

I dati della ricerca biomedica in Italia: verso la definizione di una *policy* nazionale?

Moreno Curti¹, Paola De Castro², Corrado Di Benedetto³, Rosalia Ferrara⁴, Pietro La Placa⁵, Cristina Mancini⁴, Luisa Minghetti⁶, Elisabetta Poltronieri², Filippo Santoro⁶, Franco Toni⁴, Angela Vullo⁷

¹ IRCCS, Fondazione Policlinico San Matteo, Pavia. Servizio documentazione e biblioteca scientifica

² Istituto Superiore di Sanità, Servizio conoscenza e comunicazione scientifica, Attività editoriali

³ Istituto Superiore di Sanità. Roma. Servizio controllo di gestione e informatica

⁴ Istituto Superiore di Sanità. Roma. Servizio conoscenza e comunicazione scientifica, Biblioteca

⁵ Istituto Zooprofilattico Sperimentale della Sicilia, Palermo. Servizio editoria e biblioteca

⁶ Istituto Superiore di Sanità, Roma. Servizio tecnico scientifico di coordinamento e supporto alla ricerca

⁷ Istituto Zooprofilattico Sperimentale della Sicilia, Palermo. Dipartimento sanità territoriale

Introduzione

La ricerca biomedica in Italia può contare dal 2003 sulla condivisione di risorse informative e sullo sviluppo di temi legati alla diffusione della documentazione scientifica assicurati dal sistema Bibliosan <http://www.bibliosan.it/> tramite la sua rete di biblioteche biomediche afferenti agli istituti di ricerca del Servizio Sanitario Nazionale (SSN).

Gli enti della rete Bibliosan sono 60 e includono l'Istituto Superiore di Sanità (ISS), 46 Istituti di ricovero e cura a carattere scientifico (IRCCS), 10 Istituti zooprofilattici (IZS), l'Agenzia nazionale per i servizi sanitari regionali (AGENAS), l'Agenzia italiana del farmaco (AIFA) e l'Istituto nazionale assicurazione contro gli infortuni sul lavoro (INAIL). Tali istituzioni sono tradizionalmente produttori e utilizzatori di dati della ricerca, acquisiti in forza del proprio mandato istituzionale e organizzati con diverse modalità per attuare gli obiettivi di tutela della salute e di analisi dello stato sanitario del Paese. Gli strumenti destinati a tali finalità sono principalmente:

- i registri regionali e nazionali. Sistemi di censimento sistematico di dati per monitorare e studiare fenomeni specifici, ad es. i casi di rischio per la salute legati al manifestarsi di una data patologia o alla tossicità chimica di una sostanza. Consistono in "statistiche per dati aggregati", cioè dati elaborati (non grezzi) condensati in *report* annuali. Alcuni registri sono istituiti con provvedimenti legislativi o ministeriali mentre altri sono registri 'spontanei' sorti, ad es., su iniziativa di un'istituzione per contribuire a chiarire quali fattori ambientali, genetici e stili di vita influenzino la salute psicofisica della popolazione;
- le basi di dati. Aggregazioni strutturate di dati di vario genere prodotte, ad es., nell'ambito di studi epidemiologici e di monitoraggio ambientale, condotti da un'istituzione e consultabili, a seconda dei casi, liberamente o con accesso riservato.

Accesso aperto ai risultati della ricerca

Tra gli enti del sistema Bibliosan, l'ISS ha coltivato storicamente una missione di apripista in adesione ai principi del movimento *Open Access* (OA), confluiti e ampliati nella galassia dell'*open science*, favorendo la comunicazione e lo scambio dei risultati della ricerca. Nel 2008 si è dotato di una "Politica istituzionale per il libero accesso alle pubblicazioni scientifiche" http://www.iss.it/binary/sae4/cont/policy_ISS.jpg con finalità di certificazione, libera disponibilità online e archiviazione digitale permanente delle pubblicazioni scientifiche prodotte dall'Ente e da istituzioni partner del SSN. Per l'attuazione di tali finalità, l'ISS ha istituito nel 2006 un archivio digitale in rete, DSpace ISS <http://dspace.iss.it/dspace/>, con la funzione di deposito istituzionale delle proprie pubblicazioni.

In adesione alle Raccomandazioni europee l'ISS si è attivato con iniziative e progetti nazionali e internazionali di promozione della cultura dell'accesso aperto: nel 2013 è tra i primi firmatari del *position paper* sull'accesso aperto ai risultati e ai dati della ricerca scientifica in Italia <http://www.cnr.it/sitocnr/Iservizi/Biblioteche/PositionAccessoAperto.html>, firmato dai principali enti di ricerca italiani (CNR, ENEA, INFN, ecc. e dalla CRUI) su iniziativa del progetto MedOAnet e partecipa in qualità di stakeholder della ricerca ad una consultazione della Commissione europea sugli *open data* http://ec.europa.eu/research/science-society/document_library/pdf_06/report_2013-07-open_research_data-consultation.pdf; nel 2014 firma la "Dichiarazione di Messina 2.0: la via italiana all'accesso aperto" http://decennale.unime.it/?page_id=2039 per confermare insieme agli atenei e agli enti di ricerca italiani il sostegno all'attuazione di politiche istituzionali per lo sviluppo dell'accesso aperto.

