

Webinar "Covid-19 e condivisione dei dati: perché in Italia si fa troppo poco?", seconda edizione, 16 novembre 2020



Open Science e Covid-19 Collaborare per contrastare la pandemia

WEBINAR:

Covid-19 e condivisione dei dati: perché in Italia
si fa troppo poco?

16 novembre 2020 | ore 15:00

This event was partially funded by the European Commission under the H2020 programme. ELIXIR-CONVERGE: GA 871075
EOSC-Life - GA 824087 | EOSC-Pillar: GA 857650 | OpenAIRE Advance: GA 777541 | RDA Europe 4.0: GA 777388



La pagine dell'evento sul sito CNR:

<https://www.cnr.it/it/evento/16992/open-science-e-covid-19-collaborare-per-contrastare-la-pandemia>

La pagina sul sito di OpenAIRE:

<https://www.openaire.eu/item/open-science-e-covid-19-collaborare-per-contrastare-la-pandemia>

Domande e risposte

Autori: Emma Lazzeri, Allegra Via, Graziano Pesole, Sara Casati, Gina Pavone, Francesca De Leo, Loredana Lepera.

- Zenodo o OpenAIRE sono siti internet?

Sono due strumenti finanziati dalla Commissione europea per facilitare l'implementazione dell'Open Science. Hanno entrambi i loro siti internet e sono due realtà diverse ma collegate e che collaborano.

[OpenAIRE](#) è, brevemente, una infrastruttura che mette a disposizione una serie di servizi per l'OS e l'Open Access. Per esempio [nella sezione "Explore" del sito](#), può navigare tra milioni di pubblicazioni, di cui sono raccolti i metadati e di cui potrà accedere alla versione Open Access e conoscere gli autori, le istituzioni e i progetti in cui è stata prodotta, l'ente finanziatore e il funding stream con cui è stato sostenuto. Per facilitare la "discoverability" dei contenuti si può cercare con alcune opzioni avanzate (per esempio per tipo di prodotto, per progetto eccetera). Le informazioni che si trovano sul portale di OpenAIRE e su cui si

basano i servizi offerti, sono raccolte in un grafo che deriva dai contenuti di più di 17.000 repository e altri content provider di tutto il mondo.

[Zenodo](#) è un repository Open Access, dunque un archivio digitale online dove si può depositare e conservare un qualsiasi prodotto di ricerca (pubblicazioni, dati, presentazioni, report, poster, deliverable di progetto e altro ancora). Zenodo è un repository generalista, dunque non è specifico per disciplina e non è rivolto a una comunità in particolare. Di solito si consiglia sempre di verificare l'esistenza di repository disciplinari adatti al proprio lavoro (in modo che sia più facile entrare in contatto con la comunità di riferimento). E di depositare presso i propri repository istituzionali, qualora esistano. *Su OpenAIRE e i suoi servizi si veda anche la domanda sotto.*

Si consiglia invece di ricorrere a Zenodo per adempiere ai mandati OA con modalità "green OA" - cioè autoarchiviando il proprio lavoro, facendo attenzione a rispettare i tempi di embargo richiesti dagli editori da una parte, ma al contempo anche gli eventuali obblighi di enti finanziatori.

Inoltre Zenodo va bene per i ricercatori che non dispongono di un repository disciplinare di riferimento o istituzionale.

- **L'attendibilità dei test diagnostici basati sulla RT-qPCR e/o sierologici, il COI etc. Il quadro è piuttosto confuso ma iniziano ad essere pubblicati lavori su questo argomento (Nuccetelli et al Cell Death Discovery 2020, 6, 38, <https://doi.org/10.1038/s41420-020-0275-2>)**

Questo aspetto verrà approfondito in un webinar successivo.

- **Come si potrebbe riuscire a raccogliere i dati dagli ospedali?**

È molto importante considerare come si originano i dati clinici. Innanzi tutto qualsiasi dato, anche al di là di quelli relativi alla pandemia, deve avere il consenso informato alla condivisione per scopi di ricerca, rilasciato dal cittadino o cittadina al momento delle cure. Per facilitare la condivisione dei dati andrebbe inoltre tenuta in considerazione la comunità dei clinici, altrimenti possono crearsi degli ostacoli, anche involontari.

Più in generale vanno considerati per bene tutti i passaggi che portano alla condivisione dei dati e che tutti gli attori siano considerati. Per esempio un tema importante affrontato nello specifico del biobancaggio, è quello della co-authorship, cioè di come riconoscere il contributo di tutti coloro che partecipano alla "filiera". Un altro passaggio potrebbe essere quello di armonizzare i processi di valutazione nella condivisione dei dati: per esempio in alcuni paesi nordici si sta cercando di promuovere tra i comitati etici la "RECognition", cioè una volta ottenuta la valutazione di un progetto da parte del comitato etico competente, ci si porta dietro quella valutazione su quel progetto.

