

Gestire, conservare, trasmettere i dati della scienza nei repository biomedici

Un simposio a Monza in occasione del 37° di fondazione di GIDIF-RBM

DOI: 10.3302/0392-8586-202301-062-1

Librarians can (and ought to) help create a navigable worldwide ocean of Knowledge, open to all.

Jean-Claude Guéron

Introduzione

Come viene gestita la conoscenza prodotta dalle organizzazioni scientifiche e dagli operatori del settore della sanità: ricercatori, documentalisti e bibliotecari biomedici? È questo il quesito all'origine del simposio organizzato dal GIDIF-RBM l'8 luglio 2022 e tenutosi, in presenza, a Monza. GIDIF-RBM, l'associazione italiana di bibliotecari e documentalisti biomedici, ha aperto un dibattito tra i maggiori attori sul tema per cercare di delineare lo stato dell'arte e dare, attraverso alcuni esempi ed esperienze pratiche, delle possibili soluzioni.

Il tema dei repository disciplinari o tematici – ovvero gli archivi digitali aperti che rendono accessibili pubblicazioni scientifiche, divulgative, comunicazioni, tesi e dataset prodotti dalle istituzioni – è largamente trattato in letteratura sin dai

primi anni 2000 [1]-[8], a partire dal movimento dell'Open access di cui vengono poste le basi tecnologiche già nel 1999 con la Open Archives Initiative [9], anche se, come ben descrive Iryna Solodovnik nel suo lavoro di dottorato [10], l'esigenza di condividere e rendere pubblicamente accessibili i prodotti della ricerca era molto sentita dagli studiosi già da parecchi anni prima.

Come sottolineato nell'introduzione del simposio da Fabio Venu-da, professore di biblioteconomia all'Università degli studi di Milano, il movimento di Open access sorge per diffondere i risultati della ricerca scientifica e contrastare le concentrazioni editoriali che misero la maggioranza dei periodici accademici nelle mani di una ristretta cerchia di editori. Gli editori riuscirono a imporre prezzi di abbonamento alle riviste e quindi alla produzione scientifica sempre più alti. Ridi, nella sua presentazione al libro di De Robbio [11], illustra come un paradosso il fatto per cui i ricercatori fossero disposti a cedere i propri diritti d'autore gratuitamente in cambio di una forma di

pubblicazione col maggiore impatto citazionale. Questo per favorire una rapida divulgazione dei risultati della propria ricerca su giornali autorevoli e poter migliorare la propria posizione professionale. Il risultato, sempre secondo quanto affermato da Ridi, fu che le università cominciarono a pagare due volte i risultati della propria ricerca e vedevano ridotto il loro stesso accesso all'informazione e alla letteratura scientifica per il costo degli abbonamenti. Da qui l'avvio alle pubblicazioni Open access e ad altre forme di editoria sostenibile, ma soprattutto alla creazione di Open Archives. Con il passare del tempo l'attenzione si è spostata poi non solo alle pubblicazioni vere e proprie, ma anche a tutta quella produzione di dati "grezzi" che, se messi a disposizione della comunità scientifica con le dovute attenzioni, possono essere riutilizzati per il bene pubblico.

L'innovazione e la ricerca rappresentano nondimeno una delle missioni delle organizzazioni afferenti al Ministero della Salute. Ma il tema di come gestire e diffondere i prodotti dell'attività e della ricerca non ha ancora una sua definizione nel mondo sanitario in Italia. Nella comunità degli Enti afferenti al Ministero della Salute non vi sono state né vi sono ancora indicazioni cui attenersi diversamente da quanto avvenuto nel mondo accademico in cui la CRUI ha pubblicato già nel 2009 le *Linee guida per gli archivi istituzionali*, poi perfezionate nel 2012 [12]. In mancanza di indicazioni uniformi cui attenersi, vi sono quindi state alcune iniziative a livello nazionale, come quella dell'Istituto superiore di sanità [13], e di singoli istituti [14], che non hanno però avuto continuità. Altre si stanno avviando.

Non devono essere ignorati o trascurati, inoltre, gli aspetti giuridici che regolano la proprietà intellet-

tuale, il diritto d'autore e non ultima la privacy: per la delicatezza dei dati trattati nel mondo biomedico non sono facili da governare.

