



# RAPPORTI ISTISAN 17|32

ISSN: 1123-3117 (cartaceo) • 2384-8936 (online)

## **Risultati dell'indagine BISA (Bibliosan per la Scienza Aperta) sui dati aperti per la ricerca**

Gruppo di lavoro BISA



FORMAZIONE  
E INFORMAZIONE



**ISTITUTO SUPERIORE DI SANITÀ**

**Risultati dell'indagine BISA  
(Bibliosan per la Scienza Aperta)  
sui dati aperti per la ricerca**

Gruppo di lavoro BISA

ISSN: 1123-3117 (cartaceo) • 2384-8936 (online)

**Rapporti ISTISAN  
17/32**

Istituto Superiore di Sanità

**Risultati dell'Indagine BISA (Bibliosan per la Scienza Aperta) sui dati aperti per la ricerca.**

Gruppo di lavoro BISA

2017, vi, 54 p. Rapporti ISTISAN 17/32

Il Gruppo di lavoro Bibliosan per la Scienza Aperta (BISA), nato in seno alla rete Bibliosan (Sistema bibliotecario degli enti di ricerca biomedici italiani) nel giugno 2016, ha condotto una indagine sul trattamento dei dati di ricerca prodotti dagli enti appartenenti al sistema medesimo. L'indagine, basata su un questionario online, si è rivolta a tutte le 59 istituzioni afferenti alla rete Bibliosan e le conclusioni raggiunte sono presentate in questo rapporto. I risultati dell'indagine mostrano lo scenario attuale dei ruoli e delle responsabilità coinvolti nella gestione dei dati di ricerca all'interno della comunità scientifica italiana di Bibliosan nonché le criticità che dovranno essere superate per un trattamento dei dati efficiente e sostenibile. Inoltre, le indicazioni scaturite dal questionario potranno anche contribuire a sviluppare norme, politiche e pratiche comuni a livello nazionale, adattare i servizi di dati per la diffusione e assicurare l'archiviazione dei dati di ricerca alle esigenze esistenti.

*Parole chiave:* Dati aperti; Ricerca biomedica; Indagini e questionari; Italia

Istituto Superiore di Sanità

**Results of BISA (Bibliosan per la Scienza Aperta) survey on open research data.**

Working Group BISA

2017, vi, 54 p. Rapporti ISTISAN 17/32 (in Italian)

The Working Group BISA (*Bibliosan per la scienza aperta*), set up in June 2016 within Bibliosan (Library System of Italian Biomedical Research Institutions), carried out a survey to provide an overview of the handling of research data produced by the institutions belonging to the Bibliosan system. The survey, based on an online questionnaire, was addressed to all 59 institutions involved in the Bibliosan network and the research findings are presented in this report. Survey's results show the current scenario of roles and responsibilities involved in the management of research data within the Italian scientific community of Bibliosan as well as the obstacles that need to be overcome for an efficient and sustainable data handling. Moreover, the suggestions originated from the questionnaire will help to develop standards, policies and common practices at a national level, to adapt data services for dissemination and secure archiving of research data to the existing needs.

*Key words:* Open data; Biomedical research; Surveys and questionnaires; Italy

Si ringraziano il Dott. Giovanni Leonardi (Ministero della Salute, Direzione Generale Ricerca ed Innovazione in Sanità) per la costante azione di impulso e sostegno alla realizzazione degli obiettivi dell'Indagine BISA, i Direttori scientifici IRCCS e IZS per l'adesione alle finalità dell'Indagine BISA e i referenti Bibliosan presso le biblioteche per il fattivo coinvolgimento in tutte le fasi dell'indagine, dalla somministrazione del Questionario fino alla raccolta dei risultati.

Si ringraziano Patrizia Mochi e Silvia Negrola per il prezioso lavoro di supporto tecnico-editoriale nella stesura del documento.

Per informazioni su questo documento scrivere a: [elisabetta.poltronieri@iss.it](mailto:elisabetta.poltronieri@iss.it)

Il rapporto è accessibile online dal sito di questo Istituto: [www.iss.it](http://www.iss.it).

Citare questo documento come segue:

Gruppo di lavoro BISA. *Risultati dell'Indagine BISA (Bibliosan per la Scienza Aperta) sui dati aperti per la ricerca*. Roma: Istituto Superiore di Sanità; 2017. (Rapporti ISTISAN 17/32).

---

Legale rappresentante dell'Istituto Superiore di Sanità: *Gualtiero Ricciardi*

Registro della Stampa - Tribunale di Roma n. 114 (cartaceo) e n. 115 (online) del 16 maggio 2014

Direttore responsabile della serie: *Paola De Castro*

Redazione: *Paola De Castro* e *Sandra Salinetti*

La responsabilità dei dati scientifici e tecnici è dei singoli autori, che dichiarano di non avere conflitti di interesse.



**Gruppo di lavoro BISA (Bibliosan per la Scienza Aperta)\***

ISTITUTO SUPERIORE DI SANITÀ, Roma

Servizio Conoscenza e Comunicazione Scientifica

*Attività Editoriali*

Elisabetta POLTRONIERI (coordinatore)  
Paola DE CASTRO

*Biblioteca*

Rosalia FERRARA  
Cristina MANCINI  
Franco TONI

IRCCS FONDAZIONE POLICLINICO SAN MATTEO, Pavia

*Direzione Scientifica*

Moreno CURTI (coordinatore sistema Bibliosan)

ISTITUTO ZOOPROFILATTICO SPERIMENTALE DELLA SICILIA, Palermo

*Dipartimento Sanità Territoriale*

Angela VULLO

*Servizio Editoria e Biblioteca*

Pietro LA PLACA

---

\* Il Gruppo di lavoro si è avvalso del contributo di Gaetana Cognetti (Biblioteca Digitale “Riccardo Maceratini” e Biblioteca del Paziente, IRCCS, Istituto Regina Elena per lo Studio e la Cura dei Tumori), in qualità di co-coordinatore fino al 20 febbraio 2017.



# INDICE

<b>Premessa</b> .....	v
<b>Genesi e prospettive del Gruppo di lavoro BISA (Bibliosan per la scienza aperta)</b> <i>Moreno Curti, Paola De Castro, Rosalia Ferrara, Cristina Mancini, Elisabetta Poltronieri, Franco Toni</i> .....	1
<b>Quadro europeo sui dati aperti della ricerca e contesto italiano</b> <i>Paola Gargiulo</i> .....	5
<b>Indagine BISA: risultati e analisi dei dati</b> <i>Rosalia Ferrara, Pietro La Placa, Cristina Mancini, Elisabetta Poltronieri, Angela Vullo</i> .....	12
<b>Riflessioni sugli esiti delle iniziative BISA e progetto e-Infrastructures Austria</b> <i>Paolo Budroni</i> .....	28
<b>Raccogliere la sfida della scienza aperta</b> <i>Gruppo di lavoro BISA (Bibliosan per la Scienza Aperta)</i> .....	40
<b>Appendice</b> Questionario BISA.....	43



## PREMESSA

La condivisione dei dati è una tematica di interesse strategico per il progresso della scienza e in particolare per la salute pubblica. Se ne parla ormai quotidianamente dibattendo sulla necessità di valorizzare le risorse della conoscenza, per evitare sprechi e duplicazioni e per massimizzare gli investimenti, e limitare le difficoltà di condivisione dettate soprattutto da questioni di riservatezza dei dati e di gestione dei diritti di proprietà intellettuale.

La realtà è che i dati della ricerca, memorizzati in formato digitale, crescono in maniera esponenziale e pongono problematiche che la comunità scientifica non è ancora in grado di fronteggiare completamente. Siamo all'inizio di un processo di cambiamento piuttosto complesso e le soluzioni non sono sempre facili pur riconoscendo la necessità che i dati prodotti siano aperti, cioè accessibili e disponibili per la condivisione e il riutilizzo da parte della comunità scientifica.

In tale prospettiva, l'Istituto Superiore di Sanità (ISS) si è aperto alla comunità scientifica riconoscendo, già dal 2006, le potenzialità di un archivio digitale istituzionale (DSpace ISS <http://dspace.iss.it/dspace/>) e deputandolo a divenire lo strumento di raccolta e disponibilità online delle pubblicazioni scientifiche in una logica di interoperabilità dei sistemi informativi in rete. Ricordiamo che la *policy* dell'ISS per l'accesso aperto ai risultati della ricerca è del 2008, la prima in Italia in ambito biomedico.

Oggi, oltre all'accesso alle pubblicazioni scientifiche, si richiede l'accesso ai dati della ricerca, ma permane la consapevolezza diffusa di numerose questioni ancora non risolte, legate in particolare ai ruoli di responsabilità, all'assenza di formazione adeguata, alla variegata offerta di piattaforme informatiche e non ultimo ai costi della condivisione; tutte questioni che in generale dovrebbero essere affrontate e risolte attraverso una regolamentazione condivisa che tenga in considerazione gli interessi dei diversi stakeholder.

Il dibattito internazionale e nazionale in materia è molto stimolante. La Commissione Europea, dopo una fase pilota che ebbe avvio nel 2007 prevede oggi *open access by default*, l'obbligo cioè di diffondere in accesso aperto pubblicazioni e relativi dati per tutti i progetti Horizon 2020, con deroga al deposito dei dati solo per giustificati motivi.

In Italia, il versante accademico, grazie all'iniziativa della CRUI (Conferenza dei Rettori delle Università Italiane), ha già prodotto un modello di *policy* istituzionale sui dati della ricerca da sottoporre all'approvazione dei senati accademici.

Una risposta in tal senso va quindi sollecitata anche per le istituzioni di ricerca, in particolare in ambito sanitario, che devono poter contare su una politica di gestione dei dati chiara ed efficace, oltre che su linee guida che ne attuino le disposizioni.

Naturalmente, in un disegno politico comune a più istituzioni, andranno salvaguardate le specificità delle singole organizzazioni, operando in quell'alveo di trasparenza, cooperazione e progresso della conoscenza che sostanziano il principio di scienza aperta.

L'elaborazione di un modello di *policy* sostenibile a valenza nazionale, in ambito biomedico, è quindi il percorso ottimale per dotare la comunità dei ricercatori di un quadro di riferimento utile su come trattare i dati della ricerca, anche in relazione agli adempimenti previsti dai bandi per la concessione di finanziamenti da parte degli enti erogatori di fondi.

L'impegno del Gruppo di lavoro BISA (Bibliosan per la Scienza Aperta), promosso e coordinato dall'ISS, rappresenta la volontà di più istituzioni nell'ambito del Servizio Sanitario Nazionale orientate verso una necessità di apertura e di condivisione dei dati della ricerca.

Attraverso un'apposita indagine, i cui risultati sono presentati in questo volume, si è inteso raccogliere gli elementi di valutazione necessari per operare l'auspicato cambiamento, a partire dalla conoscenza delle pratiche correnti di trattamento dei dati.

Ringrazio quindi il Gruppo di lavoro BISA e mi auguro che questa prima ricognizione possa rapidamente portare ai risultati desiderati.

Walter Ricciardi  
*Presidente dell'Istituto Superiore di Sanità*

## 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](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](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 Bibliosan 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](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](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 EAHL* 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](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](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](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-profili-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](http://www.adnkronos.com/soldi/economia/2017/05/07/cercasi-professionisti-del-digitale-ecco-figure-piu-richieste_c3ZxzHoABmn6R3zLsc63XO.html); ultima consultazione: 16/08/2017.

