

GENESI E PROSPETTIVE DEL GRUPPO DI LAVORO BISA (BIBLIOSAN PER LA SCIENZA APERTA)

Moreno Curti (a), Paola De Castro (b), Rosalia Ferrara (c), Cristina Mancini (c), Elisabetta Poltronieri (b), Franco Toni (c)

(a) *Direzione Scientifica, IRCCS, Fondazione Policlinico San Matteo, Pavia*

(b) *Servizio Conoscenza e Comunicazione Scientifica, Attività Editoriali, Istituto Superiore di Sanità, Roma*

(c) *Servizio Conoscenza e Comunicazione Scientifica, Biblioteca, Istituto Superiore di Sanità, Roma*

Nato nel 2003, il sistema Bibliosan (<http://www.bibliosan.it/>), dopo quasi 15 anni di continua attività dedicata alla cooperazione e alla costruzione di una rete tra le biblioteche biomediche degli Istituti di ricerca afferenti al Ministero della Salute, è diventato l'asse portante per la condivisione delle risorse informative e lo sviluppo di temi legati alla diffusione della documentazione scientifica biomedica nel nostro Paese.

Accanto alle ormai consolidate attività di condivisione e potenziamento delle risorse documentarie esistenti e di acquisto collettivo di risorse in formato elettronico (*e-journals*, banche dati biomediche e altro), uno dei principali impegni che Bibliosan si è assunto fin dalle origini è stato la costituzione al suo interno di una serie di gruppi di studio che assicurassero ai colleghi bibliotecari e ai ricercatori dei diversi enti il supporto su questioni e aspetti di particolare importanza e attualità connessi con la vasta galassia dell'informazione scientifica. Così, nel corso degli anni si è dato corso a gruppi di lavoro che hanno affrontato e dato risposte concrete a temi quali il problema copyright, gli abbonamenti elettronici consortili, la formazione e l'aggiornamento continuo, le banche dati, la comunicazione.

L'ultimo in ordine temporale di questi gruppi si è costituito nel 2016 per affrontare quello che nell'immediato appare come il tema di più stringente attualità e bisognoso di risposte concrete nel campo della ricerca in generale e di quella biomedica in particolare: il libero accesso all'informazione scientifica e il trattamento dei dati che costituiscono la base di ogni attività di ricerca. Sono ormai alcuni anni che questo argomento dai molteplici risvolti – l'utilizzo, la diffusione, l'accesso, la conservazione dei dati – è al centro dell'attenzione delle comunità scientifiche e dei governi che ne devono regolare l'uso. Si tratta, come facilmente comprensibile, di un compito di non semplice risoluzione e per il quale non sempre è possibile trovare risposte univoche e soddisfacenti. Su di un punto tuttavia sembra esserci un accordo generalizzato: la necessità di rendere i dati della ricerca il più aperti e disponibili possibile, un principio riaffermato con forza sia all'interno dei bandi Horizon 2020 dell'Unione Europea (UE), volti ad attuare disseminazione, riuso e riproducibilità dei risultati della ricerca finanziata dall'UE (1), sia dalle decisioni assunte dai governi (finora pochi) che hanno affrontato il problema.

L'Italia, come sovente purtroppo accade, sconta un ritardo anche in questo campo e solo grazie alle iniziative del mondo bibliotecario accademico e degli istituti di ricerca si è cominciato ora ad affrontare la questione e a sensibilizzare ricercatori e mondo politico sull'indifferibilità di trovare soluzioni soddisfacenti e valide. In questo scenario è sembrato imprescindibile che all'interno di Bibliosan si venisse a formare il Gruppo di lavoro BISA (Bibliosan per la Scienza Aperta), con il compito di promuovere i principi dell'*Open Science* e specificamente degli *Open Research Data* nell'ambito della ricerca biomedica. Il primo impegno assunto dai componenti – provenienti dall'Istituto Superiore di Sanità (ISS), dagli Istituti di Ricovero e Cura a Carattere Scientifico (IRCCS) S. Matteo di Pavia e Regina Elena di Roma e dall'Istituto Zooprofilattico Sperimentale

