibile a tutti. Questo approccio enfatizza la condivisione di dati, metodi e risultati con la comunità scientifica e il pubblico al fine di facilitare la verifica dei risultati e consentire ad altri di sviluppare e replicare il lavoro. La scienza aperta aiuta a promuovere la trasparenza scientifica, la responsabilità e, dove possibile, la riproducibilità, e ha il potenziale per accelerare il ritmo della scoperta e dell'innovazione.</p> <p style="text-align: right;">Fonte: DataCite</p>
<p>PID</p>	<p>PI o PID (<i>Persistent Identifier</i>) sono riferimenti durevoli e univoci assegnati a una risorsa digitale. Sono fondamentali per identificare con esattezza, trovare e citare un lavoro, un dataset, un report, un campione o qualsiasi altro prodotto della ricerca. Sono identificatori persistenti e risolvibili dalla macchina con uno schema di metadati associato.</p> <p style="text-align: right;">Fonte: DataCite</p>
<p>re3data</p>	<p>Re3data (<i>REgistry of REsearch Data Repository</i>) è il registro mondiale dei repository dedicati ai dati della ricerca di diverse discipline accademiche</p> <p style="text-align: right;">Fonte: re3data</p>
<p>Repository</p>	<p>All'interno di DataCite i repository rappresentano l'unità di riferimento che contiene un ruolo chiave nel modello di appartenenza a DataCite e sono definiti come un servizio gestito da organizzazioni di ricerca, in cui i materiali di ricerca vengono archiviati, gestiti e resi accessibili.</p> <p><u>Un repository è un'unità singola.</u></p> <p>Un account repository rappresenta un archivio in cui tutti i DOI per un gruppo di risorse sono registrati e rimarranno insieme. Hanno un prefisso univoco e sono utilizzati esclusivamente per la registrazione DOI. Un account repository può appartenere a un membro diretto o a un'organizzazione consorziata.</p> <p style="text-align: right;">Fonte: DataCite</p>

Research output	Un risultato di ricerca (research output) copre un'ampia gamma di risorse prodotte tramite la ricerca. Alcuni esempi includono: set di dati, software, campioni, tesi, immagini e pubblicazioni. <i>Fonte: DataCite</i>
Schema di metadati	Uno schema di metadati è una struttura o modello predefinito che specifica come i metadati devono essere organizzati, descritti e formattati per un insieme di dati. Stabilisce quali elementi o campi devono essere inclusi (ad esempio, titolo, autore, data di pubblicazione) e come devono essere rappresentati. Gli schemi di metadati aiutano a garantire coerenza e interoperabilità tra diversi sistemi di gestione dei dati, facilitando lo scambio e l'accesso ai dati attraverso piattaforme e settori diversi. <i>Fonte: Dublin Core Metadata Initiative</i>