## QUADRO EUROPEO SUI DATI APERTI DELLA RICERCA E CONTESTO ITALIANO

Paola Gargiulo

Consorzio Interuniversitario per Il Calcolo Automatico (Cineca), Milano

La scienza moderna, per sua propria natura, è aperta al confronto, allo scambio, alla condivisione e oggi può avvalersi, grazie al digitale, di strumenti e di tecnologie che le permettono di essere effettivamente aperta. Eppure tali opportunità sono ostacolate da resistenze culturali, interessi economici, barriere giuridiche e in misura minore tecnologiche, al punto tale che è stato coniato il termine *open science*, scienza aperta, apparentemente una tautologia, per definire l'approccio alla ricerca scientifica basato sulla collaborazione, sull'apertura, sulla trasparenza, sulla condivisione degli strumenti medesimi e dei risultati e sulla disseminazione in accesso aperto di questi ultimi.

Quando si parla di scienza aperta, non si fa riferimento solo ai dati della ricerca (*open research data*) e alle pubblicazioni (*open access publications*) risultanti dai progetti finanziati prevalentemente con fondi pubblici, ma anche alle metodologie usate nel corso della ricerca (*open methodology*), ai software utilizzati o sviluppati (*open source*), alle attività di revisione da parte dei pari, alle pubblicazioni in modalità aperta (*open peer reviewing*) e anche ai materiali per scopo didattico (*open educational resources*). Si tratta dell'intero processo della ricerca e dei suoi esiti che viene reso accessibile e possibilmente fruibile nelle modalità più aperte e della disseminazione dei risultati nelle comunità scientifiche, della loro conservazione, ma anche della loro divulgazione con mezzi e strumenti appropriati. Quest'ultimo aspetto è connesso al ruolo della scuola, ma anche a quello dei decisori politici locali, nazionali, internazionali e alla società in generale, perché tutti possano trarre vantaggio dalla condivisione della conoscenza e operare delle scelte consapevoli nell'interesse collettivo.

Non è questa la sede per discutere in modo diffuso delle implicazioni di questo approccio; certamente va detto che richiede un grosso cambiamento culturale, soprattutto per quanto attiene ai criteri di incentivazione, premialità e valutazione, che oggi determinano l'assegnazione di fondi, la progressione della carriera dei ricercatori medesimi. I modelli attuali, ancora di più delle politiche editoriali dei grandi gruppi editoriali internazionali, costituiscono un notevole impedimento al cambiamento. Si rinvia al recente saggio di Roberto Caso per una trattazione più completa delle criticità che attraversano la scienza oggi e sui benefici che un sistema aperto apporterebbe alla scienza curandola dai mali che l'affliggono (1).

Nel contesto europeo, alcuni finanziatori della ricerca, quali la Commissione Europea - Direzione Ricerca e Innovazione, lo *European Research Council*, enti finanziatori nazionali in diversi Paesi europei, finanziatori privati (*Wellcome Trust*, e, in Italia, *Telethon* e *Cariplo*) oltre, in alcuni casi, gli stessi governi nazionali, sono impegnati in un sostegno concreto alla scienza aperta che prevede interventi su più piani, dai cambiamenti nei modelli di comunicazione scientifica nei sistemi di validazione e valutazione, alla creazione di nuovi servizi e infrastrutture.

Di seguito, vengono illustrate brevemente le iniziative sui dati aperti della ricerca e più in generale sulla scienza aperta portate avanti dalla Commissione Europea, da alcuni Paesi europei e in Italia.

La Commissione Europea, in qualità di decisore politico, finanziatore e *capacity builder*, svolge da diversi anni una funzione di guida, di leadership nella promozione della scienza aperta. Da oltre un decennio commissiona studi, indagini, crea gruppi di lavoro con esperti, elabora

documenti, raccomandazioni, *policy* e finanzia progetti volti a implementare infrastrutture per la ricerca e infrastrutture di servizi tecnologici a supporto della scienza aperta. Sono, inoltre, numerosi i progetti finanziati sia nel Settimo Programma Quadro sia in *Horizon 2020* che hanno visto l'elaborazione di indagini, raccomandazioni, studi, materiali, workshop, attività di *training* e di *helpdesk* con la partecipazione di diversi Paesi europei e, in qualche caso, anche extraeuropei. Non possiamo elencarli tutti, ci limitiamo a segnalare il progetto OpenAIRE (<http://www.openaire.eu>), l'infrastruttura europea fondata sulla partecipazione di tutti i Paesi dell'Unione, che fornisce l'accesso, il monitoraggio delle pubblicazioni e dei dati della ricerca dei progetti sostenuti dai programmi di finanziamento della ricerca (FP7, *Horizon 2020*) e anche di progetti e programmi nazionali che aderiscono alla piattaforma OpenAIRE.

La Commissione, in qualità di finanziatore della ricerca, ha stabilito, alla fine del 2013, che tutti i beneficiari dei progetti di ricerca afferenti a *Horizon 2020* (2) che pubblicano i risultati in riviste *peer reviewed* sono tenuti a disseminarli in accesso aperto, secondo due strategie: la via verde (deposito del *post-print* dell'articolo in un *repository* aperto) o la via oro (pubblicazione su riviste *peer reviewed* ad accesso aperto). Dal 2017 anche i dati della ricerca che validano i risultati presentati nelle pubblicazioni (*underlying data*) vanno depositati, resi accessibili e fruibili, secondo i principi FAIR (*Findable, Accessibile, Interoperable, Reusable*), principi di rintracciabilità, accessibilità, interoperabilità e riusabilità/riproducibilità dei dati della ricerca, elaborati da esperti di diversi domini con la collaborazione del gruppo FORCE11 e dotati di un piano di gestione (3).

In un documento pubblicato nel 2013, un *concept paper* (4) sulla scienza digitale in *Horizon 2020*, la Commissione afferma che l'epoca digitale ha causato una radicale trasformazione della scienza e dell'innovazione grazie all'integrazione dell'ICT nel processo della ricerca e all'affermazione della cultura dell'*openness* e della condivisione che caratterizza Internet. Questa cultura richiede trasparenza nella metodologia sperimentale, nella raccolta dei dati, affidabilità e riusabilità pubblica dei dati scientifici e dei prodotti della ricerca, accessibilità, trasparenza nei processi della comunicazione scientifica e utilizzo di strumenti web basati sulla condivisione. La Direzione Ricerca e Innovazione della Commissione porta avanti questa visione della scienza - giova ricordare le tre "O" del motto dell'attuale Commissario alla Ricerca, all'Innovazione Carlos Moedas: *Open Science, Open Innovation, Open to the World* - consapevole delle sfide e dei rischi. Bisogna dire che le posizioni di altre Direzioni della Commissione sui temi della proprietà intellettuale e del mercato unico digitale sono ancora alquanto distanti da questa visione della scienza e pertanto persistono contraddizioni all'interno delle politiche della Commissione e del Parlamento europeo che purtroppo indeboliscono e rallentano il processo del cambiamento e rischiano di vanificare l'immane sforzo profuso. È molto importante pertanto che ci siano a livello nazionale e istituzionale azioni allineate con le posizioni espresse dalla Commissione, in modo assicurare una piena affermazione della scienza aperta.

La Direzione Ricerca e Innovazione ha definito 5 priorità:

1. sviluppo di infrastrutture di ricerca per la scienza aperta al fine di migliorare l'*hosting*, l'accesso e la *governance* dei dati della ricerca, attraverso lo sviluppo di un quadro di riferimento comune e la creazione dello *European Open Science Cloud*;
2. sostegno e promozione delle politiche a favore dell'accesso aperto ai dati della ricerca e alle pubblicazioni scientifiche;
3. rimozione delle barriere alla scienza aperta che richiede una revisione dei criteri relativamente alle carriere dei ricercatori e l'inclusione di incentivi e premi per coloro che si impegnano nel fare scienza aperta;
4. promozione e sostegno dell'integrità della ricerca, per l'affermazione di quei principi, valori etici, doveri deontologici e standard professionali che tutti coloro che partecipano

alla ricerca (dai ricercatori ai finanziatori, ai valutatori ecc.) sono tenuti a rispettare e che oggi sono fortemente minacciati (vedi il dibattito sulla cattiva scienza e sulla crisi della riproducibilità dei dati) e che rischiano di inficiare soprattutto la qualità, ma anche la reputazione e l'immagine pubblica della scienza;

5. creazione di incentivi a favore della scienza aperta tramite la promozione delle migliori pratiche e l'apertura della conoscenza a settori più ampi della società coinvolti nel progresso della conoscenza scientifica (*citizen science*).

Per realizzare queste priorità la Commissione ha istituito la *Open Science Policy Platform* (OSPP) e ha affidato agli esperti, scelti tramite una selezione pubblica, il compito di assisterla nella formulazione delle politiche e nell'implementazione delle stesse. In particolare, l'OSPP ha il compito di fornire raccomandazioni sulle azioni da intraprendere per migliorare la qualità e l'impatto della scienza europea e contribuire alla attuazione della *Open Science Agenda*, definita in 8 punti:

1. cambiamento dell'attuale sistema di premialità dei ricercatori;
2. individuazione di nuove misure per valutare la qualità e l'impatto della ricerca;
3. creazione di nuovi modelli economici per l'editoria scientifica;
4. impegno a rendere i dati della ricerca conformi ai principi della rintracciabilità, accessibilità, interoperabilità e riusabilità/riproducibilità;
5. realizzazione dello *European Open Science Cloud*;
6. integrità della ricerca;
7. *citizen science*;
8. *open education* e relative competenze da acquisire.

La Commissione ha inoltre avviato l'*Open Science Monitor* (<https://ec.europa.eu/research/openscience/index.cfm?pg=home&section=monitor>) che rileva la transizione alla scienza aperta tramite l'utilizzo di alcuni indicatori, raccogliendo dati provenienti da fonti affidabili relativamente alle politiche, ai *repository* esistenti, alle pubblicazioni, ai *pre-print*, ai dati delle ricerche, ai fattori che favoriscono la scienza aperta e che la ostacolano, nonché alle posizioni dei ricercatori sull'accesso aperto.

Attualmente, è in corso di valutazione la creazione di una piattaforma per la pubblicazione degli articoli, l'*Open Access Publishing Platform* (<https://ec.europa.eu/research/openscience/index.cfm?pg=home>), allo scopo di fornire una serie di servizi editoriali di cui potranno usufruire i beneficiari dei progetti, tra cui il supporto alla pubblicazione, servizi di *peer reviewing* e l'utilizzo di metriche alternative. Si tratterebbe di un'iniziativa simile a quella realizzata dal *Wellcome Trust* (5). Il gruppo di lavoro dell'OSPP, a sua volta suddiviso in sottogruppi, ha già provveduto a redigere le *Recommendations on Open Science Publishing* (6) e, di recente, ha rilasciato un *report* sulla valutazione delle carriere dei ricercatori e le pratiche della scienza aperta, curato dal sottogruppo di lavoro *Rewards* (7) e un *report* sulle competenze dei ricercatori e la pratica della scienza aperta a cura del sottogruppo *Skills* (8).

Sicuramente l'iniziativa di maggiore respiro sui dati della ricerca è la *European Open Science Cloud* (EOSC) (<https://ec.europa.eu/research/openscience/index.cfm?pg=open-science-cloud>) che risponde all'esigenza di creare una federazione delle infrastrutture di dati della ricerca esistenti e future e di tutti i servizi ad esse connessi. Tale bisogno è stato portato all'attenzione della Commissione da alcune comunità di ricerca europee, in particolare dal settore delle scienze della vita. È in questo ambito che da anni si avverte fortemente la necessità di condividere i dati e soprattutto di trovare soluzioni ai problemi di rintracciabilità, accessibilità, interoperabilità, riusabilità dei dati con i conseguenti vantaggi di debellare malattie, individuare cure appropriate in modo più veloce, senza sprechi e duplicazioni, puntando sulla collaborazione e sulla condivisione piuttosto che sulla competizione.