(IZS) della Sicilia – è stato quello di promuovere questa indagine ricognitiva per cercare di capire come i diretti interessati, i ricercatori, approccino il tema; quali siano i loro orientamenti; come concretamente cerchino di fornire risposte ai diversi problemi che incontrano nel difficile percorso del trattamento dei dati; quali comportamenti e soluzioni si aspetterebbero dal sistema Bibliosan. I risultati dell'indagine BISA, illustrati nel corso del Convegno "I dati aperti cemento della scienza: risultati dell'indagine Bibliosan per la scienza aperta", organizzato dall'ISS il 15 maggio 2017 (2) vengono descritti nello studio analitico a questo scopo condotto.

Del resto, strumenti conoscitivi diretti a sondare pratiche, attitudini e aspettative dei ricercatori circa la gestione dei dati generati all'interno delle proprie istituzioni sono stati predisposti da varie organizzazioni internazionali (3-4), per poi avviare una pianificazione organica di risorse e investimenti finalizzati a realizzare servizi di supporto al trattamento dei dati.

Il dibattito sui dati aperti della ricerca coinvolge attivamente anche la comunità dei bibliotecari e degli esperti dell'informazione biomedica a livello internazionale, come dimostrano i recenti contributi pubblicati a più riprese, tra 2016 e 2017, attraverso canali di informazione professionale (5).

L'esito del questionario BISA, illustrato anche in sede internazionale in occasione della conferenza annuale dell'EAHIL (*European Association of Health Information and Libraries*) svoltasi a Dublino nel giugno 2017 (6), costituisce l'elemento base per costruire un principio di conoscenza e un'opera di sensibilizzazione degli organi politici. Solo così si potrà auspicabilmente dare luogo all'adozione di una *policy* a livello nazionale, già realizzata nei più avanzati stati membri dell'UE e che invece, in Italia, è ancora in fase di incubazione. È un lavoro che va portato avanti unitamente ai colleghi esperti di *open access* dei gruppi di lavoro della CRUI (Conferenza dei Rettori delle Università Italiane) che agiscono nell'ambito del MIUR (Ministero Ministero dell'Istruzione, dell'Università e della Ricerca), di per sé già attivo nell'elaborazione di un modello di *best practice* con un Gruppo di lavoro *ad hoc* (7), e che deve prevedere il coinvolgimento degli organi del Ministero della Salute che si occupano della ricerca e della gestione dei dati. Solo attraverso il conseguimento di una unità di intenti tra i due Ministeri principalmente interessati alle attività di ricerca sarà possibile pervenire alla definizione di una linea di azione unitaria a livello nazionale. Riuscire in questo intento deve essere il principale obiettivo a breve termine, sul quale incentrare impegno e attenzione per le prossime attività del Gruppo.

Una raccolta organica delle risorse documentali in tema di accesso, trattamento e politiche di indirizzo è disponibile sul portale *OA Italia/Risorse sugli open research data* (http://wikimedia.sp.unipi.it/index.php/OA_Italia/Risorse_sugli_open_research_data) che offre un panoramica articolata e aggiornata di contenuti con un ampio sguardo anche alle iniziative internazionali. A tale proposito, la Commissione Europea *in primis*, in riferimento a tutti i progetti Horizon 2020 che raccolgono o generano dati, individua nella definizione di un *Data Management Plan* (DMP) lo strumento principe per una gestione ottimale dei dati e ne impone il deposito, salvo deroghe per motivate ragioni (es. copertura brevettuale e salvaguardia di riservatezza e sicurezza dei dati), in conformità ai criteri standard FAIR (*Findable, Accessible, Interoperable, Reusable*) (8).

In definitiva, si pone la necessità di tracciare direttrici guida ispirate alle buone pratiche di gestione dei dati della ricerca, valorizzando così il patrimonio di produzione scientifica di ciascuna istituzione. A sostegno di tale orientamento, esistono già validi strumenti (es. modelli di *policy*) predisposti dalla comunità accademica e di ricerca, a livello nazionale e internazionale, per facilitare le procedure di raccolta e gestione dei dati, in base a criteri uniformi per garantirne la riproducibilità (9-12). Occorre ora integrare questi strumenti con una politica istituzionale che rifletta le specificità della ricerca biomedica e che risponda alle aspettative dei singoli enti operanti nel sistema Bibliosan.