L'obiettivo perseguito da GIDIF-RBM nella prima parte dei lavori è stato quello di offrire ai partecipanti un contributo conoscitivo da utilizzare negli ambiti decisionali rendendo la funzione del documentalista e bibliotecario biomedico mediatore culturale attivo, attore centrale della domanda di contenuti della comunità scientifica nelle biblioteche e negli archivi digitali di ultima generazione.

Stato dell'arte dei repository dei dati della scienza

Il primo contributo offerto da Simona Turbanti (Dipartimento di Studi storici - UNIMI) ha ricostruito lo stato dell'arte dei repository dei dati della scienza, cercando di individuare i problemi ancora aperti nella gestione dell'archiviazione dei dati prodotti dalla scienza. Simona Turbanti [6] ha preso in considerazione nella sua presentazione il delicato tema della valutazione della produzione scientifica ampliando lo sguardo alle metriche alternative e agli indicatori bibliometrici non convenzionali. Le metriche alternative sono strumenti di sicura utilità se impiegati in modo complementare alle metriche tradizionali. Al problema della valutazione della ricerca si aggiunge la domanda di visibilità e di conservazione delle informazioni depositate negli archivi digitali. Simona Turbanti ha presentato IRIS, il progetto sviluppato da CINECA per le università.

Il nucleo fondamentale della soluzione IRIS è il Repository dei prodotti della ricerca [15] quindi, sia per la gestione in ambito universitario, sia per tutti gli Enti di Ricerca pubblici o privati. IRIS è un

sistema qualificato nel Catalogo dei servizi Cloud per la Pubbliche Amministrazioni (Cloud Marketplace AgID che consente l'archiviazione, la consultazione e la valorizzazione dei prodotti scaturiti dalle attività di ricerca. È un sistema unico e integrato con le altre soluzioni CINECA, in grado di dialogare con i sistemi centrali nazionali e internazionali per la gestione e la disseminazione delle pubblicazioni, conforme ai requisiti del MIUR e della Commissione europea per l'Open access. Il repository è basato sulla piattaforma tecnologica internazionale DSpace. Il sistema è integrato con i più importanti provider di metadati editoriali e informazioni bibliometriche internazionali (Web Of Science, Scopus, CrossRef, PubMed). Questo sistema integrato consente di fornire altri servizi al ricercatore che vanno oltre il semplice deposito d'informazione, consentendo alle biblioteche di ampliare i propri servizi. Le competenze trasversali dei bibliotecari diventano un punto chiave per la disseminazione e conservazione dei dati scientifici. Simona Turbanti ha concluso la sua presentazione con una citazione di Jean-Claude Guéron [16]: "Librarians can (and ought to) help create a navigable worldwide ocean of Knowledge, open to all".

Repository dei dati della ricerca. FAIR compliance, infrastrutture e tecnologie

È il titolo della presentazione di Susanna Mornati, Direttore di 4Science. Dopo aver illustrato le caratteristiche dell'azienda per cui lavora, Susanna Mornati ha spiegato come ogni anno grandi quantità di dati prodotti dalla ricerca e dalla sperimentazione vanno perduti. A partire da questo dato di fatto, un team di scienziati canadesi ha studiato la probabilità che un dataset di ricer-

ca continui a essere reperibile negli anni successivi alla pubblicazione, per scoprire che tale probabilità diminuisce del 17% all'anno [17]. Si è quindi evidenziata la necessità – ha sostenuto Mornati – di creare un ambiente sicuro per il deposito dei dati che garantisca: a) fiducia rispetto alla provenienza dei contenuti; b) conservazione e accessibilità nel lungo termine; c) strumenti per la ricercabilità dei dati e la facilità di reperimento (metadati e protocolli standard) e per l'identificazione univoca della risorsa (Persistent Identifiers PIDs); d) facilità di condivisione dei dati e possibilità di riuso per la scienza futura.

Le soluzioni, secondo Mornati, partono dalla considerazione che le istituzioni che si occupano di ricerca dovrebbero dare supporto ai propri utenti, ad esempio tramite i servizi di biblioteca e/o documentazione:

- identificando e coordinando la fornitura e l'accesso a servizi e risorse per i dati di ricerca, tra cui: istruzione e formazione; linee guida e best practice; servizi interni o esternalizzati per la cura e la gestione dei dati;
- fungendo da canale di comunicazione formale per i requisiti e le attività del ciclo di vita dei dati di ricerca, compresa la promozione di eventi, servizi e risorse, e delle politiche istituzionali relative ai dati di ricerca.