Nell'aprile 2016, la Commissione ha presentato una bozza di progetto sulla base del lavoro svolto dall'*High Level Expert Group* (HLEG) che aveva nominato alcuni mesi addietro. Il Gruppo ha pubblicato un rapporto nell'ottobre dello stesso anno, nel quale vengono esaminati la fattibilità di questa iniziativa, i benefici e le criticità e formulate precise raccomandazioni. La realizzazione di questo progetto richiede la partecipazione fattiva di vari attori: governi, enti finanziatori, istituzioni di ricerca, comunità di ricerca, infrastrutture di ricerca, gestori dei *data repository*, infrastrutture di rete e di supporto, fornitori di servizi. Per questo motivo, il 12 giugno 2017, la Commissione ha organizzato il primo EOSC Summit al quale hanno partecipato i diversi attori che hanno esposto le loro posizioni. I risultati di questa consultazione, così come quelli del workshop organizzato alla fine di giugno sugli aspetti relativi alle modalità di governo e di finanziamento di EOSC, saranno presenti nella RoadMap che la Commissione presenterà agli stati membri alla fine di quest'anno.

Nel frattempo ha preso il via l'EOSC Pilot (<https://eoscpilot.eu/>), il progetto pilota che sostiene l'iniziativa nella prima fase di sviluppo, tramite la formulazione di proposte concrete e operative sulla *governance* di EOSC, lo sviluppo di "dimostratori" per sperimentare in concreto l'integrazione e l'interoperabilità tra i *data repository* di diverse comunità scientifiche e la promozione di politiche sui dati sulla ricerca. Su questi ultimi, la Commissione, in qualità di decisore politico, ha deliberato la politica di accesso aperto sopracitata, relativa ai dati della ricerca generati e utilizzati nei progetti finanziati in Horizon 2020 e, in qualità di erogatore di fondi, ha finanziato diversi progetti focalizzati sui dati della ricerca di seguito brevemente descritti:

1. RECODE (*Policy Recommendations for Open Access to Research Data in Europe*) (<http://recodeproject.eu/>), finanziato dalla Commissione nel Settimo Programma Quadro. Tale progetto ha generato le *Policy Guidelines on Open Access and Data Dissemination and Preservation* che raccolgono raccomandazioni pratiche rivolte agli enti finanziatori, alle istituzioni di ricerca, ai gestori dei dati e agli editori;
2. LEARN (*Leaders Activating Research data management Network*) (<http://learn-rdm.eu/en/about/>), finanziato nel programma Horizon 2020 e terminato qualche mese fa. Anche in questo caso sono stati prodotti utili strumenti, tra cui un ricco Toolkit of Best Practice for Research Management che raccoglie raccomandazioni relativamente alle politiche e al piano di gestione dei dati, diversi studi di casi e best practice oltre a una griglia per la valutazione del piano di gestione;
3. EDISON (*Building the Data Science Profession*) (<http://edison-project.eu/>) per la creazione, l'analisi, la gestione, la cura e la conservazione dei dati, tutti processi che richiedono nuove competenze professionali; tale progetto, appena terminato, ha affrontato il tema della formazione, del curriculum del *data scientist* circa le competenze da acquisire nell'analisi dei dati e in materia di *data science engineering* e delle conoscenze di dominio.

Inoltre, all'interno di FP7 (*Seventh Framework Programme*) e Horizon 2020, sono stati finanziati, oltre al sopracitato OpenAIRE, altri progetti su infrastrutture e servizi quali EUDAT (<http://www.eudat.eu/>), l'iniziativa paneuropea finalizzata alla creazione di un'infrastruttura interdisciplinare e transnazionale per i dati della ricerca allo scopo di fornire servizi per l'accesso e la conservazione dei dati e, più recentemente, OpenMintEd (<http://openminted.eu/>) sulla creazione di un'infrastruttura per il supporto del *text* e *data mining*.

Spostandoci sul piano nazionale, sono vari i Paesi impegnati nell'elaborare politiche sulla scienza aperta e sui dati aperti della ricerca. Di recente SPARC Europe, in collaborazione con il Digital Curation Centre (DCC), ha pubblicato il rapporto *Open Data and Open Policy in Europe* che verrà regolarmente aggiornato (9). Il documento censisce quanto avviene a livello nazionale nei singoli Paesi e fornisce tutti i link necessari per gli approfondimenti. Il quadro che emerge

presenta una situazione in movimento, molto articolata, che riflette le differenze tra i Paesi nonché le diverse sensibilità e modalità di approccio su questi temi.

Sono pochi i Paesi dotati di avanzate infrastrutture nazionali di supporto alla gestione e alla cura dei dati della ricerca, dotati di strategie nazionali. Alcuni hanno approvato politiche nazionali e sono impegnati nella loro attuazione (Norvegia, Cipro), altri hanno optato per le RoadMap (Austria, Belgio, Croazia, Finlandia). La Danimarca ha elaborato una strategia e ha creato un forum per promuovere un approccio partecipativo che prevede il coinvolgimento di tutti i soggetti interessati. La Francia ha approvato una legge nazionale sulla repubblica digitale che affronta la questione del riutilizzo dei dati. In Portogallo, il principale ente finanziatore della ricerca ha approvato una policy sulla gestione dei dati in corso di attuazione; in Germania, il DFG (Deutsche Forschung Gesellschaft), il principale ente finanziatore della ricerca, ha approvato le linee guida sulla gestione dei dati della ricerca e la Helmholtz Association, che raccoglie 18 enti di ricerca tedeschi in ambito tecnico scientifico e biomedico, ha adottato un *Position paper on the management of research data*. In Gran Bretagna, i Research Councils, principali finanziatori della ricerca, hanno approvato gli *RCUK Common Principles on Data Policy* pubblicati nel 2011 e rivisti nel 2015, che sono alla base delle politiche individuali adottate dai medesimi enti. Di recente, è stato approvato un interessante documento, *Concordat on Open Research Data*, sottoscritto dai finanziatori della ricerca pubblici e privati britannici e dalle università, che affronta e propone soluzioni su problemi sollevati dalla condivisione dei dati. I Paesi Bassi hanno confermato il loro impegno a sostegno della scienza aperta durante la Presidenza dell'EU nel primo semestre del 2016, inserendo l'*open science* nell'agenda politica; di recente, hanno dato vita al *National Plan Open Science*, a lungo discusso dagli enti finanziatori della ricerca, dai ricercatori, dalle università e dai centri di ricerca e al National Coordination Point for Research Data Management. Altri Paesi (Bulgaria, Grecia, Irlanda, Islanda, Lituania, Lettonia, Lussemburgo, Malta, Romania, Repubblica Ceca, Polonia, Slovacchia, Spagna, Svezia, Svizzera, Ungheria) non hanno ancora concretizzato iniziative di respiro nazionale. L'Italia fa parte di questo ultimo gruppo.

In Italia abbiamo una situazione molto frammentaria; disponiamo di infrastrutture nazionali per alcune specifiche comunità di ricerca, siamo membri di diverse infrastrutture di ricerca europee afferenti ad ERIC (*European Research Infrastructure Consortium*), abbiamo consolidate infrastrutture di servizi quali Cineca (Consorzio Interuniversitario per Il Calcolo Automatico), GARR (Gruppo per l'Armonizzazione delle Reti della Ricerca) e partecipiamo ad organizzazioni e comunità internazionali che si occupano di dati della ricerca come Research Data Alliance (<https://www.rd-alliance.org/>). Disponiamo di *expertise* e competenze, ma siamo privi di coordinamento, di politiche e strategie nazionali che dovrebbero poi riflettersi nella creazione di servizi, strutture e supporto a livello istituzionale (10). Non si profilano per ora azioni concrete da parte del Governo. Vi sono state enunciazioni da parte del MIUR in più occasioni; tra le più recenti quelle riportate nel recente documento sui big data, redatto da un Gruppo di lavoro del Ministero: “È utile ricordare in proposito che l'attuale Programma Nazionale della Ricerca Italiano intende favorire lo sviluppo e la diffusione di *open science* e *big data* nel senso qui indicato. La previsione è di adottare una *policy* nazionale relativa al deposito, l'accesso aperto, la verificabilità e il riutilizzo dei prodotti e dei dati della ricerca, anche in ottica *big data*. Al riguardo, un gruppo di lavoro a questo dedicato definirà e proporrà in tempi brevi le strategie, le linee di indirizzo, piani di attuazione e strumenti condivisi a livello inter-istituzionale, europeo e internazionale per realizzare in Italia l'*open science*, l'*open access* e la *digital preservation*” (11).

In compenso sono in corso alcune iniziative dal basso che da una parte promuovono la necessaria sensibilità e attenzione sui temi complessi legati ai dati della ricerca e della scienza aperta e, dall'altra, affrontano in modo pratico la gestione dei dati.

Nella primavera del 2016 si è costituito un gruppo informale di responsabili dell'ufficio ricerca, di informatici, di bibliotecari e di responsabili degli archivi aperti della ricerca, appartenenti al Politecnico di Milano e agli atenei di Milano Statale, Torino, Trento e Venezia Ca' Foscari, con l'obiettivo di dare risposte concrete e supporto ai ricercatori nella gestione dei dati e di fornire raccomandazioni e linee guida alle rispettive amministrazioni. Per ora il gruppo ha elaborato un modello di politica istituzionale sui dati aperti e una griglia di supporto nella redazione del piano di gestione dei dati; è in corso di completamento un'analisi comparata delle infrastrutture esistenti. I documenti, disponibili in rete ([http://wikimedia.sp.unipi.it/index.php/OA\\_Italia/Risorse\\_sugli\\_open\\_research\\_data#Italy](http://wikimedia.sp.unipi.it/index.php/OA_Italia/Risorse_sugli_open_research_data#Italy)), sono stati inviati al Gruppo di lavoro sull'Open Access della CRUI e auspicabilmente saranno adottati dalla CRUI.

La costituzione del Gruppo di lavoro BISA (Bibliosan per la scienza aperta), all'interno della rete Bibliosan, è un altro esempio di consapevolezza della necessità di affrontare le questioni aperte poste dal trattamento, dalla conservazione, dall'accessibilità e dal riuso dei dati della ricerca in ambito biomedico. L'indagine promossa da BISA allo scopo di indagare pratiche e aspettative della comunità di ricerca biomedica, circa la collaborazione nell'individuare soluzioni sostenibili da fornire ai decisori politici per la definizione di politiche di gestione dei dati della ricerca, costituisce un ottimo punto di partenza. I risultati dell'indagine illustrati in questa pubblicazione forniscono un quadro articolato sulle necessità di intervento sul piano infrastrutturale, tecnico, giuridico, formativo, tutti elementi sufficienti per elaborare un documento di indirizzo e intraprendere un piano di azione che veda coinvolti la comunità di ricerca della rete Bibliosan, il personale tecnico che si è fatto promotore dell'iniziativa e il Ministero della Salute.

Dal 2015 si è costituita AISA (Associazione Italiana per la Promozione della Scienza Aperta) (<http://aisa.sp.unipi.it/>) cui afferiscono alcuni atenei italiani e singoli docenti, ricercatori, nonché esperti di *open access*. AISA nasce dalla consapevolezza dell'assenza di una cultura diffusa su proprietà intellettuale, comunicazione della scienza e scienza aperta nel mondo accademico e della ricerca in Italia e dalla convinzione che si debba colmare questo vuoto. Inoltre, l'investimento nella formazione di studenti, scienziati, bibliotecari e personale amministrativo delle strutture accademiche e di ricerca è indispensabile per superare le barriere culturali, giuridiche, economiche che oggi affliggono la scienza. Partendo da questa prospettiva, AISA si impegna a condurre una campagna di informazione, promozione, *advocacy*, presso atenei, enti di ricerca e organi di governo delle singole istituzioni nazionali, tramite la pubblicazione di articoli e saggi, l'organizzazione di un convegno annuale, la partecipazione a convegni, workshop, progetti nazionali e internazionali e attraverso attività di formazione a cura dei propri membri.