Un peso rilevante nell'economia di una gestione consapevole dei flussi di dati della ricerca e di una regolamentazione condivisa, da attuare attraverso un documento di indirizzo istituzionale, assume la questione dei dati personali. Nella ricerca biomedica gli aspetti della *privacy* sono intimamente legati al trattamento dei dati sensibili e dovranno quindi essere analizzati attentamente alla luce della nuova normativa in materia. Si tratta di valutare le implicazioni, circa l'identificazione, l'utilizzo, la tutela della sicurezza, i servizi al cittadino, del nuovo Regolamento europeo sulla protezione, sul trattamento e sulla circolazione dei dati personali, emanato il 27 aprile 2016, cui gli stati membri sono tenuti ad adeguarsi entro il 25 maggio 2018 (13).

Senza dubbio, spostandoci su un altro fronte, le molteplici attività connesse con il trattamento digitale dei dati della ricerca, inquadrato nel più vasto ambito dei *big data*, implicano l'affermazione di nuove figure professionali di supporto al mondo accademico e della ricerca, quali, ad esempio, *data scientist*, *data modeler*, *data analyst*, *data architect* e *data engineer*. Tali attività rappresentano inoltre uno stimolo costante allo sviluppo e alla competitività del mercato del lavoro – con riflessi su pianificazione e investimenti nel campo dei sistemi informativi e del reclutamento del personale - nonché al progresso della società civile. In particolare, nel comparto *Information Technology* (IT), all'interno delle organizzazioni aziendali pubbliche e private, si stanno delineando profili professionali e relative iniziative di formazione che richiedono capacità, creatività e conoscenze innovative multidisciplinari e interrelate; un mix trasversale di competenze matematiche, statistiche, informatiche e di economia sta alimentando una cultura dinamica di impresa nell'era della cosiddetta *data science* (14), a sostegno dei decisori politici nell'organizzazione di adeguati flussi informativi interni che prevedono processi di estrazione, trasformazione e caricamento dei dati (15). Infatti, in un contesto di elementi integrati, le abilità tecniche riguardanti l'allestimento, la gestione e il controllo delle infrastrutture chiamano direttamente in causa i saperi legati alla comunicazione scientifica e alla gestione di progetti per la realizzazione di piattaforme istituzionali interoperabili.

Bibliografia

1. European Commission, Directorate-General for Research & Innovation. *H2020 Programme. Guidelines to the rules on open access to scientific publications and open access to research data in Horizon 2020*. Version 3.2, Brussels: European Commission; 2017. Disponibile all'indirizzo: http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf; ultima consultazione 17/08/2017.
2. *Le presentazioni dei relatori al Convegno: I dati aperti, cemento della scienza: risultati dell'indagine Bibliosani per la Scienza Aperta (BISA)*. Roma, 15 maggio 2017. Disponibile all'indirizzo: http://www.bibliosan.it/ftp/bisa_atti_15052017/bisa_15_05_2017.html; ultima consultazione 24/11/2017.
3. Bauer B, Ferus A, Gorraiz J, Gründhammer V, Gumpenberger C, Maly N, Mühlegger JM, Preza JL, Sánchez Solís B, Schmidt N, Steineder C. *Researchers and their data. Results of an Austria survey – Report 2015*. Vienna: University of Vienna; 2015. Disponibile all'indirizzo: <https://services.phaidra.univie.ac.at/api/object/o:409318/diss/Content/download>; ultima consultazione 07/12/2017.
4. European Council of Doctoral Candidates and Junior Researchers. *EURODOC Survey I: The first Eurodoc Survey on doctoral candidates in twelve European countries – Descriptive Report*. Brussels: Eurodoc; 2011. Disponibile all'indirizzo: http://eurodoc.net/oldwebsite/Eurodoc_survey_I_report_2011.pdf; ultima consultazione 22/11/2017.
5. Napolitani F. Open data: another step towards open science. *Journal of EAHIL* 2017;13(2):2.
6. Poltronieri E, Cognetti G, La Placa P, Vullo A, De Castro P, Ferrara R, Mancini C, Toni F. Let your open data smoothly settle in your publishing habits. In: *Diversity in Practice: integrating, inspiring*

