

Gestione dati e pratiche aperte nelle infrastrutture di ricerca sociale

Destinatari: personale di ricerca e tecnologo CNR

Durata: 11 ore complessive

Modalità: Online (piattaforma di formazione GESFOR GoToWebinar)

Partecipanti: Max 100

Format: 2 incontri da 5 ore e 6 ore:

- **23 Febbraio:** Modulo 1 ore 10:00 – 12:30; Modulo 2 ore 14:00 – 16:30
- **24 Febbraio:** Modulo 3 ore 10:00 – 13:00; Modulo 4 ore 14:00 – 17:00

Obiettivo generale: fornire ai ricercatori competenze e pratiche aggiornate su gestione FAIR dei dati, integrità della ricerca e uso dei servizi delle infrastrutture di ricerca CNR del contesto nazionale ed europeo (CESSDA).

Risultati attesi:

- Comprensione dei principi FAIR e della loro applicazione pratica.
- Maggiore consapevolezza sull'integrità e la condivisione responsabile dei dati.
- Capacità di utilizzare i principali strumenti di CESSDA e infrastrutture correlate.
- Creazione di una rete di contatti tra ricercatori CNR che operano nel dominio.

Attestazione corso: è previsto il rilascio di un attestato di partecipazione al corso e la somministrazione di un questionario di gradimento.

Per chi fosse interessato ad argomenti specifici, è possibile seguire singolarmente il/i modulo/i di interesse

23 Febbraio

10.00 – 12:30

Modulo 1 – FAIR Data Management nelle Scienze Sociali (2,5h)

Obiettivi:

- Introduzione e motivazioni, vantaggi della gestione FAIR dei dati
- Comprendere i principi FAIR e la loro applicazione nei contesti delle infrastrutture di dati sociali
- Riconoscere il ruolo delle policy europee (EOSC, Horizon Europe, ...)
- Applicare concetti di metadati, licenze e di data management plan

Contenuti:

- Introduzione al ciclo di vita dei dati di ricerca: perchè è importante gestire correttamente i dati – what's in it for me
- I principi FAIR
- Esempi di buone pratiche di gestione dei dati. Concetto di proprietà dei dati
- Esercitazione interattiva: analisi di casi concreti (via Mentimeter o breakout rooms)

Relatori: Emma Lazzeri – Ufficio Agenda Digitale e Processi della Direzione Generale CNR

14.00 – 16:30

Modulo 2 – Etica della Ricerca e trattamento dei dati personali (2,5h)

Obiettivi:

- Promuovere buone prassi nella gestione e condivisione dei dati nel settore delle scienze sociali
- Trasmettere i principi di riferimento in materia di etica e integrità nella ricerca
- Analizzare i problemi etici in progetti nel settore delle scienze sociali connessi alla liceità della raccolta e al riuso dei dati
- Analizzare i diritti dei soggetti interessati o donatori di dati e misure di protezione del dato nella condivisione per finalità di ricerca

Contenuti:

- Funzioni e attività della Commissione per l'Etica e l'Integrità nella Ricerca del CNR e del Centro Interdipartimentale per l'Etica e l'Integrità nella Ricerca (CID Ethics-CNR)
- Etica della ricerca, integrità nella ricerca e qualità del dato
- Procedure per la richiesta di un parere etico con valore autorizzativo per progetti di ricerca di scienze sociali che utilizzano servizi di un'infrastruttura
- Ethics by design e privacy by design
- Ethical issues: coinvolgimento dei partecipanti, consenso informato, informativa privacy (GDPR), situazioni di vulnerabilità, uso di incentivi, incidental findings

Relatori: Roberta Martina Zagarella e Elena Mancini, con introduzione/moderazione di Marco Annoni – Centro Interdipartimentale per l'Etica e l'Integrità nella Ricerca (CID Ethics) del CNR

24 Febbraio

10.00 – 13:00

**Modulo 3 – Strumenti e Servizi delle infrastrutture di ricerca per le Scienze Sociali –
La Data Discovery (3h)**

Obiettivi:

- Comprendere il ruolo delle infrastrutture di ricerca (DASSI e CESSDA) nei processi di scoperta, accesso e riuso dei dati nelle scienze sociali
- Individuare i dataset rilevanti per le proprie attività di ricerca utilizzando il CESSDA Data Catalogue
- Utilizzare vocabolari controllati e standard di metadattazione per affinare le strategie di ricerca e interpretare correttamente i metadati
- Valutare la qualità, copertura e riusabilità dei dati attraverso l'analisi della documentazione associata ai dataset
- Accedere e scaricare i dati nel rispetto delle condizioni di accesso, licenze e responsabilità d'uso definite dai data provider
- Acquisire consapevolezza

Contenuti:

3.1 Data Discovery nelle infrastrutture di ricerca per le scienze sociali

- Il concetto di data discovery nel ciclo di vita dei dati di ricerca.
- Perché i metadati sono centrali per la scoperta e il riuso dei dati
- Ruolo di CESSDA e DASSI nell'ecosistema europeo (EOSC)