Degna di attenzione e di sostegno è l'iniziativa di AISA di modifica della legge sul diritto d'autore (12) per contribuire ad abbattere le barriere alla diffusione della conoscenza, allineando l'Italia ad altri Paesi europei che si sono attivati in questa direzione (Francia, Germania, Paesi Bassi). Si tratterebbe di aggiungere un articolo al dettato normativo attuale, con la finalità di attribuire un diritto di ripubblicazione in accesso aperto agli autori di opere scientifiche quali articoli apparsi in riviste, capitoli pubblicati in libri collettanei e monografie finanziate con fondi di ricerca.

Dopo questa breve panoramica sulla scienza aperta e sui dati della ricerca, possiamo concludere che le iniziative sono tante, come altrettante sono le sfide. Queste ultime si potranno vincere in Italia e altrove con la partecipazione attiva di tutti gli attori e attraverso il cambiamento dei modelli culturali correnti, in particolare dei sistemi di valutazione e di premialità; i benefici saranno enormi. È dunque auspicabile in tal senso la collaborazione tra i gruppi di lavoro che si occupano di queste tematiche, nel caso specifico tra BISA e il gruppo informale costituitosi tra alcuni atenei, o anche tra altri gruppi. Sono necessarie azioni congiunte presso gli organi di governo per realizzare una visione coordinata e strategica a livello nazionale. Al medesimo tempo,

l'eliminazione delle barriere culturali e giuridiche, il cambiamento culturale e l'acquisizione di nuove competenze, l'implementazione delle infrastrutture e dei servizi a supporto della scienza aperta in Italia, e i relativi finanziamenti, potranno ricevere una notevole spinta se gli obiettivi di AISA saranno sostenuti da parte delle istituzioni e dei singoli ricercatori.

## Bibliografia

1. Caso R. *Open Science*. Geneva: CERN; 2017. Disponibile all'indirizzo: <https://zenodo.org/record/838986#.Wa6fS8ZLe1s>; ultima consultazione 31/8/2017.
2. European Commission Directorate-General for Research & Innovation. *2020 Programme – Guidelines to rules on Open Access to Scientific Publications and Open Access to Research Data in Horizon 2020, vers.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](http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf); ultima consultazione 31/8/2017.
3. Wirkinson MD, Dumontier M, Aalbersberg IJ, Appleton G, Axton M, Baak A, *et al.* The FAIR guiding principles for scientific data management and stewardship. *SciData* 2016;3:160018. Disponibile all'indirizzo: <https://www.nature.com/articles/sdata201618#bx2>; ultima consultazione 21/08/2017.
4. European Commission. *Digital Single Market. Digital Science in Horizon 2020*, Brussels: European Commission; 2013. Disponibile all'indirizzo: [http://ec.europa.eu/information\\_society/newsroom/cf/dae/document.cfm?doc\\_id=2124](http://ec.europa.eu/information_society/newsroom/cf/dae/document.cfm?doc_id=2124); ultima consultazione 31/8/2017.
5. Butler D. Gates Foundation announces open-access publishing venture. *Nature* 2017;543(7647):599. Disponibile all'indirizzo: [http://www.nature.com/polopoly\\_fs/1.21700!/menu/main/topColumns/topLeftColumn/pdf/nature.2017.21700.pdf](http://www.nature.com/polopoly_fs/1.21700!/menu/main/topColumns/topLeftColumn/pdf/nature.2017.21700.pdf); ultima consultazione 06/12/2017.
6. Leonelli S, Lawrence R, Edmund J. *Recommendations on open science publishing. Adopted by open science platform on 25 April 2017*. Brussels: European Commission; 2017. Disponibile all'indirizzo: [https://ec.europa.eu/research/openscience/pdf/ospp\\_open\\_access\\_publishing\\_report.pdf](https://ec.europa.eu/research/openscience/pdf/ospp_open_access_publishing_report.pdf); ultima consultazione 31/08/2017.
7. European Commission, DG for Research and Innovation Open Science and ERA policy. *Evaluation of research career fully acknowledging open science practices. Rewards, incentives and/or recognition for researchers practicing Open Science*. Brussels: European Commission; 2017. Disponibile all'indirizzo: [https://cdn1.euraxess.org/sites/default/files/policy\\_library/os-rewards-wgreport-final\\_integrated\\_0.pdf](https://cdn1.euraxess.org/sites/default/files/policy_library/os-rewards-wgreport-final_integrated_0.pdf); ultima consultazione 31/8/2017.
8. European Commission, DG for Research and Innovations Open Science and ERA policy. *Providing researchers with the skills and competencies they need to practice open science. Open Science Skills Working Group Report*. Brussels: European Commission; 2017. Disponibile all'indirizzo: [https://ec.europa.eu/research/openscience/pdf/os\\_skills\\_wgreport\\_final.pdf#view=fit&pagemode=none](https://ec.europa.eu/research/openscience/pdf/os_skills_wgreport_final.pdf#view=fit&pagemode=none); ultima consultazione 5/9/2017.
9. SPARC Europe. *Open data and open policy in Europe*. Apeldoorn (NL): SPARC Europe; 2017. Disponibile all'indirizzo: <http://sparceurope.org/open-data-open-science-policy-europe/>; ultima consultazione 31/8/2017.
10. Galimberti P. Open access, Open science. L'Italia, un Paese in grave ritardo. *ROARS. Return On Academic Research* 2017: 3 pag. Disponibile all'indirizzo: <https://www.roars.it/online/open-access-open-science-litalia-un-Paese-in-grave-ritardo/>; ultima consultazione 31/08/2017.
11. MIUR. *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.
12. AISA. *Diritto di ripubblicazione in ambito scientifico*. Trento: AISA; 2016. Disponibile all'indirizzo: <http://aisa.sp.unipi.it/attivita/diritto-di-ripubblicazione-in-ambito-scientifico/>; ultima consultazione 31/8/2017.

## INDAGINE BISA: RISULTATI E ANALISI DEI DATI

Rosalia Ferrara (a), Pietro La Placa (b), Cristina Mancini (a), Elisabetta Poltronieri (c), Angela Vullo (d)  
(a) Servizio Conoscenza e Comunicazione Scientifica, Biblioteca, Istituto Superiore di Sanità, Roma  
(b) Servizio Editoria e Biblioteca, Istituto Zooprofilattico Sperimentale della Sicilia, Palermo  
(c) Ufficio Conoscenza e Comunicazione Scientifica, Attività Editoriali, Istituto Superiore di Sanità, Roma  
(d) Direzione Generale, Istituto Zooprofilattico Sperimentale della Sicilia, Palermo

### Questionario

L'indagine BISA (Bibliosan per la Scienza Aperta) si è avvalsa di un questionario (Appendice A), elaborato dal Gruppo di lavoro BISA e articolato in 26 domande, di cui 9 a risposta multipla, 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.

Il testo del questionario è stato caricato sulla piattaforma Google Forms e trasmesso online alle Direzioni scientifiche degli enti Bibliosan, attraverso il *workflow* del Ministero della Salute.

La rete Bibliosan è costituita da 60 enti:

- Istituto Superiore di Sanità (ISS);
- 46 Istituti di Ricovero e Cura a Carattere Scientifico (IRCCS);
- 10 Istituti Zooprofilattici Sperimentali (IZS);
- Agenzia Nazionale per i Servizi Sanitari Regionali (AGENAS);
- Agenzia Italiana del Farmaco (AIFA);
- Istituto Nazionale Assicurazione contro gli Infortuni sul Lavoro (INAIL).

La diffusione del questionario all'interno della rete Bibliosan è stata curata dai referenti delle rispettive biblioteche.

La somministrazione del questionario è avvenuta nel periodo 31 gennaio-5 marzo 2017.

Sono state registrate 2.433 risposte (con un tasso di risposta pari al 15%) provenienti da 58 su 60 degli enti coinvolti nell'indagine, cui non hanno aderito l'AIFA e l'INAIL (vedi Allegato al capitolo).

La popolazione di riferimento raggiunta è stata di 15.000 unità.

In Figura 1 sono riportate la frequenza delle risposte e le percentuali relative agli enti coinvolti nell'indagine. Le prime 31 istituzioni elencate rappresentano circa l'85% del totale delle risposte.

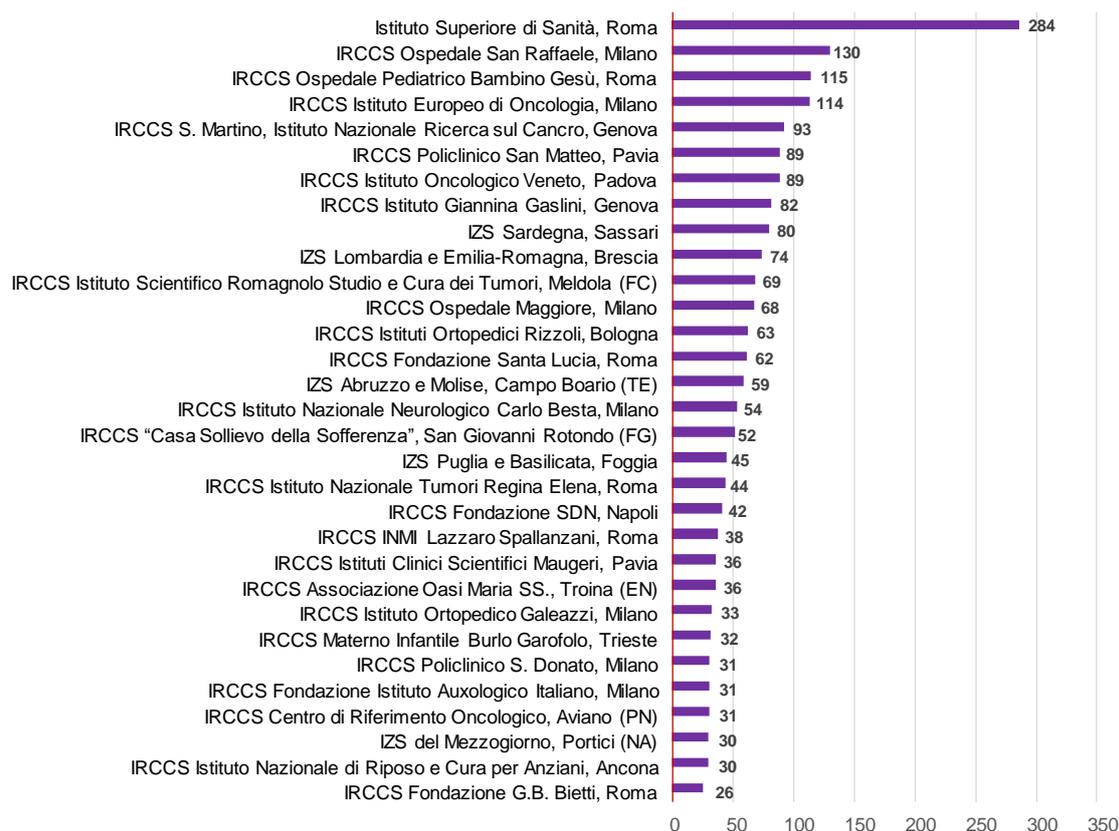


Figura 1. Istituzioni rispondenti all'Indagine BISA 2017 (85% delle risposte)

## Risultati

### Dati generali

La Tabella 1 riporta i dati identificativi dei rispondenti, relativamente ai ruoli professionali e alle fasce di età, nonché al genere dei partecipanti all'indagine, rispetto alla professione.