Si sta inoltre cercando di arrivare a definire la linea di indirizzo per il riutilizzo per scopi scientifici dei dati che si raccolgono inizialmente per scopi clinici. La cosiddetta legge Lorenzin era accompagnata da un decreto attuativo del 2019 secondo cui l'Istituto Superiore

di Sanità con il supporto di BBMRI.it, sentiti il Garante è il Centro di Coordinamento Nazionale dei Comitati Etici Territoriali, avrebbe dovuto scrivere la linea di indirizzo. La tempistica prevista è stata stravolta dalla pandemia, tuttavia si prevede la finalizzazione del documento entro pochi mesi.

Per quanto riguarda le infrastrutture digitali necessarie, si sta lavorando a un modello evoluto di sviluppo negli ospedali, primi tra tutti gli IRCCS, che sia anche integrabile a livello di sistema paese.

- **Gli svedesi hanno il loro portale per la condivisione dei dati sanitari, noi no. Esiste una procedura europea per mettere a disposizione, supponiamo, i dati dei tamponi? Se sì, e avessimo contatti con dirigenze ASL potremmo diffondere qualche materiale in particolare?**

No, per il momento in Italia non esiste. Il portale svedese si occupa perlopiù di dati molecolari (sequenze, proteomica, etc...) ma anche nella parte dei dati sanitari, non c'è nulla sulla condivisione dei materiali dei tamponi.

- **Difficile ottenere dati omogenei dalle varie strutture ospedaliere. Come si può superare questo problema?**

È evidente una frammentazione a monte, per esempio tutta questa materia è di competenza del ministero della salute e del ministero della ricerca. Sarebbe auspicabile uno sforzo "top-down" - anche se condiviso a vari livelli - nell'impostazione di un lavoro di integrazione. In altre parole l'integrazione dei dati non può essere fatta su iniziative sporadiche o spontanee: va organizzata.

- **Su quale storage di quale sistema di calcolo italiano immagazzinare, indipendentemente dal metodo, i dati raccolti?**

In Italia le istituzioni di ricerca offrono spesso soluzioni di storage dedicate ai propri ricercatori.

Il GARR offre soluzioni Cloud personalizzabili per istituzioni e singoli ricercatori (<https://www.garr.it/it/infrastrutture/infrastruttura-cloud/infrastruttura-cloud>), sia nella classica modalità IaaS (macchine virtuali e virtual data centre) che PaaS, disponibili in prova senza impegno.

All'interno di ICDI (Italian Computing and Data Infrastructure), il forum nazionale delle infrastrutture di ricerca, esiste una task force che sta lavorando ad una soluzione cloud federata su scala nazionale per offrire:

- la condivisione di risorse di calcolo
- il data sharing
- la federazione delle identità

maggiori informazioni sulle attività della task force ICDI consultate [questa pagina](#).

- **Potete dirci di più sugli strumenti di biobancaggio? Quali sono? Con che modalità? (solo "commerciale" oppure anche "istituzionali")?**

BBMRI è l'infrastruttura di ricerca europea per il biobancaggio e raccoglie a livello europeo 676 biobanche partner. La biobanca per definizione è una struttura di servizio senza scopo di lucro, costituita in un'istituzione pubblica o privata che nel pieno dei diritti dei soggetti coinvolge, garantisce e gestisce l'intero processo del biobancaggio, dal prelievo all'accesso secondo standard di qualità internazionali

Con la pandemia è diventato evidente quanto sia importante conservare in qualità i campioni biologici e con i requisiti etico-legali determinanti per un loro utilizzo scientifico nel tempo per questo è stato messo a punto un consenso informato semplificato per la ricerca Covid, e reso noto anche al garante.

La situazione italiana presenta alcuni problemi che denotano come la cultura del biobancaggio di ricerca non sia fino in fondo patrimonio di tutti, anche se sono in atto importanti azioni di sistema affinché la distribuzione per scopi di ricerca dei campioni sia dentro un iter istituzionale tracciato e documentato dal Material/Data Transfer Agreement, in cui la biobanca ha una funzione di terzietà e di garanzia fondamentale. Non casualmente, molte biobanche sono state in prima linea ed operative raccogliendo campioni e dati per la ricerca COVID, rendendo pubblica la loro disponibilità direttamente, nel catalogo italiano come in quello europeo di BBMRI, ove c'è una sezione COVID dedicata, dove non sono pochi i campioni, anche italiani, che sono stati resi disponibili.