3.2 I Data Catalogue: DASSI e CESSDA

- Struttura e contenuti del Data Catalogue
- Tipologie di risorse: studi, dataset, serie longitudinali
- Strategie di ricerca: ricerca libera; filtri tematici, geografici e temporali; ricerca per metodologia e popolazione
- Lettura critica della scheda di un dataset: descrizione dello studio; copertura spaziale e temporale; metodologia; condizioni di accesso e licenza
- Collegamento tra metadati e FAIR principles

3.3 Vocabolari controllati e servizi semantici

- Perché usare vocabolari controllati nella ricerca dei dati.
- Introduzione ai CESSDA Controlled Vocabularies: concetti chiave (topic, data type, analysis unit, time method, ecc.)
- Impatto dei vocabolari sulla qualità della ricerca e sull'interoperabilità

3.4 Comprendere i dati prima di scaricarli

- Importanza della documentazione (questionari, codebook, note di archivio)
- Come valutare se un dataset è adatto al proprio obiettivo di ricerca
- Licenze, condizioni di accesso e responsabilità dell'utente
- Buone pratiche per il riuso corretto dei dati (citazione, attribuzione)

24 Febbraio

Esercitazione:

I partecipanti saranno guidati attraverso un'esercitazione pratica in cui, partendo da una domanda di ricerca concreta, impareranno a trasformare un'esigenza di ricerca in una strategia efficace di esplorazione dei dati. Si utilizzeranno la piattaforma DASSI e il CESSDA Data Catalogue per individuare i dataset potenzialmente pertinenti, applicando parole chiave, filtri tematici e vocabolari controllati. Verrà dedicata attenzione all'analisi della documentazione associata ai dataset selezionati per comprenderne metodologia, copertura e condizioni di accesso, valutandone l'idoneità all'obiettivo di ricerca. In conclusione, ogni gruppo presenterà brevemente le scelte effettuate e le motivazioni, evidenziando punti di forza e limiti delle soluzioni individuate, promuovendo una riflessione collettiva sulla qualità delle strategie di data discovery e sulle implicazioni per la ricerca.

Relatori: Domingo Scisci – Università degli Studi di Milano Bicocca

14.00 – 17:00

Modulo 4 – Strumenti e Servizi delle infrastrutture di ricerca per le Scienze Sociali– DEPOSITO (3h)

Obiettivi:

- Imparare a preparare i dati di ricerca per la condivisione
- Esplorare i servizi e gli strumenti per il deposito dei dati della infrastruttura DASSI
- Applicare buone pratiche di gestione e documentazione dei dati in conformità con i principi FAIR
- Valorizzare i propri dati di ricerca accrescendo la loro visibilità e riutilizzabilità

Contenuti:

- I vantaggi di archiviare e condividere i dati di ricerca attraverso DASSI: modalità di conservazione dei dati in modo sicuro, maggiore visibilità e riutilizzabilità, assegnazione di un Digital Object Identifier (DOI) per rendere i dati citabili e accessibili
- Panoramica sulle buone pratiche e requisiti per la preparazione dei dati al deposito: formati, strumenti e servizi per l'archiviazione
- Standard disciplinari per la documentazione dei dati nell'ambito delle scienze sociali: lo standard DDI (Data Documentation Initiative) e il DASSI Metadata model (DaMM)
- Conformità etico-legali per la condivisione dei dati di ricerca, con particolare focus sulle possibili modalità di accesso ai dati, e sulle licenze di riutilizzo
- Esercitazione interattiva: applicazione pratica del workflow di preparazione e deposito dei dati attraverso l'infrastruttura di ricerca DASSI
- Valutazione finale del corso via Mentimeter (survey di valutazione del corso da condividere con i partecipanti nell'ultimo quarto d'ora di questo modulo)

24 Febbraio

Esercitazione:

Durante il modulo è prevista un'esercitazione interattiva dedicata alla simulazione del deposito dei dati di ricerca nell'infrastruttura DASSI. I partecipanti lavoreranno su dati reali già in loro possesso, sperimentando in prima persona le fasi principali del processo: preparazione e documentazione dei dati, definizione dei metadati, scelta delle condizioni di accesso e delle licenze, attenzione agli aspetti etici e di protezione dei dati. L'attività, svolta individualmente, è pensata per fornire competenze pratiche immediatamente applicabili e favorire il confronto tra ricercatori sulle buone pratiche di gestione e condivisione dei dati nelle scienze sociali.

Requisiti:

Per una partecipazione efficace all'esercitazione, è richiesto ai partecipanti di portare con sé dati reali (anche non definitivi) derivanti dalla propria attività di ricerca, corredati, se disponibile, da documentazione di supporto (ad es. questionari, tracce di interviste, note metodologiche, Data Management Plan, ecc.). Questa modalità consente di adattare l'attività a diversi profili di ricerca e di favorire lo scambio di esperienze e competenze differenti.

Relatori: Carlo Pisano – Università degli Studi di Milano Bicocca e Emma Lazzeri – Ufficio Agenda Digitale e Processi della Direzione Generale CNR