Nell'identificare una soluzione appropriata che si rifaccia ai principi FAIR (Findable, Accessible, Interoperable, Reusable), Mornati ha illustrato alcuni progetti: the Dataverse Project [18], un'applicazione web ad accesso aperto per condividere, preservare, citare, esplorare e analizzare i dati di ricerca; il progetto GO FAIR [20], un'iniziativa dal basso verso l'alto, guidata dagli stakeholder e autogestita, che mira ad attuare i principi dei dati FAIR,

rendendo i dati reperibili, accessibili, interoperabili e riutilizzabili. GOFAIR offre un ecosistema aperto e inclusivo per individui, istituzioni e organizzazioni che lavorano insieme attraverso le reti di implementazione. Fare in modo che i principi FAIR siano una realtà è il loro scopo. Passando alle soluzioni tecnologiche, Mornati si è soffermata sul progetto Dataverse, piattaforma software gratuita e open-source per implementare un repository di dati, nella quale è possibile depositare, condividere, conservare, citare, esplorare e analizzare i dati della ricerca. Nasce dall'Institute for Quantitative Social Science (IQSS) di Harvard [19] con l'obiettivo di facilitare la condivisione e il riuso dei dati, l'attribuzione di crediti accademici e la visibilità sul web di ricercatori, istituzioni, editori. Di particolare interesse sono le caratteristiche del software, che fornisce una grande quantità di servizi utili: branding; internazionalizzazione (multilingual); il protocollo OAI-PMH per l'interoperabilità, che espone i metadati in formati standard quali DC Dublin Core, DDI Data Documentation Initiative (standard internazionale per descrivere indagini, questionari, file di dati statistici e informazioni a livello di studio sulle scienze sociali), OpenAIRE, ecc. Fornisce, inoltre, diverse API per la ricerca, il deposito, l'accesso ai dati, le metriche; librerie client per interagire tramite altri linguaggi di programmazione; integrazione con Datacite per l'attribuzione del DOI e con diversi sistemi di autenticazione/autorizzazione (SSO, OAuth2, OIDC, ORCID) e con strumenti esterni quali il Data Curation Tool. Offre altri servizi utili quali: il caricamento di file gerarchici e ausiliari; la preview di file testuali e multimediali; ricerca e analisi di file tabulari; possibilità di attivare workflow di revisione e personalizzare gli stati di *curation versioning* dei

dataset; integrazione con vocabolari controllati; possibilità di definire regole di accesso: ristretto, embargo, licenze CC o custom; libro degli ospiti; controllo di integrità dei file; citazione in diversi formati (End-Note XML, RIS, BibTeX) e molti altri servizi utili. Soprattutto, sostiene Mornati, Dataverse è davvero FAIR *compliant*.

Dalle tecnologie Mornati passa a illustrare alcune infrastrutture di condivisione. Fra queste:

- l'EOSC (European Open Science Cloud) [21], un ambiente per i dati di ricerca a sostegno della scienza dell'Unione europea che fornisce a ricercatori, aziende e cittadini europei un ambiente multidisciplinare, federato e aperto in cui pubblicare, trovare e riutilizzare dati, strumenti e servizi per scopi di ricerca, innovazione e istruzione. L'EOSC mira a sviluppare un web di dati e servizi FAIR per la scienza in Europa su cui costruire servizi a valore aggiunto, dalla visualizzazione e analisi alla conservazione a lungo termine delle informazioni o al monitoraggio dell'adozione di pratiche di scienza aperta. Nel sito EOSC, filtrando per "Scienze mediche e della salute", si possono consultare i repository tematici e istituzionali attivi in questo campo.
- OpenAIRE [22], una partnership non profit, derivante da un progetto europeo e costituitasi come entità legale per garantire un'infrastruttura permanente di comunicazione scientifica aperta a supporto della ricerca europea. È un'infrastruttura partecipativa, una comunità eterogenea presente in 37 Paesi europei, che fornisce linee guida, training e servizi a supporto della missione condivisa con EOSC. Fornire un accesso aperto, illimitato e senza barriere ai

risultati della ricerca finanziata con fondi pubblici in Europa. OpenAIRE mette a disposizione molta documentazione fra cui le *Linee guida per la gestione e implementazione degli archivi dei dati della ricerca* [23], uno strumento davvero utile per la comunità degli operatori.

