

Open Science Webinar Series

Serie 2

2.1 Research Data Management e Politiche Europee sui dati Webinar 15 Novembre 2018 ore 12.00

Elena Giglia ed Emma Lazzeri

Domande e Risposte

Vorrei chiedere di approfondire il rapporto tra OS/OD e Terza Missione dell'Università. Grazie

Gli Open Research Data sono parte della Open Science. Spesso la Open Science viene definita un "concetto ombrello" che comprende appunto testi, dati, software, protocolli, risorse educative aperte. Il concetto è che ogni passo del processo di ricerca può essere aperto, come si vede nell'arcobaleno della Open Science di Jeroen Bosman e Bianca Kramer (<https://zenodo.org/record/1195653>).

Quanto alla Terza missione, la Open Science è il modo migliore che ha un Ateneo per aprirsi al territorio: mettere a disposizione di tutti dati e testi significa aprire le porte della ricerca a Piccole e Medie Imprese, start up, professionisti e semplici cittadini. Open Science ha anche un legame profondo con Citizen Science (almeno all'estero!), per cui è davvero il modo migliore per coinvolgere l'intera società.

**Ciao Elena, secondo te cosa si intende per repository pubblico di dati? deve necessariamente essere OPEN o i dati in esso contenuti possono essere aperti su richiesta?
quindi repository pubblico=repository OPEN?**

I repositories open sono tutti liberamente accessibili, come lo sono tutti i record contenuti, a livello di metadati (dati bibliografici su autore, titolo del dataset, provenienza...).

L'elenco si trova in Re3data (<https://www.re3data.org/>)

Quanto ai singoli dataset depositati, possono anche essere ad accesso riservato (cioè chiusi) per ragioni legate a sicurezza, privacy ecc. Ricordiamo però che la A fi FAIR significa Accessible, non Open: ovvero, devo sapere dove poter trovare i dataset, e qualora fossero chiusi, a chi chiedere autorizzazione per l'accesso.

Buongiorno vorrei chiedere un chiarimento sulla proprietà dei dati: " i dati della mia ricerca non sono miei". Di chi è allora la proprietà dei dati? dell'ente che ha finanziato la mia ricerca (pubblico o privato che sia) ?

I dati grezzi non sono di nessuno (né singolo né ente) in quanto non coperti da copyright/diritto d'autore. Il diritto d'autore infatti tutela la particolare forma espressiva di un'idea. Le informazioni grezze non sono coperte da diritto d'autore.

Su questo potete vedere il contributo di Simone Aliprandi

(<https://aliprandi.org/books/fenomeno-opendata/> e anche

<https://www.slideshare.net/simonealiprandi/il-licensing-di-dati-e-le-principali-licenze-open-data>) o il factsheet di Creative Commons (<https://doi.org/10.5281/zenodo.840651>)

E' il responsabile del progetto che deve ottemperare agli obblighi di H2020?

No, è responsabilità di tutti i partners perché tutti hanno firmato il Grant Agreement in cui sono contenute le norme "open"

I dati non sono di nessuno, ma è altamente probabile che comunque chi li ha prodotti/raccolti li organizzi in una qualche forma. In tal caso non scatta il diritto sui generis?

Il diritto sui generis (20 anni invece dei 70 del diritto d'autore) si applica ai databases (banche dati) strutturate, non ai semplici datasets.

L'articolo di Lusoli è molto chiaro:

<https://www.linkedin.com/pulse/repeat-me-research-data-mine-wainer-lusoli>

Se i miei dati sono tracce audio/video/trascrizioni di interviste sono considerati dati grezzi? e se gli intervistati hanno firmato un consenso informato, non sono quindi io/ente proprietario di quei dati?

No, io sono il "responsabile del trattamento" di quei dati, non il proprietario.

Su casi specifici possiamo sentire i data stewards di Delft, o le colleghe di DANS olandese, che lavorano quotidianamente su casi concreti. Difficile dare una risposta univoca sui dati: ogni dataset è unico e ha esigenze particolare. Per questo servono i datastewards, che aiutano nella corretta gestione dei dati nelle singole discipline.

Una buona base è il Data Wizard (<https://dsw.fairdata.solutions/>) segnalato nelle slides, che pone domande concrete tipo "stai utilizzando dati preesistenti" e apre il capitolo del libro di Barend Mons relativo all'argomento in oggetto. Mi pare che ci sia anche la questione delle liberatorie.

In fase di controllo su metadati e set di dati all'interno di un repository per la pubblicazione non è possibile capire se i dati sono grezzi o meno. Come stabiliamo se va applicata un CC0 o altra licenza? [...]

Intanto la licenza va applicata dall'autore, quindi chi meglio di lui/lei sa se i dati sono grezzi o no?

In ogni caso, con la licenza CC 0 non si sbaglia mai (vedete anche Thomas Margoni alla Open science Fair di Atene:

<https://www.slideshare.net/OSFair/osfair2017-training-open-science-check-list-for-repositorie-s-and-publishers>)

Bisogna ricordare che l'uso di una CC0 (dedicazione al pubblico dominio) Non esime dal corretto comportamento accademico, ossia citare la fonte, in questo caso chi ha messo insieme il dataset.

Sul sito del DCC c'è anche una parte curata dall'ottimo Alex Ball su come licenziare i dati:
<http://www.dcc.ac.uk/resources/how-guides/license-research-data>

[...] sì, tra poco apriamo un repository di dati. Non avevano proprio considerato la cc0 di default

Per i dati è quella più appropriata ed è perfettamente in linea con le richieste di Horizon2020 che chiede la licenza più aperta possibile per il riuso.

Ok, quindi in teoria gli editori dovrebbero mettere open i supplementary materials?

Questo succede già: i supplementary materials sono sempre a disposizione liberamente anche qualora l'articolo sia chiuso dietro abbonamento.

Nel caso di un software sviluppato quando si è membri di un ente di ricerca, si può andare contro il diritto dell'ente se si rende pubblico il codice?

Sul software non sono in grado di rispondere.

Consiglierei di nuovo Simone Aliprandi.

<http://aliprandi.blogspot.com/2016/04/contratti-software-ingegneri-modena.html> e anche <https://www.slideshare.net/simonealiprandi/conoscere-e-gestire-le-licenze-open-source-marzo-2017>

Leggere l'elenco delle eccezioni che consentono di non rendere accessibili i dati (benché depositati in un database interpellabile) mi fa nascere una domanda: l'impresa vale la spesa?

Assolutamente sì!!! Le eccezioni sono appunto eccezioni, ricordiamoci che il principio è "as open as possible, as closed as necessary". Quelle contenute nel Grant Agreement sono quasi tautologiche e riguardano sempre sicurezza, privacy ecc ecc (motivi per cui, a ben considerare, non si pubblica nemmeno...).

In ogni caso le ragioni di opt out vanno discusse e argomentate con il project officer, quindi occorre avere motivazioni ben solide.

E ci sono mille mila vantaggi nel rendere aperti i dati...

Grazie ancora a tutti, e se avete bisogno scrivete.

Elena e Emma

noad-it@openaire.eu

<https://rd-alliance.org/group/rda-italy/wiki/open-science-webinar-series-2018>