Verso una policy di gestione dei dati della ricerca

A giugno 2016, sotto il coordinamento dell'ISS è stato istituito, in seno al sistema Bibliosan, il Gruppo di lavoro BISA "Bibliosan per la scienza aperta", con l'obiettivo di elaborare un questionario sulle pratiche di gestione dei dati della ricerca in ambito biomedico in Italia. Un più ampio traguardo è costituito dall'elaborazione di una *policy* di gestione dei dati della ricerca, per allineare la comunità di ricerca italiana agli standard internazionali di condivisione dei dati.

Il questionario BISA si è articolato in 26 domande distribuite nelle seguenti sezioni tematiche: dati generali; tipologia e formato dei dati; archiviazione dei dati; aspetti etici e legali; accessibilità e riuso; infrastrutture e servizi; desiderata, idee. L'indagine, svoltasi nel periodo 31 gennaio – 5 marzo 2017, ha registrato 2433 risposte provenienti da 58 su 60 degli enti coinvolti, con un tasso di risposta pari al 15%, su una popolazione di riferimento di 15.000 unità.

Dall'indagine BISA sono emerse esplicite richieste che delineano una comunità di ricerca consapevole delle problematiche e delle aspettative legate alla gestione dei dati aperti della ricerca; il 96% delle risposte al quesito "Quale ulteriore azione ti aspetti dalla tua Istituzione?" evidenzia, infatti, una forte aspettativa dei ricercatori nei confronti della propria istituzione a fruire di linee guida o politiche di indirizzo per gestire i dati della ricerca e individuare, al contempo, professionalità *ad hoc*.

Entrambi i fattori, strutturali (strumentazioni e tecnologie) e gestionali (linee di indirizzo e coordinamento centrale), sono risultati i motori per attivare la cultura della condivisione dei singoli ricercatori, al riparo da pratiche burocratiche. In definitiva, l'orientamento è verso la richiesta di attuazione di una legislazione nazionale competitiva che superi i legalismi e promuova concretamente il trattamento dei dati, soprattutto nella casistica di studi di coorte, caso controllo, multicentrici e di sorveglianza epidemiologica che rappresentano il nerbo della ricerca biomedica. Una volta analizzate, le criticità di gestione dei dati della ricerca andrebbero eliminate attuando una regolamentazione che ne disciplini il trattamento e il riutilizzo. Un'ipotesi potrebbe essere quella di istituire un database dei dati generati dalle ricerche, ripartiti in settori tematici. Si attuerebbe così la catalogazione del patrimonio di dati raccolto onde favorirne il riutilizzo e l'adeguata valorizzazione. Attuando un'infrastruttura informatica deputata a ospitare i dati a livello nazionale, si offrirebbe la possibilità di massimizzare la circolazione dell'informazione scientifica, liberando così il potenziale della ricerca biomedica a beneficio della più vasta platea di ricercatori e scienziati.

Attualmente l'ISS sta elaborando una bozza di *policy* sulla gestione dei dati della ricerca, con l'obiettivo di recepire le aspettative dei ricercatori dell'Ente, di convalidare le buone pratiche in uso e di costituire una infrastruttura organica interna per la registrazione, la protezione e la condivisione dei dati. Si tratta di agevolare i responsabili di progetto aderenti ai bandi di ricerca Horizon 2020 e tenuti a depositare ad

accesso aperto, tramite un *Data Management Plan*, i cosiddetti “underlying data”, raccolti o generati durante la ricerca e necessari per validare i risultati presentati negli articoli.

In tale prospettiva, come Gruppo di lavoro BISA e in collaborazione con altre componenti dell’ISS (Settore Informatica e Servizio tecnico scientifico di coordinamento e supporto alla ricerca), si è proceduto ad esplicitare le tipologie di dati prodotti e utilizzati e a impostare le procedure di descrizione, registrazione e archiviazione dei dati, in conformità agli standard europei (principi FAIR). In tale fase si è posta particolare attenzione agli aspetti di responsabilità dei ricercatori e dell’Ente nella raccolta e nella gestione dei dati. Sono stati considerati inoltre gli eventuali vincoli derivanti da accordi con terze parti (es. protocolli d’intesa e brevetti). Questa iniziativa preliminare dell’ISS, volta a delineare i contenuti di una *policy* sui dati della ricerca, si è avvalsa della stesura approntata in sede CRUI da un Gruppo di lavoro *ad hoc* ed è stata integrata da elementi che riflettono la specificità della *mission* istituzionale dell’ISS. Nell’iter di definizione di tale documento si prevede di acquisire la valutazione da parte dei rappresentanti dei ricercatori ISS, per poi pervenire all’approvazione dei vertici dell’Ente, anche in raccordo con gli orientamenti e le aspettative del Ministero della Salute. Esaurita questa fase, si potranno accogliere, su una base di valutazione condivisa, le integrazioni al documento da parte delle comunità di ricerca degli enti Bibliosin. Sarà molto proficuo istituire una base di confronto con il mondo accademico e la comunità di ricerca internazionale per approdare a un documento inclusivo di tutti gli aspetti e le implicazioni del trattamento dei dati della ricerca.