Tabella 1. Professione dei rispondenti per fasce di età e genere, Indagine BISA 2017

Professione	Età (anni)			Genere		Totale
	<30	30-50	>50	Femmina	Maschio	
Biologo	44	485	248	590	187	777
Medico	9	343	362	312	402	714
Altro	47	225	57	209	120	329
Veterinario	4	85	65	78	76	154
Tecnico	20	73	36	81	48	129
Psicologo	6	72	13	74	17	91
Chimico	4	41	21	43	23	66
Infermiere	17	34	19	59	11	70
Fisico	0	35	22	25	32	57
Bibliotecario/Documentalista	1	19	16	30	6	36
Editore scientifico (editor)	0	4	6	7	3	10
<b>Totale</b>	<b>152</b>	<b>1.416</b>	<b>865</b>	<b>1.508</b>	<b>925</b>	<b>2.433</b>

## Tipologia e formato dei dati

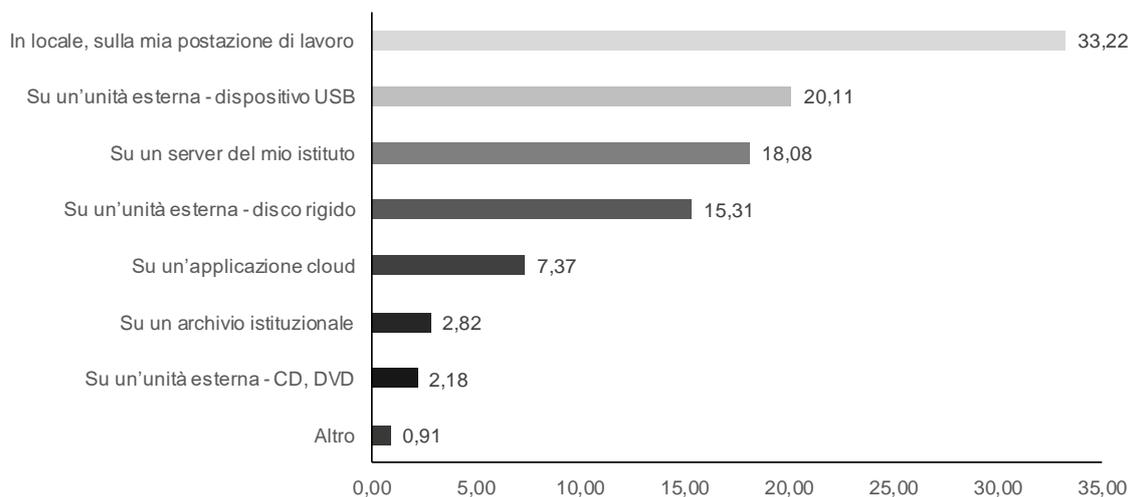
Dalla Tabella 2 si evince che la maggior parte dei formati di produzione dei dati, selezionati dai rispondenti al questionario tra le varie opzioni proposte, sono costituiti da tre categorie: documenti di testo (29%), grafici e immagini (25%) e fogli di calcolo (22%). L'impiego di differenti formati incide sull'archiviazione a lungo termine e, in base al volume dei dati generati, costituisce un indicatore per valutare le tecnologie da utilizzare ai fini di comprimere i dati, in modo da ottimizzare gli spazi di memoria.

**Tabella 2. Formato dei dati di ricerca, Indagine BISA 2017**

Formato dei dati	Frequenza	%
Documenti di testo	2.283	29,16
Grafici/immagini	1.929	24,64
Fogli di calcolo (es. fogli Excel)	1.709	21,83
Materiale multimediale (video, foto, audio)	812	10,37
Banche dati	777	9,92
Programmi applicativi (software)	292	3,73
Altro	27	0,34
<b>Totale</b>	<b>7.829</b>	<b>100</b>

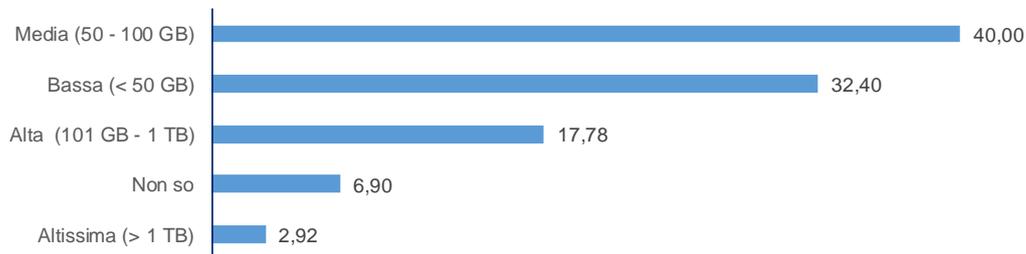
## Archiviazione dei dati

Si è ritenuto importante individuare i supporti utilizzati per salvare i propri dati di ricerca. In particolare, si trattava di scoprire se a tale scopo si utilizzano strumenti resi disponibili a livello istituzionale o strumenti personali, in uso al singolo ricercatore. Come mostrato in Figura 2, il risultato si è rivelato molto frammentario, con risposte eterogenee in cui prevale una preferenza per gli strumenti personali – postazione di lavoro in locale e dispositivi USB – che, sommati, superano il 50% delle risposte.



**Figura 2. Supporti di archiviazione dei dati della ricerca, Indagine BISA 2017**

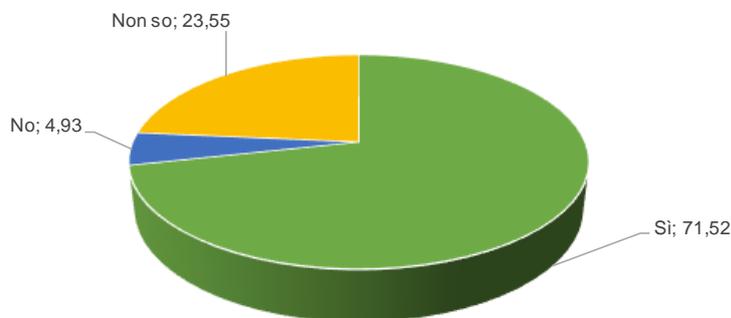
Un ulteriore fattore conoscitivo delle pratiche di archiviazione dei dati è stato il calcolo della media annuale del volume di dati gestiti. L'entità più consistente, in termini di fabbisogno di spazio, è risultata quella di fascia intermedia, pari a 50-100 GB. Se si associa a questo dato quello relativo al volume medio annuale dichiarato come basso, si raggiunge oltre il 70% delle risposte riferite a volumi medi e bassi (Figura 3).



**Figura 3. Quantità media annuale dei dati di ricerca gestiti, Indagine BISA 2017**

I dati ottenuti in risposta sono significativi ai fini di pianificare il fabbisogno di spazio di memoria occorrente per l'immagazzinamento e la conservazione dei dati.

Un fattore essenziale per la reperibilità dei dati della ricerca è il corredo di metadati ad essi associati, che dovrebbero rispondere a criteri condivisi di descrizione. Quasi tre quarti dei rispondenti ha dichiarato in proposito di attenersi a modelli descrittivi standard per la metadattazione. Anche in questo caso, a fronte di una presumibile varietà di sistemi disponibili, quali vocabolari e ontologie specifici in uso nelle varie discipline, occorrerebbe definire, in linee guida *ad hoc*, le convenzioni da adottare (Figura 4).



**Figura 4. Adozione di criteri standard di descrizione dei dati della ricerca, Indagine BISA 2017**

Quanto alla responsabilità della conservazione dei dati, occorrerebbe definire *a priori*, al momento di pianificare le attività di produzione, raccolta e gestione dei dati, i ruoli dei soggetti coinvolti nelle attività istituzionali. Va da sé, infatti, che individuare tali ruoli costituisce il prerequisito essenziale per stabilire poi chi sarà abilitato all'utilizzo dei dati. Se, come risulta, sono i singoli operatori a identificarsi in maggioranza (52%) come depositari della conservazione dei dati, ciò implica dover adottare e attenersi a standard comuni per il ri-uso dei dati e quindi prevedere l'investimento di risorse nella formazione di esperti in materia (Figura 5).

Un elemento cruciale nella pratica di gestione dei propri dati è costituito dall'eventualità di perderli per cause molteplici. Considerati gli alti costi connessi alla creazione dei dati, diviene imprescindibile, ai fini di razionalizzare la spesa, attivarsi preventivamente con contromisure di salvaguardia, come ad esempio procedure frequenti di salvataggio (*backup*). Tra i rispondenti al questionario il fenomeno sembra relativamente contenuto, attestandosi a circa il 20%, un dato che resta comunque da non sottovalutare (Tabella 3).

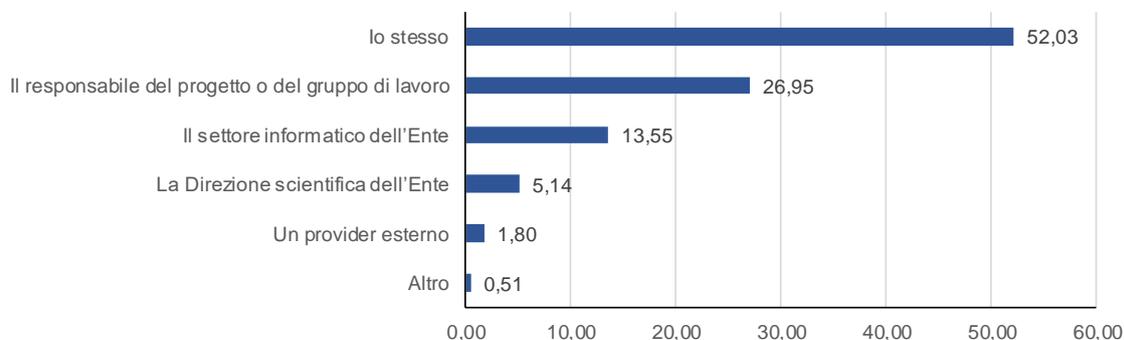


Figura 5. Soggetti responsabili della conservazione dei dati, Indagine BISA 2017

Tabella 3. Perdita dei dati di ricerca, Indagine BISA 2017

Perdita dei dati di ricerca	Frequenza	%
No	1.929	79,28
Si	504	20,72
<b>Totale</b>	<b>2.433</b>	<b>100</b>

Molteplici le motivazioni riportate dai rispondenti al questionario circa le specifiche cause che hanno comportato l'impossibilità improvvisa di accedere ai propri dati (Figura 6). In prima istanza si colloca la cancellazione accidentale con circa il 33% delle risposte. Questo risultato induce a ragionare sulle misure di archiviazione garantita dei dati (es. creazione di copie di sicurezza), considerati i costi che la loro produzione e gestione comporta.

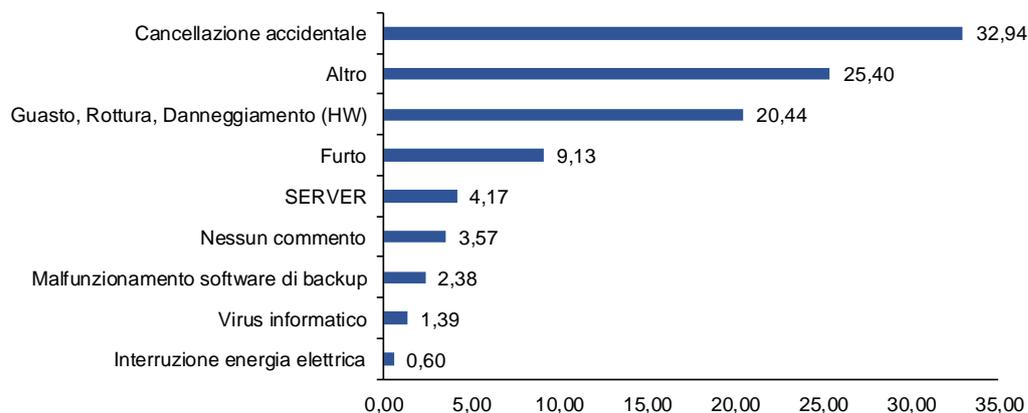


Figura 6. Perdita dei dati: cause, Indagine BISA 2017

## Aspetti etici e legali

Ai fini di misurare il grado di cooperazione nell'ambito della comunità scientifica, è stata formulata una domanda relativa all'eventuale utilizzo di dati esterni. La maggior parte dei rispondenti (58%) dichiara di non avvalersene (Tabella 4). È utile precisare che per dati esterni devono intendersi quelli che il ricercatore non ha generato o raccolto personalmente nel corso della sua ricerca. Le risposte riconducibili al "no" possono riferirsi a diverse motivazioni, quali:

- non esistono dati disponibili;
- esistono dati, ma non si possono utilizzare;
- non si conosce l'esistenza di dati che potrebbero essere utilizzati;
- difficoltà di citare correttamente i dati esterni e rischio di violazione involontaria delle norme di copyright.