Fra le altre risorse di OpenAIRE, Mornati cita OpenAIRE Graph che è una risorsa aperta che aggrega una raccolta di proprietà dei dati di ricerca (metadati, link) da oltre 70.000 fonti di comunicazione accademica di tutto il mondo, fra cui repository istituzionali ad accesso aperto, archivi di dati e riviste. Tutti i record di metadati sono riuniti in un data lake [24] insieme ai record di Crossref, Unpaywall, ORCID ecc. e alle informazioni sui progetti fornite da finanziatori nazionali e internazionali. Algoritmi di inferenza dedicati, applicati ai metadati e ai testi integrali delle pubblicazioni ad accesso aperto, arricchiscono il contenuto del data lake [24] con collegamenti tra risultati di ricerca e progetti, affiliazioni degli autori, classificazione dei soggetti, collegamenti a voci di database specifici. Le organizzazioni e i risultati duplicati vengono identificati e uniti per ottenere una risorsa pubblica, aperta e affidabile che consente di esplorare il panorama della comunicazione accademica.

Nel 2021 4Science ha rilasciato due servizi per l'utilizzo dell'OpenAIRE Research Graph nei repository: OpenAIRE ELD [25]-[26].

L'impegno di 4Science è quello di essere promotori sistemici dell'innovazione, dello scambio di conoscenze e della connessione fra culture.

Lo sviluppo open-source è un potente meccanismo di collaborazione per raggiungere questi obiettivi, in quanto favorisce la trasparenza e l'efficienza dei costi, quindi l'in-

novazione continua per migliorare gli strumenti collettivi di gestione delle informazioni e dei dati. L'auspicio è che anche noi bibliotecari e documentalisti si entri a far parte di questa comunità opensource. A questo proposito Mornati cita il gruppo di lavoro RDM, Research Data Management di LIBER [27], che nasce per aiutare le biblioteche a garantire un ruolo fondamentale nell'infrastruttura dell'informazione accademica del XXI secolo e per riconoscere i dati della ricerca come preziose risorse che devono essere gestite, condivise e conservate per promuovere la ricerca e la scienza.

L'esperienza del San Raffaele Open Research Data Repository

Marco Soriano e Marco Morelli, ricercatori con formazione diversa dell'Università Vita-Salute San Raffaele e dell'IRCSS Ospedale San Raffaele (UniSR), hanno presentato l'esperienza della loro realtà relativamente alla gestione e condivisione dei dati biomedici. La timeline inizia nel 2017, raccontano Soriano e Morelli, con la creazione di un Research Integrity Office (OSR), one-person, avente il compito di garantire la qualità, l'affidabilità e la riproducibilità della ricerca. L'obiettivo era quello di garantire elevati standard di integrità e qualità nella scienza e perseguire tutte le accuse fondate di cattiva condotta della ricerca. In che modo? Con attività educative per migliorare la presentazione e la riproducibilità dei dati scientifici; lo screening del manoscritto per evitare la possibilità di plagio; fornendo un punto di contatto per verificare l'integrità della ricerca e le questioni di cattiva condotta. L'impatto di tale ufficio è stato quello di far capire che la scienza ha qualche problema, ma

l'istituzione se ne prende cura garantendo supporto, monitoraggio e formazione invece che soltanto regole.

Nel 2018-2019 viene creato un gruppo di lavoro informale di Research Data Management in cui è coinvolta anche la biblioteca che si occupa della gestione della produzione scientifica, e si organizzano incontri su argomenti specifici. L'obiettivo del gruppo è di creare un novo repository per la gestione dei dati della scienza e se ne valutano le diverse opzioni e relativi pro e contro:

- Repository fatto-in-casa: comporta un notevole sforzo informatico.
- Repository open source (ad es. Zenodo): manca il controllo sui dati e sulla manodopera.
- Repository commerciale, chiavi in mano: ha un costo gestibile e richiede uno sforzo minimo.