Per la raccolta di campioni e dati COVID-19 in qualità ci sono strumenti [ELSI \(che affrontano le questioni etico-legali-sociali in gioco\)](#), sviluppati dal "common service ELSI" del nodo nazionale di BBMRI: sono fruibili direttamente alla sezione del sito <https://www.bbmri.it/nodo-nazionale/elsi-covid/>. Questi strumenti di buona pratica (dal consenso informato al MTA) sono frutto di gruppi di lavoro nazionali ELSI inclusivi di tutti gli attori coinvolti nel biobancaggio di ricerca. Ogni anno il nodo nazionale BBMRI.it propone gruppi di lavoro sulle tematiche cogenti e aperte, al fine di sviluppare linee di indirizzo e strumenti di buona pratica comuni; per partecipare le porte sono aperte ed è molto semplice: direttamente sul sito ci si può iscrivere alla Newsletter o ai gruppi di lavoro, quando le call di partecipazione sono aperte. Per quanto riguarda problematiche concrete, sono operativi un Helpdesk per le questioni Etico-Legali-Sociali scrivendo a elsi@bbmri.it, così come per le questioni di qualità scrivendo a italy@bbmri.it. Per saperne di più invece dei servizi europei e fruire del catalogo internazionale il sito è www.bbmri-eric.eu

[A questo link ci sono le raccomandazioni dell'ISS per raccolta, trasporto e conservazione di campioni biologici COVID-19 \(versione 15 aprile 2020\).](#)

- Sono un data scientist, posso fare il volontario?

Forse conosce [ICDI](#) - Italian Computing and Data Infrastructure, un "tavolo di lavoro" a cui siedono diverse infrastrutture di ricerca e infrastrutture digitali italiane, con l'obiettivo di coordinare la partecipazione italiana alle iniziative rilevanti nell'ambito della scienza aperta, tra cui per esempio l'EOSC. All'interno del [Competence Center di ICDI](#) è attivo una Task Force di esperti. Chiunque sia interessato può controllare se la propria istituzione fa parte di ICDI: se sì può entrare nella TF; in caso contrario invece ci può contattare per ulteriori informazioni (noad-it@openaire.eu).

- **Esiste un metodo per evidenziare gli articoli con risultati in contraddizione in OpenAIRE? Esiste un sistema di peer review trasparente in OpenAIRE?**

OpenAIRE non è una piattaforma di pubblicazione e non si occupa di valutazione della ricerca in senso stretto, per cui non dispone di informazioni sulla peer review dei contenuti né segnala contraddittorietà o altre “caratteristiche” riguardo al contenuto dei prodotti della ricerca.

OpenAIRE si occupa di raccogliere informazioni sui prodotti della ricerca, tra cui: autori del prodotto, organizzazione di appartenenza, progetto di ricerca per cui è stato realizzato, ente finanziatore e funding stream. Queste informazioni sono poi organizzate in un grafo, cioè vengono riprodotti in questo tipo di struttura le connessioni esistenti nella realtà tra i vari elementi coinvolti nella produzione scientifica (prodotti della ricerca, istituzioni, enti finanziatori, eccetera).

Questo grafo può essere usato dai vari attori in vario modo:

1. il ricercatore per cercare altri prodotti della ricerca (discoverability)
2. l'ente finanziatore per monitorare l'esito e i frutti delle risorse distribuite, e l'ottemperanza ai mandati su Open Science da parte dei beneficiari
3. il coordinatore di progetto per avere una misura dell'impatto prodotto dal suo progetto in termini di risultati.
4. l'ente di ricerca per avere delle statistiche a proposito della sua produzione e avere così “il polso” della situazione (per esempio potrebbe voler controllare quanta della sua produzione è depositata nel proprio repository istituzionale e quanta invece si trova all'esterno).

- **La proliferazione di riviste scientifiche (così come di convegni scientifici) è uno dei by-products dell'errato sistema di valutazione della ricerca. Se si riuscisse a cambiare il meccanismo di valutazione della ricerca probabilmente anche il numero di riviste e convegni diminuirebbe, e di conseguenza i soldi pubblici sarebbero meglio spesi. La mia domanda è: quali sono le proposte per modificare il sistema di valutazione della ricerca?**

Nella comunità al lavoro sui temi dell'Open Science e dell'Open Access è diffusa la consapevolezza della necessità di elaborare nuove modalità di valutazione della ricerca, e soprattutto di superare gli attuali metodi basati sugli indici bibliometrici quali impact factor e h-index. Una delle iniziative specializzate che lavora in questa direzione è [DORA \(Declaration on Research Assessment\)](#). Inoltre la Commissione Europea come ente finanziatore ha annunciato che con Horizon Europe, il nuovo programma di finanziamento che partirà da gennaio 2021, si dovranno soddisfare alcuni principi dell'Open Science per esempio si richiederà ai beneficiari dei fondi di non cedere il proprio diritto d'autore, in modo da poter soddisfare gli obblighi di Open Access; i dati prodotti dalla ricerca finanziata

dovranno essere FAIR e aperti “by default” (fatte salve le eccezioni, per esempio in vista di applicazioni commerciali) eccetera. [A questo link si può leggere qualcosa più in dettaglio.](#) Recentemente, DORA, EUA e SPARC EUROPE, hanno rilasciato una serie di case studies di istituzioni che hanno mosso i primi passi nel cambiamento della valutazione: <https://sfdora.org/dora-case-studies/>