Ad una successiva richiesta (cosiddetta domanda di controllo) per specificare con quale modalità venissero utilizzati i dati esterni, si è evidenziato che dei 1021 rispondenti che avevano affermato in precedenza di utilizzarli, 33 hanno poi dichiarato di non utilizzarli affatto; di conseguenza, gli effettivi rispondenti, utilizzatori di dati esterni, sono risultati 988. Questi ultimi, relativamente alle modalità di utilizzo dei dati esterni, hanno risposto, nel 74% dei casi, di acquisirli rielaborandoli e nel 26% senza operarne alcuna rielaborazione.

**Tabella 4. Utilizzo di dati esterni, Indagine BISA 2017**

Utilizzo di dati esterni	Frequenza	%
No	1.412	58,04
Si	1.021	41,96
<b>Totale</b>	<b>2.433</b>	<b>100</b>

Tra i fattori legati alla gestione dei dati della ricerca assume particolare rilevanza la questione del loro utilizzo legale; si è voluta quindi rilevare la consapevolezza dei ricercatori rispetto a questo delicato aspetto. Circa tre quarti dei rispondenti si è dichiarato al corrente dei risvolti legali associati all'uso di tali dati, come dimostrato dalla quasi totalità (99%) di risposte negative circa eventuali problemi legali occorsi nell'utilizzo di dati esterni (Tabella 5).

**Tabella 5. Problemi legali nell'utilizzo di dati esterni, Indagine BISA 2017**

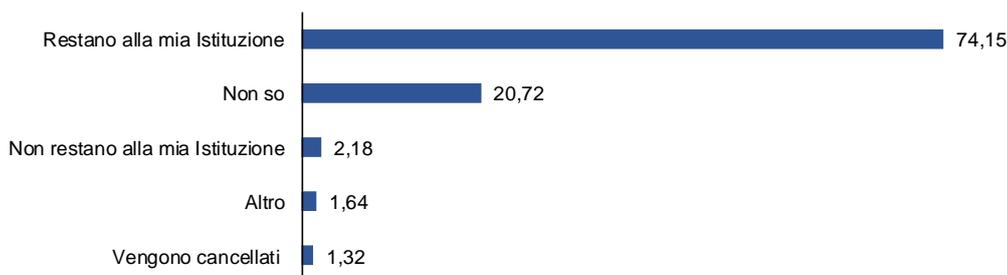
Problemi legali	Frequenza	%
No	2.419	99,42
Si	14	0,58
<b>Totale</b>	<b>2.433</b>	<b>100</b>

In merito alle questioni di ordine giuridico associate all'utilizzo di dati esterni, occorre predisporre strumenti (es. accordi con terze parti) che regolino l'accesso e il riuso degli stessi. A tal fine, gli uffici legali già attivi nelle varie istituzioni dovrebbero anche agire a supporto dei ricercatori per orientare le loro azioni nel rispetto della legalità. Tali strutture dovrebbero prevedere personale con profilo professionale specifico.

Un argomento di estremo interesse ai fini della preservazione e della riproducibilità dei dati della ricerca è quello della loro tracciabilità nel tempo. In particolare, si pone il problema dei

cosiddetti “dati orfani” che spinge ad indagare sul destino dei dati una volta che il ricercatore lascia l’istituzione di appartenenza per ragioni di fine rapporto o migra da un’istituzione all’altra.

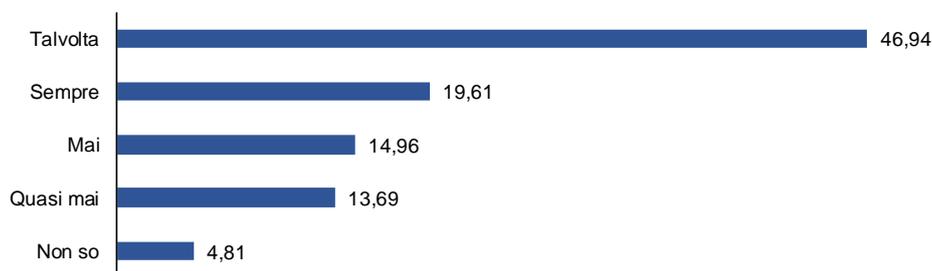
A tale proposito, il 74% dei rispondenti ha dichiarato che i dati rimangono archiviati presso l’istituzione di appartenenza. Sorprendentemente, nel 21% dei casi, si è registrata la risposta “non so” relativamente al destino dei dati. La motivazione di tale risposta potrebbe essere interpretata, da un lato, come quella di chi non si è mai posto la questione e, dall’altro, come quella di chi ha lavorato continuativamente presso un unico ente o di chi non risulta strutturato nell’organico dell’ente (Figura 7).



**Figura 7. Tracciabilità dei dati, Indagine BISA 2017**

Un punto focale della pratica di ricerca in ambito biomedico è rappresentato dalla gestione dei dati sensibili e/o riservati. In un contesto di scienza aperta, l’aspettativa è quella di un accesso ai dati *tout court*, nel segno della massima disponibilità alla condivisione e al riuso. Tuttavia, ragioni di etica della ricerca, in accordo con la normativa di salvaguardia della privacy, impongono di proteggere i dati di natura sensibile per evitare di risalire direttamente o indirettamente all’identità degli individui e rivelarne caratteristiche della sfera personale come lo stato di salute (es. dati di suscettibilità genetica). Sono orientate in tal senso anche le Raccomandazioni della Commissione Europea che prevedono l’accesso ai dati “quanto più aperto possibile e chiuso solo quando necessario,” cercando di bilanciare la cultura della trasparenza dei dati con il diritto alla riservatezza dei dati personali, anche a tutela degli interessi di eventuali terze parti (es. dati coperti da brevetto).

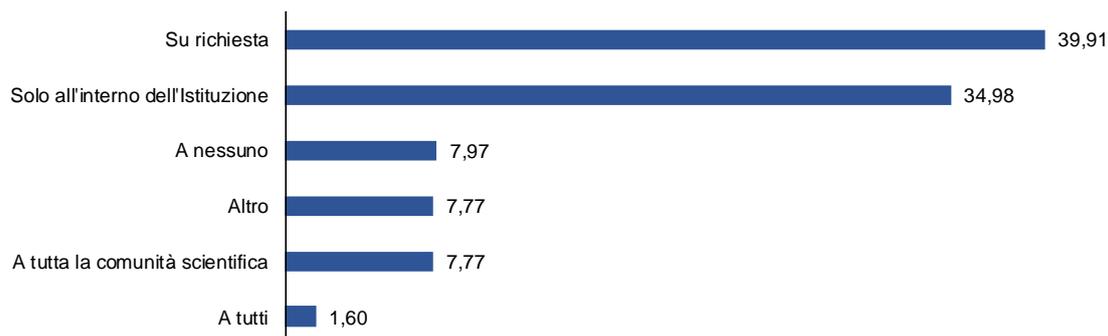
Alla domanda circa l’utilizzo di dati di ricerca sensibili o riservati, il 47% dei partecipanti al questionario ha dichiarato di trattarli, alcune volte (Figura 8). Questo risultato richiama l’attenzione sulla necessità per il ricercatore di riferirsi a linee guida che ne dettino il comportamento in materia o di poter contare sulla consulenza di esperti per evitare conseguenze di ordine legale.



**Figura 8. Produzione o utilizzo di dati sensibili o riservati, Indagine BISA 2017**

## Accessibilità e riuso

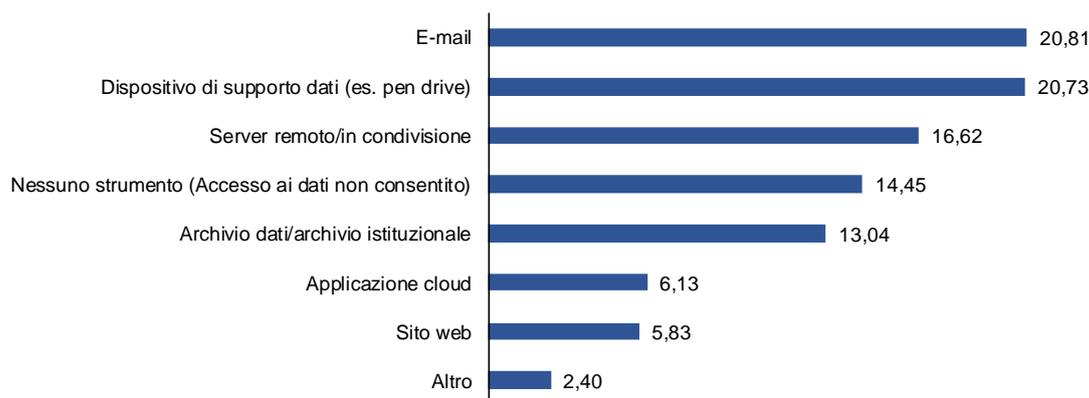
Con le domande di questa sezione del questionario l'intento era di individuare l'*audience* di destinazione dei dati della ricerca, con riferimento ai vari bacini di utenza, dall'ambito ristretto dell'istituzione di appartenenza alla platea più vasta del largo pubblico dei cittadini. Il presupposto è che i dati siano valutabili non solo per chi li produce, ma anche per chi intenda riutilizzarli. Le risposte più numerose si sono concentrate sull'accesso ai dati consentito "su richiesta" (40%) e sull'accesso previsto "solo all'interno dell'istituzione" (35%). Complessivamente, dunque, nel 75% dei casi, le risposte raccolte denunciano un'offerta selettiva di accesso, che ha diretto impatto sulla condivisione e sul riutilizzo dei dati stessi (Figura 9).



**Figura 9. Soggetti abilitati all'accesso ai dati, Indagine BISA 2017**

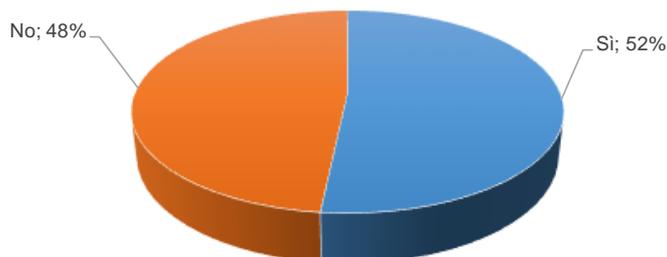
Probabilmente tale orientamento all'apertura "moderata" è motivato dalla natura stessa dei dati trattati che, se grezzi, cioè non anonimizzati, restano suscettibili di essere associati all'identità di singoli soggetti con pregiudizio della privacy.

Argomento correlato all'individuazione dei soggetti che beneficiano dell'accesso ai dati è quello delle modalità previste per accedervi. Nelle risposte in merito sono prevalsi lo scambio per e-mail (21%) e il ricorso alla memorizzazione su chiavetta USB (21%), il che induce a ritenere che restino privilegiati gli strumenti di contatto personali, anziché spazi virtuali istituzionali previsti *ad hoc*. Le soluzioni di archiviazione in remoto (*cloud*) e i siti web hanno totalizzato, infatti, percentuali di risposta inferiori, pari al 6% (Figura 10).