È stata scelta la terza opzione e creato un team apposito – Open Science Team – per la sua gestione, di cui i due relatori fanno parte, che fa capo all'area Ricerca e sviluppo dell'UniSR. L'obiettivo dell'Open Science Team è quello di promuovere e sostenere l'adozione di Open Science e pratiche di gestione dei dati di ricerca (RDM).

Le sue attività principali sono:

- supporto dedicato su Open Science & RDM;
- formazione e tirocinio su Open Science & RDM;
- gestione del repository istituzionale per i dati della ricerca.

Per quanto riguarda l'attività di supporto, essa si concretizza nel mettere a disposizione risorse e strumenti per implementare la scienza aperta nell'attività di ricerca; nel management dei dati della ricerca (RDM) e condivisione delle problematiche nelle varie fasi del ciclo di vita del progetto di ricerca. Sono state inoltre elaborate delle Linee

guida per rendere i dati FAIR (reperibili, accessibili, interoperabili, riutilizzabili) e si attua una revisione dei Piani di gestione dei dati (DMP). Oltre a formazione e training specifico di tutti i professionisti coinvolti, l'altra attività si è concentrata nella creazione di un nuovo repository. È stato scelto Mendeley Data [28], uno strumento proprietario di Elsevier, fondamentalmente per ragioni di sostenibilità economica. È un repository comune basato su cloud gratuito in cui è possibile archiviare i dati, garantendo tutte le opzioni necessarie con la piattaforma Digital Commons Data [29], che fornisce gli strumenti necessari per gestire i dati di ricerca. Si è poi proceduto a un rebranding del repository che, grazie a Digital Commons Data, garantisce le funzionalità che ricercatori, amministratori e curatori di dati richiedono per archiviare, gestire, curare, condividere e conservare i dati, cui ha fatto seguito la creazione di nuove funzionalità di personalizzazione, ad esempio il nome – San Raffaele Open Research Data Repository (ORDR) [30] – e una nuova vetrina.

L'Unione europea, con il programma Horizon 2020, ha posto come un obbligo quello di depositare i dati in un archivio di fiducia, il più presto possibile, garantire l'accesso aperto secondo il principio "as open as possible, as closed as necessary", aperto quanto possibile, chiuso quanto necessario, utilizzando il più possibile licenze aperte [31].

L'ORDR garantisce i requisiti richiesti dalla UE in quanto ha: identificatori persistenti (DOI) per consentire riferimenti e citazioni; set completo di campi di metadati per consentire la scoperta e riutilizzo; versioning per tenere traccia di ogni modifica ai dati nel tempo; opzioni di condivisione multiple (Aperto vs accesso limitato ai contenuti); pre-moderazione: il gruppo esamina ogni set di dati prima che venga messo a

disposizione del pubblico; disposizioni per la conservazione a lungo termine del materiale depositato. In conclusione, il messaggio chiave, affermano Morelli e Soriano, è: avere un repository istituzionale è necessario ma non è abbastanza. È necessario garantire un continuo supporto tecnico e non a tutti i professionisti coinvolti. Da ciò la necessità di un team con figure diverse e complementari, che non può prescindere dalla stretta collaborazione con la biblioteca di riferimento e dalla figura del bibliotecario.

Aspetti giuridici dei dati della scienza nei repository istituzionali

Giovanni Ziccardi, professore di filosofia del diritto all'Università degli studi di Milano, nel suo intervento, ha dimostrato che sono diverse le sfide che il professionista, giovane o anziano che sia, affronta oggi in una società tecnologica che sta mutando a velocità incredibili. Alfabetizzazione informatica e cultura digitale, innanzitutto, sono le prime due: devono essere poste al centro di ogni azione nel tentativo di superare un analfabetismo diffuso e una situazione delle infrastrutture, in molti casi, inadeguata. Un'attenzione particolare deve essere data, poi, al ruolo del diritto e delle norme, con interventi che da un lato non alterino il delicatissimo ecosistema tecnologico e non criminalizzino inutilmente le tecnologie e, dall'altro, che mettano però al centro i valori dell'essere umano e la protezione sua e dei suoi dati. Vi è la tendenza, oggi, da parte dello Stato di disciplinare l'informatica con un approccio liberticida, per una sorta di "paura ancestrale" che il legislatore si porta dietro nei confronti del digitale. Vi è, come ulteriore aspetto, un problema di responsabilità ed etica