- and innovative. *Book of abstracts of the 12th International Congress on Medical Librarianship and the 2017 EAHIL Workshop*. Dublin (Ireland), June 12-16, 2017. p. 20-1. Disponibile all'indirizzo: <http://eahil2017.net/wp-content/uploads/2017/06/Abstracts-ICML-EAHIL-2017.pdf>; ultima consultazione 22/11/2017.
7. Gruppo di lavoro del Ministero dell'Istruzione, dell'Università e della Ricerca. *Big Data@MIUR. Rapporto del gruppo di lavoro*. Roma: MIUR; 2016. Disponibile all'indirizzo: <http://www.istruzione.it/allegati/2016/bigdata.pdf>; ultima consultazione 31/8/2017.
 8. H2020 Programme. *Guidelines on FAIR Data Management in Horizon 2020*. Brussels: European Commission; 2016. Disponibile all'indirizzo: http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf; ultima consultazione:17/08/2017.
 9. GdL Dati della ricerca. *Modello di Policy sulla gestione dei dati della ricerca. Documento elaborato dal gruppo di lavoro informale sui dati della ricerca costituito da Politecnico di Milano, Università di Milano, Università di Torino, Università di Trento, Università di Venezia Ca' Foscari*. Milano: Università degli studi di Milano; 2017. Disponibile all'indirizzo: http://wikimedia.sp.unipi.it/images/RDMpolicyresearchdata27_03_2017.pdf; ultima consultazione 17/08/2017.
 10. University College of London, Information Services Division. *UCL Research Data & Network Services. UCL Research Data Policy*. London: UCL; 2013. Disponibile all'indirizzo: <http://www.ucl.ac.uk/isd/services/research-it/documents/uclresearchdatapolicy.pdf>; ultima consultazione:17/08/2017.
 11. EU-funded LEARN project. *LEARN Toolkit of Best Practice for Research Data Management*. 2017. Disponibile all'indirizzo: <http://learn-rdm.eu/wp-content/uploads/RDMToolkit.pdf>; ultima consultazione:17/08/2017.
 12. Harvard Library, Office for Scholarly Communication. *Model open access policy*. Cambridge, MS: University of Harvard; 2015. Disponibile all'indirizzo: https://osc.hul.harvard.edu/assets/files/model-policy-annotated_12_2015.pdf ; ultima consultazione:17/08/2017.
 13. Europa. Regolamento (UE) 2016/679 del Parlamento Europeo e del Consiglio del 27 aprile 2016 relativo alla protezione delle persone fisiche con riguardo al trattamento dei dati personali, nonché alla libera circolazione di tali dati e che abroga la direttiva 95/46/CE (regolamento generale sulla protezione dei dati). *Gazzetta ufficiale dell'Unione europea* L 119, 4 maggio 2016.
 14. Giacchino A. *Data Science: grandi opportunità di lavoro per vari profili professionali*. Pavia: BCI Italia; 2017. Disponibile all'indirizzo: <http://www.itware.com/blog-itware/big-data-analytics-data-management/item/1343-data-science-grandi-opportunita-di-lavoro-per-vari-profil-professionali>; ultima consultazione 16/08/2017.
 15. ADN Kronos, Economia. Cercasi professionisti del digitale, ecco le 10 figure più richieste. Roma: ADN Kronos; 2017. Disponibile all'indirizzo: http://www.adnkronos.com/soldi/economia/2017/05/07/cercasi-professionisti-del-digitale-ecco-figure-piu-richieste_c3ZxzHoABmn6R3zLsc63XO.html; ultima consultazione: 16/08/2017.