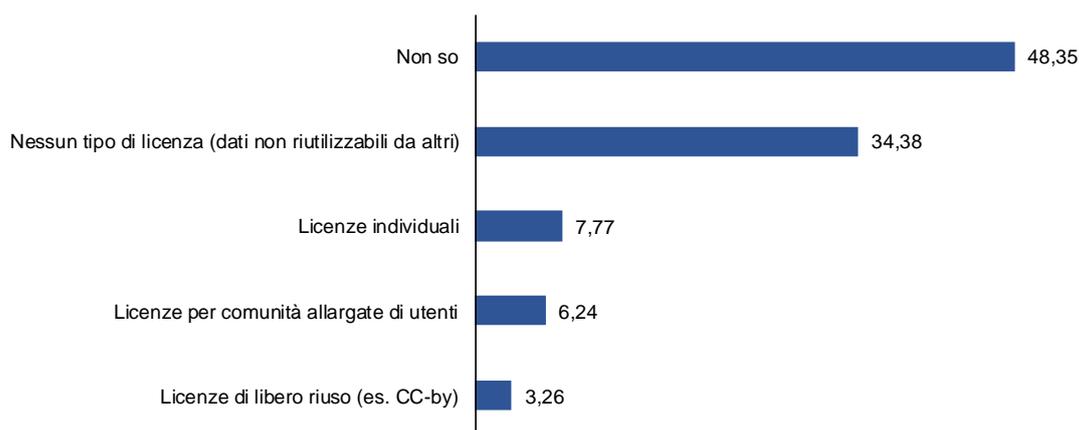
**Figura 10. Strumenti di accesso ai dati, Indagine BISA 2017**

Dall'accesso ai dati al riutilizzo degli stessi il passaggio è di quelli che denotano la piena comprensione o meno del concetto di disponibilità immediata dell'informazione scientifica, in un paradigma di scienza aperta. Il riuso è un principio che va ben oltre l'accesso ai dati, in quanto implica sapere di poter beneficiare di dati resi disponibili; circostanza possibile qualora il riuso sia stato predefinito in origine, grazie ad esempio all'applicazione di licenze libere come le *Creative Commons*, che regolano varie condizioni di riutilizzo determinate da chi detiene i diritti sui dati prodotti. L'analisi delle risposte ottenute sul tema del riuso dei dati ha mostrato percentuali quasi equivalenti tra chi lo autorizza (52%) e chi non è invece orientato a rendere disponibili i dati (48%), come mostrato in Figura 11.



**Figura 11. Disponibilità a rendere riutilizzabili i dati, Indagine BISA 2017**

I partecipanti al questionario hanno risposto anche in merito al tipo di accordi e di licenze d'uso adottati per regolare il riuso dei propri dati. Si trattava di sondare, a questo riguardo, il grado di consapevolezza dei ricercatori circa le implicazioni in merito. La maggiore proporzione di risposte si è concentrata sulla risposta "non so" (48%), a riprova del livello di incertezza ancora consistente in materia di strumenti regolatori del riuso dei dati e a dimostrazione di quanto sia imprescindibile pianificare e avviare iniziative di formazione mirate all'applicazione di strumenti giuridici atti a disciplinare una pratica di riuso corretta (Figura 12).



**Figura 12. Accordi e licenze per il riuso dei dati, Indagine BISA 2017**

Inoltre, considerando che il 61% dei rispondenti al questionario è costituito da biologi e medici e che è molto alta la percentuale (82,73%) derivata dalla somma delle risposte riferite alla mancanza di consapevolezza su accordi e licenze per il riuso dei dati (48,35) e alla non applicazione di alcun tipo di licenza (34,38), si ricava che occorre contare su politiche nazionali di impulso alle pratiche di condivisione e riutilizzo dei dati a beneficio dell'intera comunità scientifica.

Un sistema efficace per favorire la condivisione dei dati si può impiantare grazie ad un paniere di incentivi che rendono questa pratica convincente per i ricercatori, aiutandoli a guadagnare visibilità e impatto della ricerca. I due maggiori stimoli alla condivisione sono risultati quelli di tipo intellettuale, legati alla prospettiva di avanzamento della conoscenza scientifica (28%) e all'opportunità di stabilire nuovi rapporti di cooperazione per valorizzare il capitale della ricerca (25%). Con rilievo decisamente minore (6%) si sono posizionati gli incentivi economici (Figura 13). Quanto, invece, alle ragioni contrarie alla condivisione dei dati, le motivazioni prevalenti sono risultate il rischio di incorrere in interpretazioni improprie dei dati o addirittura in falsificazioni (34%) e quello di andare incontro a una violazione delle misure di sicurezza nel trattamento dell'informazione (26%), nel caso in cui, ad esempio, si trattino dati grezzi cioè non aggregati (Figura 14).



**Figura 13. Incentivi alla condivisione dei dati, Indagine BISA 2017**



**Figura 14. Disincentivi alla condivisione dei dati, Indagine BISA 2017**

## Infrastrutture e servizi

Il corredo dei servizi a supporto di una gestione efficace dei dati della ricerca prevede una serie di linee di intervento parallele che la singola istituzione dovrebbe attivare investendo risorse umane e finanziarie. Su questa tematica, i rispondenti al questionario hanno individuato in maniera marcata una serie di opzioni che potrebbero costituire un paniere integrato di misure, da attuare a livello centrale e in raccordo con quelle della rete Bibliosan (Figura 15). In particolare, per l'acquisizione di una infrastruttura tecnica (archivio di deposito dei dati, tecnologia avanzata, hardware e software aggiornati) si sono pronunciati il 35% dei rispondenti, mentre il 27% ha sottolineato la necessità di avvalersi di attività di formazione specifica per ruoli professionali gestionali. In ordine discendente di rilevanza si sono evidenziati il vantaggio derivato da un servizio legale (16%) e il sostegno di una struttura di *helpdesk* (12%).

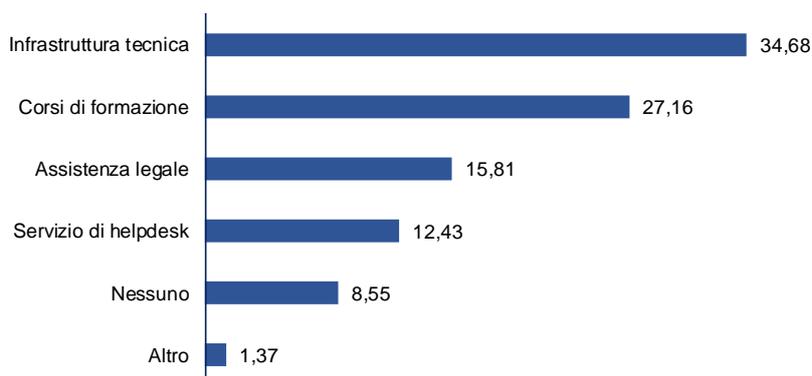


Figura 15. Tipologia di supporto richiesto per la gestione dei dati della ricerca, Indagine BISA 2017

Quanto alle politiche istituzionali da adottare per una valida gestione dei dati della ricerca, sono state individuate le azioni maggiormente attese dai propri vertici aziendali. In pratica, le risposte si sono polarizzate tra le due esigenze giudicate prioritarie: la redazione di documenti di indirizzo (politiche e linee guida) che definiscano ambiti di applicazioni, procedure e strumenti per il trattamento dei dati della ricerca (49%) e la previsione di profili professionali specifici per la gestione dei dati della ricerca (47%) (Tabella 6).

Tabella 6. Politiche istituzionali sulla gestione dei dati e iniziative attese, Indagine BISA 2017

Azione istituzionale attesa	Risposte	%
Linee guida/politiche di indirizzo	1.537	48,75
Previsione di una professionalità <i>ad hoc</i> per la gestione dei dati della ricerca	1.494	47,38
Altro	122	3,87
<b>Totale</b>	<b>3.153</b>	<b>100</b>

## Desiderata, idee

In chiusura di questionario è stata prevista una domanda a risposta libera, preordinata all'acquisizione di commenti e suggerimenti da parte dei partecipanti all'Indagine BISA circa i molteplici aspetti inerenti i dati della ricerca. Molte le considerazioni espresse, anche di valenza critica costruttiva, accompagnate ad auspici per una concreta e razionale gestione istituzionale del trattamento dei dati, soprattutto nel senso di una integrazione di competenze, servizi e piani di investimento da realizzare internamente a ciascun ente e in misura allargata con realtà istituzionali similari in ambito Bibliosan.

Le tematiche oggetto di riflessione da parte di rispondenti al questionario hanno spaziato dalle infrastrutture, intese come tecnologie e servizi a supporto della creazione e del mantenimento di piattaforme informatiche, alla necessità di dotarsi di politiche di indirizzo sulla gestione e condivisione dei dati.

Sul fronte delle infrastrutture, occorrono competenze di *data management* per definire, comprendere e gestire il dato della ricerca e il ciclo di vita dei dati, al fine di pianificarne la raccolta iniziale, prepararne l'analisi e provvedere infine alla pubblicazione dei risultati, favorendone la condivisione e la gestione a lungo termine. La cultura del saper gestire i dati consente ad altri ricercatori di scoprire, interpretare e riutilizzare i dati stessi, ne sostiene il valore consentendo ad altri di verificarli, una volta pubblicati, e ne facilita, infine, l'accesso e la conservazione a lungo termine. Inoltre, è da tenere presente che un numero crescente di agenzie di finanziamento, editori di riviste scientifiche e istituzioni di ricerca richiedono esplicitamente ai ricercatori una gestione controllata dei dati della ricerca, mediante l'uso di politiche di gestione e di piattaforme di collaborazione e condivisione realizzate con software specifici, per garantirne autenticità e riconoscimento della paternità intellettuale (diritti di copyright) oltre che conservazione, condivisione e diffusione.

L'aspetto della conservazione del dato e delle problematiche ad esso connesse sembra essere la criticità maggiormente rilevata. La prassi già consolidata in letteratura impone che, per tutto il corso della ricerca, sia necessario assicurarsi di archiviare i dati di ricerca in modo sicuro e di avere copie di *backup* in almeno tre spazi di memoria. L'archiviazione consigliata è l'unità di rete dell'istituzione; i computer personali e i computer portatili non dovrebbero essere utilizzati per memorizzare copie master dei dati e non si dovrebbero utilizzare dispositivi di memorizzazione esterni per l'archiviazione a lungo termine. Sembra che questo aspetto gestionale risulti quasi totalmente disatteso dall'istituzione di appartenenza dei ricercatori rispondenti al questionario. L'auspicio ricorrente tra i rispondenti, del resto, è l'adozione di sistemi istituzionali centralizzati, validati dalle direzioni aziendali, di facile utilizzo e basati su applicativi aggiornati, anche nell'intento di non ricorrere a piattaforme esterne a pagamento. Alla carenza marcata di strutture *ad hoc* per la conservazione e l'accessibilità dei dati in base a criteri condivisi fa eco, d'altra parte, l'esigenza di definire figure professionali specifiche interne alle istituzioni. Personale adeguatamente formato dovrebbe sovrintendere alla raccolta e all'analisi dei dati sensibili, alla gestione delle modalità di stoccaggio e alla delineazione di linee d'indirizzo circa la conservazione dei dati.