di tutti i soggetti coinvolti nello sviluppo delle tecnologie che hanno cambiato, o che stanno cambiando, il nostro mondo. Il ritorno dell'attenzione verso la computer ethics, la responsabilità d'impresa e lo sviluppo di prodotti sostenibili è un segnale molto interessante. L'attenzione all'aspetto etico è, oggi, essenziale, vista anche la potenza delle tecnologie più moderne. La grande onda del digitale ha travolto la società e le professioni tradizionali, tra cui quelle legate al diritto. Il quadro di crisi attuale ci porta a ripensare il modo di lavorare nel nuovo mercato digitale, acquisendo competenze lontane dai domini di conoscenza tradizionali. Temi quali l'intelligenza artificiale, il coding, il legal tech, la cybersecurity, la profilazione, il family tech e l'internet delle cose devono oggi far parte delle nostre conoscenze.

Il legislatore europeo ha introdotto, nella lunga elencazione (che va dall'art. 15 all'art. 22 del GDPR), nuove prerogative riconosciute agli interessati dal trattamento dei dati, tenendo in considerazione l'attuale sviluppo delle nuove tecnologie che, potenzialmente, può determinare nuovi pericoli e rischi per i diritti e le libertà degli individui.

Il primo diritto dell'interessato è quello di sapere se un soggetto tratti o meno i suoi dati e di accedere, se del caso, a quelle informazioni.

La conoscenza del periodo di conservazione, le possibilità di rettifica e cancellazione e la "minaccia" di rivolgersi al garante in caso di insoddisfazione sono altri tre pilastri dell'esercizio dei diritti. Il diritto all'oblio, o alla cancellazione, sta diventando un elemento centrale nella moderna società dell'informazione, soprattutto con riferimento ai dati trattati dalle piattaforme e dai motori di ricerca. Anche il diritto alla portabilità ha un aspetto tecnico molto marcato: occorre infatti far sì che tutti i dati presenti in un

archivio siano, per così dire, "impacchettati" e scaricabili dall'utente, sia che voglia custodirli (quindi la portabilità è intesa come possibilità per l'utente di archiviare in proprio i dati), sia che voglia che siano trasmessi a un altro gestore di dati. L'importanza della cifratura degli archivi: un data breach su un archivio di dati cifrato allo stato dell'arte può evitare grandi responsabilità in capo al titolare e annullare qualsiasi effetto dannoso dell'evento nei confronti di chiunque, titolare, interessati e terzi.

In sintesi, Ziccardi, autore di molte pubblicazioni su questi argomenti [32], ha sottolineato che nella gestione dei dati trattati nelle pubblicazioni o nei dati supplementari è necessario che:

- l'informativa sul loro trattamento preveda un consenso dinamico;
- sia esplicitato il periodo di conservazione;
- sia posta attenzione all'anonimizzazione, diversa dalla pseudonimizzazione (ogni scelta deve essere motivata);
- i dati siano cifrati;
- sia verificato il rapporto con i responsabili esterni (quale cloud si usa, per esempio).

Ziccardi conclude il suo intervento segnalando a tutti noi che intendiamo farci portatori di queste novità di mediazione che l'obiettivo primario deve rimanere sempre il bene pubblico, il bene della persona e che la finalità di utilizzo o riutilizzo dei dati della scienza deve essere sempre quello della crescita della salute e del benessere dell'umanità.

Discussione e conclusioni

Dopo le presentazioni che hanno ben evidenziato i problemi sul tavolo, i due moderatori – Francesca Gualtieri per il GIDIF-RBM e Moreno Curti per Bibliosans – sottolineano

no le questioni aperte: il tema della sostenibilità appare centrale, sia quella finanziaria che quella umana. Il simposio, com'era nelle intenzioni, si conclude con alcuni problemi aperti e con alcuni messaggi su cui riflettere e continuare a lavorare. I più importanti possono essere così riassunti:

1. le organizzazioni del mondo della salute devono gestire accuratamente i dati e i prodotti della propria attività e ricerca per conservare, preservare e trasmettere la conoscenza prodotta;
2. per gestire tali dati, i repository costruiti su misura con software open source sono più in linea con i principi della scienza aperta ma non sempre sostenibili sul piano finanziario
3. i modelli di repository – mono-istituzionali o consortili – devono comunque garantire ad un'organizzazione il controllo dei propri dati;
4. le organizzazioni hanno la necessità di confrontarsi sulle policy di gestione di tali repository;
5. ogni organizzazione deve stabilire cosa è obbligatorio e cosa è meritevole di essere immesso nel proprio deposito digitale, oltre alle pubblicazioni scientifiche, sicuramente ciò che è utile per la propria comunità di riferimento;
6. sarà sempre necessario garantire il rispetto dei diritti di proprietà intellettuale insieme all'accesso aperto ed avere estrema cura, competenza ed etica, nel trattare i dati;
7. sono necessarie adeguate risorse per garantire il ciclo di vita dei documenti digitali.

Si evince che la sostenibilità è fondamentale così come il ruolo dei documentalisti biomedici. La professionalità del documentalista è centrale ma è altresì importante

che tali professionisti siano coinvolti in tutte le fasi del processo di ideazione, gestione, conservazione e trasmissione dei repository istituzionali o tematici.

Il documentalista e/o bibliotecario biomedico conferma la centralità della sua mediazione culturale trasversalmente a tutti gli ambiti della scienza “aperta” per il bene della comunità non solo scientifica

Ringraziamenti

GIDIF-RBM ringrazia tutti coloro che sono intervenuti e hanno reso possibile la realizzazione del simposio in occasione del 37° dalla fondazione dell'Associazione, in particolare un sentito ringraziamento va al prof. Venuda per la moderazione e l'attenzione ai contenuti per aver stimolato con puntuale criticità l'analisi dei temi proposti.

**FRANCESCA GUALTIERI,
IVANA TRUCCOLO, CHIARA
FORMIGONI, FULVIA MERLINI,
SILVIA MOLINARI**

GIDIF-RBM Working Group
direttivo@gidif-rbm.org

BIBLIO-SITOGRAFIA

1. V. Comba, *Institutional Repositories e riviste Open Access danno più impatto alla ricerca: le nuove metriche di valutazione*, 2006. <http://eprints.rclis.org/7811>.
2. V. Comba, *Il progetto AEPIC (Academic E-Publishing Infrastructures - CILEA): le esigenze degli autori e lo stato dell'arte dei servizi*, Bibliotime, anno VI, numero 3, 2003.
3. De Robbio, *Archivi aperti e comunicazione scientifica*, Napoli: Clio-Press, 2007. [Online]. Disponibile su: <https://core.ac.uk/download/pdf/11913307.pdf>.
4. Venturato, *Gli archivi istituzionali*

ad accesso aperto delle università italiane, ott. 2009, Consultato: 4 novembre 2022. [Online]. Disponibile su: https://www.academia.edu/1039446/Gli_archivi_istituzionali_ad_accesso_aperto_delle_università_italiane_Unindagini.

5. P. Bellini, *Repositories istituzionali e valutazione della ricerca*, presentato al 54 Congresso nazionale AIB. Le politiche delle biblioteche in Italia. Il sistema bibliotecario nazionale, Firenze, nov. 2007. [Online].
6. S. Turbanti, *Repository istituzionali e social network mondi (soltanto) contrapposti*, 2017, [Online].
7. M. Tammam, *I depositi istituzionali ed i nuovi paradigma della comunicazione scientifica*, *Institutional repositories and the new paradigm of scholarly communication*, mag. 2007. [Online].
8. M. Cassella, *Repository istituzionali e CRIS ovvero la valutazione serve alla causa della scienza aperta?* *Biblioteche oggi*, vol. 36, pp. 19–28, set. 2018, doi: 10.3302/0392-8586-201806-019-1.
9. Open Archives Initiative. <http://www.openarchives.org>.
10. Solodovnik, *Repository Istituzionali Open Access e strategie Linked Open Data: Per una migliore comunicazione dei prodotti della ricerca scientifica*, vol. 167. Florence: Firenze University Press, 2016. doi: 10.36253/978-88-6655-928-3.
11. R. Ridi, *Presentazione*, in *Archivi aperti e comunicazione scientifica / Antonella De Robbio*, Napoli: Clio-Press, 2007, pp. ix–xii. [Online]. Disponibile su: <https://core.ac.uk/download/pdf/11913307.pdf>.
12. CRUI - Conferenza dei Rettori delle Università italiane, *Linee guida per gli archivi istituzionali*. <https://www.cru.it/linee-guida-per-gli-archivi-istituzionali.html>.
13. P. De Castro e E. Poltronieri, *Proceedings of the Conference on Institutional archives for research: experiences and projects in Open Access*. Istituto Superiore di Sanità. Rome, 30 No-

- vember- dicembre 2006. Roma: Istituto Superiore di Sanità, *RAPPORTI ISTISAN*, vol. 7, fasc. 12, pp. 1–112, 2007, [Online]. Disponibile su: <https://www.iss.it/documents/20126/45616/07-12.1183020754.pdf/33ddf7f4-e14e-17f4-e-27a-ba719ddfd46?t=1581094804854>.
14. M. Mazzocut, J. Piscanc, N. Gruarin, R. Ricci, R. Trampus, e I. Trucolo, *The inside out library: CRO OpenDocuments digital institutional repository*, *ArTS Archivio della ricerca di Trieste*, 2016. <https://arts.units.it/handle/11368/2877808>.
 15. IRIS: la gestione della Ricerca, *Cineca*. http://www.cineca.it/index.php/sites/default/files/IRIS_Cineca_web.pdf.
 16. J.C. Guedon, *ARL Proceedings 138: In Oldenburg's Long Shadow: Librarians, Research Scientists, Publishers, and the Control of Scientific Publishing*. The Association of Research Libraries, 2021. [Online]. Disponibile su: http://eprints.rclis.org/6375/1/ARL_Proceedings_138_In_Oldenburg%27s_Long_Shadow%2C_by_Guedon.htm.
 17. T. H. Vines *et al.*, *The Availability of Research Data Declines Rapidly with Article Age*, *Current Biology*, vol. 24, pp. 94–97, gen. 2014, doi: 10.1016/j.cub.2013.11.014.
 18. The Dataverse Project - Dataverse.org. <https://dataverse.org/home>.
 19. G. King, *An Introduction to the Dataverse Network as an Infrastructure for Data Sharing*, *Sociological Methods and Research*, vol. 36, pp. 173–199, 2007.
 20. GO FAIR Initiative. <https://www.go-fair.org>.
 21. EOSC Marketplace. <https://marketplace.eosc-portal.eu>.
 22. Iatropoulou, OpenAIRE, *OpenAIRE*. <https://www.openaire.eu>.
 23. OpenAIRE Guidelines for Data Archives — OpenAIRE Guidelines documentation. <https://guidelines.openaire.eu/en/latest/data/index.html>.
 24. Che cos'è un data lake?, *Google Cloud*. <https://cloud.google.com/learn/what-is-a-data-lake?hl=it>.
 25. Pavlidou, *OpenAIRE ELD-Enrich Local Data via the OpenAIRE Graph*, *OpenAIRE*. <https://www.openaire.eu/openaire-eld-enrich-local-data-via-the-openaire-graph>.
 26. Bollini, S. Mornati, G. Digilio, J. Piščanc, M. Artini, e C. Atzori, *Results from the OpenAIRE Call for Innovation: Enrich local data via the OpenAIRE Graph*, presentato al Open Repositories 2021 (OR2021), 10 giugno 2021. doi: 10.5281/zenodo.4923451.
 27. LIBER Research Data Management Working Group, *LIBER Europe*. <https://libereurope.eu/working-group/research-data-management>.
 28. Elsevier, Mendeley Data. <https://data.mendeley.com>.
 29. «Digital Commons Data: for your institution's RDM journey». <https://www.elsevier.com/solutions/digital-commons/data>.
 30. San Raffaele Open Research Data Repository. <https://ordr.hsr.it/research-data>.
 31. Vieira, *How to comply with Horizon Europe mandate for RDM*, *OpenAIRE*. <https://www.openaire.eu/how-to-comply-with-horizon-europe-mandate-for-rdm>.
 32. Università degli Studi di Milano, [Pubblicazioni di Giovanni Ziccardi], *IRIS AIR*. <https://air.unimi.it/simple-search?query=giovanni+ziccardi>.