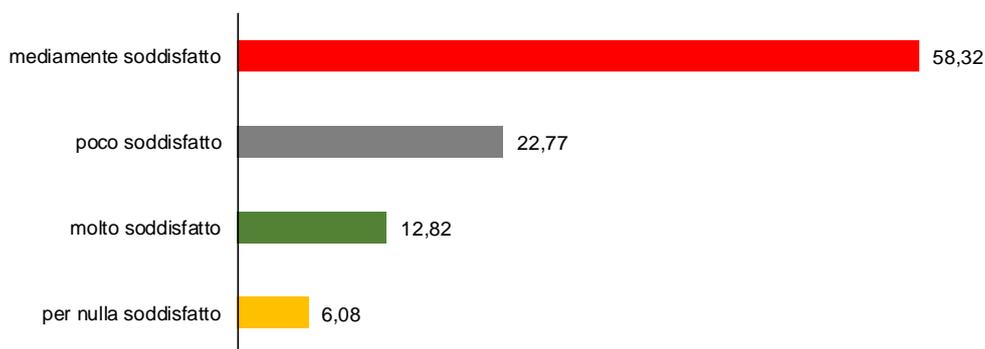
Per quel che concerne la condivisione dei dati, ferma restando la consapevolezza dei vantaggi legati a questa pratica, permane una certa reticenza a condividere i propri dati. Sono comunemente riconosciuti come elementi di spinta alla condivisione le seguenti motivazioni: trasparenza, rendimento dell'investimento per la ricerca che viene finanziata pubblicamente, supporto alla verifica e alla replica dei risultati originali, promozione di nuove ricerche, grazie alla sperimentazione di metodi nuovi o alternativi, incoraggiamento alla collaborazione tra i ricercatori, riduzione dei costi legati alla gestione dei dati, possibilità di evitare duplicazioni dei dati raccolti, protezione da dati non validati o fraudolenti, incremento di visibilità e impatto dei

progetti di ricerca. D'altra parte, la reticenza alla condivisione si spiega con una serie di considerazioni legate al trattamento e al riutilizzo dei dati, come rilevato nei commenti rilasciati in proposito dai rispondenti al questionario: "il tempo e lo sforzo per rendere i dati condivisi", "i rischi percepiti che accompagnano la perdita del controllo dei dati", "la presenza di informazioni riservate o sensibili nei set di dati", "la non chiara attribuzione dei diritti di proprietà dei dati" e infine la "mancanza di incentivi per la condivisione di dati".

Una notazione a margine va riservata ad alcune osservazioni critiche riportate dai rispondenti al questionario BISA. Sarebbe stato utile fornire una definizione puntuale di cosa si intendeva per "dati della ricerca", in quanto, a seconda dei casi, è possibile riferirsi a differenti tipologie di dati – dati primari grezzi, oppure dati aggregati, cioè strutturati in tabelle – o a entità quantitativamente diverse di dati (informazioni di piccola entità come il risultato di un saggio biochimico o molecolare o una tabella riportante un'analisi genomica).

Alcuni rispondenti inoltre hanno espresso l'esigenza di inquadrare la tematica dei dati della ricerca nel contesto più ampio dell'accesso aperto ai risultati della ricerca. In tal senso, l'aspettativa era di poter esplicitare modalità e perplessità legate all'adesione ai modelli di pubblicazione in accesso aperto.

Infine, il questionario ha rivelato il grado di soddisfazione dei rispondenti circa l'attuale stato dei servizi erogati dall'istituzione di appartenenza in materia di raccolta, conservazione, diffusione e protezione dei dati della ricerca. A questo riguardo (Figura 16), la maggioranza dei rispondenti, pari a più della metà (58%), ha espresso un grado medio di apprezzamento per le iniziative istituzionali, seguito da quanti si sono dichiarati poco soddisfatti (23%), da chi si è detto molto soddisfatto (13%) e infine da coloro che hanno affermato di non essere affatto soddisfatti (6%).



**Figura 16. Grado di soddisfazione dei servizi istituzionali circa la gestione dei dati della ricerca, Indagine BISA 2017**

## Conclusioni

Dall'indagine BISA sono emerse richieste esplicite e circostanziate che delineano una comunità di ricerca attiva e consapevole circa le problematiche e le aspettative legate alla gestione dei dati aperti della ricerca. Se da una parte rimane costante l'esigenza storica di un supporto guidato alla metodologia di ricerca ed accesso all'informazione scientifica, dall'altra s'impone

unanime il bisogno dei ricercatori di essere coinvolti in un processo istituzionale di organizzazione, conservazione e diffusione dei dati della ricerca.

Su questo scenario si innesta la previsione di una struttura lavorativa aziendale multidimensionale, basata su competenze distinte ma dinamicamente integrate. Gli enti Bibliosan potrebbero entrare nella logica di ridisegnare le attuali strutture di supporto alla ricerca, potenziando le attività delle loro biblioteche e centri di documentazione. Una volta dotate di personale adeguatamente formato, si potrebbero creare le condizioni per usufruire di idonei sistemi informativi, in stretto raccordo con altre componenti istituzionali (situazione di *co-working* attivo), quali centri elaborazione dati, uffici statistici e strutture di formazione, in modo da rendere ai ricercatori servizi effettivi, grazie alla collaborazione con i settori amministrativi e di controllo di gestione interni.

Entrambi i fattori, strutturali (strumentazioni e tecnologie) e gestionali (linee di indirizzo e coordinamento centrale) rappresentano dunque il volano per attivare la cultura della condivisione dei singoli ricercatori, al riparo di eccessive pratiche burocratiche. È dunque alla sensibilità dei vertici aziendali che occorre fare appello, anche nella prospettiva di raccogliere, valorizzare e diffondere il patrimonio di conoscenza scientifica generato nei singoli enti. 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 in campo biomedico.

Sembra questa, quindi, la logica di intervento più appropriata, anche per attuare concretamente le finalità di condivisione dei dati della ricerca promossa dall'Unione Europea per il progresso del sapere scientifico fondato su riproducibilità dei dati, metodi uniformi di descrizione, pratiche di riuso e conservazione a lungo termine. Una volta analizzate, le criticità di gestione dei dati della ricerca andrebbero eliminate attuando una regolamentazione che ne disciplini il trattamento e il riutilizzo nell'ambito della comunità scientifica nazionale.

Un'ipotesi potrebbe essere quella di istituire un database dei dati generati dalle ricerche e ripartiti in settori tematici. Tale progetto avrebbe la finalità di avviare la catalogazione del patrimonio di dati raccolto e di favorirne il riutilizzo e l'adeguata valorizzazione. Attuando una infrastruttura informatica deputata a ospitare i dati a livello nazionale, si offrirebbe, alle figure professionali in grado di interpretare i dati rielaborandoli, 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 per il progresso effettivo della conoscenza scientifica.

## Bibliografia di riferimento

- Australian National Data Service. *What is research data*. Canberra: ANDS; 2017. Disponibile all'indirizzo: [http://www.ands.org.au/\\_\\_data/assets/pdf\\_file/0006/731823/What-is-research-data.pdf](http://www.ands.org.au/__data/assets/pdf_file/0006/731823/What-is-research-data.pdf); ultima consultazione 30/11/2017.
- Data erasure. *Wikipedia, the free encyclopedia* (last edited on 19 October 2017). Disponibile all'indirizzo: [https://en.wikipedia.org/wiki/Data\\_erasure](https://en.wikipedia.org/wiki/Data_erasure); ultima consultazione 30/11/2017.
- DDI Alliance. *Data Documentation Initiative*. Data Documentation Initiative Alliance; 2017. Disponibile all'indirizzo: <http://www.ddialliance.org/>; ultima consultazione 30/11/2017.
- Digital Curation Centre. *About DMPOnline*. Edinburgh: University of Edinburgh; 2017. Disponibile all'indirizzo: [https://dmponline.dcc.ac.uk/about\\_us](https://dmponline.dcc.ac.uk/about_us); ultima consultazione 30/11/2017.

- Finch L, Webster J. *Caring for CDs and DVDs*. London: The British Library; 2008. (NPO Preservation Guidance. Preservation in Practice Series). Disponibile all'indirizzo: [http://www.bl.uk/aboutus/stratpolprog/collectioncare/faqs/cddvd/caring\\_for\\_cds\\_dvds.pdf](http://www.bl.uk/aboutus/stratpolprog/collectioncare/faqs/cddvd/caring_for_cds_dvds.pdf); ultima consultazione 30/11/2017.
- Government of Canada. *Digital preservation toolkit*. Ottawa: Government of Canada; 2017. Disponibile all'indirizzo: <http://canada.pch.gc.ca/eng/1443203443442>; ultima consultazione 30/11/2017.
- Inter-university Consortium for Political and Social Research. *Guide to social science data preparation and archiving: best practice throughout the data life cycle*. 5th ed. Ann Arbor, MI: ICPSR; 2012. Disponibile all'indirizzo: <http://www.icpsr.umich.edu/files/ICPSR/access/dataprep.pdf>; ultima consultazione 30/11/2017.
- Lyon L. *Dealing with data: roles, rights, responsibilities and relationships. Consultancy Report*. Bath: UKOLN - University of Bath; 2007. Disponibile all'indirizzo: [http://www.ukoln.ac.uk/ukoln/staff/e.j.lyon/reports/dealing\\_with\\_data\\_report-final.pdf](http://www.ukoln.ac.uk/ukoln/staff/e.j.lyon/reports/dealing_with_data_report-final.pdf); ultima consultazione 30/11/2017.
- MIT Libraries. *Data management: organize your files*. Cambridge, MA: Massachusetts Institute of Technology. Disponibile all'indirizzo: <http://libraries.mit.edu/data-management/store/organize>; ultima consultazione 30/11/2017.
- UK Data Service. *Document your data*. Colchester: University of Essex; 2017. Disponibile all'indirizzo: <http://www.data-archive.ac.uk/create-manage/document>; ultima consultazione 30/11/2017.
- UK Data Service. *Research data lifecycle*. Colchester: University of Essex; 2017. Disponibile all'indirizzo: <http://www.data-archive.ac.uk/create-manage/life-cycle>; ultima consultazione 30/11/2017.
- UK Data Service. *Store your data*. Colchester: University of Essex; 2017. Disponibile all'indirizzo: [www.data-archive.ac.uk/create-manage/storage](http://www.data-archive.ac.uk/create-manage/storage); ultima consultazione 30/11/2017.
- UK Data Service. *Citing data*. Colchester: University of Essex; 2017. Disponibile all'indirizzo: [ukdataservice.ac.uk/use-data/citing-data.aspx](http://ukdataservice.ac.uk/use-data/citing-data.aspx); ultima consultazione 30/11/2017.
- University of California, Digital Library. *About DMPTool*. Oakland, CA: California Digital Library. Disponibile all'indirizzo: <https://dmptool.org/about>; ultima consultazione 30/11/2017.
- Vellani S, Lepore V. FAIR data management: workshop di Research Data Alliance (RDA) a Firenze. In: *Workshop Fair Data Management: Best practices and open issues*. Firenze, 14-15 novembre 2016. p. 34-41. Disponibile all'indirizzo:

## Allegato al capitolo

### Istituzioni a cui è stato somministrato il questionario BISA

#	Istituzione di appartenenza	Risposte	
		n.	%
1	Istituto Superiore di Sanità, Roma	284	11,67
2	IRCCS Ospedale San Raffaele, Milano	130	5,34
3	IRCCS Ospedale Pediatrico Bambino Gesù, Roma	115	4,73
4	IRCCS Istituto Europeo di Oncologia, Milano	114	4,69
5	IRCCS San Martino, Istituto Nazionale per la Ricerca sul Cancro, Genova	93	3,82
6	IRCCS Istituto Oncologico Veneto, Padova	89	3,66
7	IRCCS Policlinico San Matteo, Pavia	89	3,66
8	IRCCS Istituto "Giannina Gaslini", Genova	82	3,37
9	Istituto Zooprofilattico Sperimentale delle Venezie, Padova	80	3,29
10	Istituto Zooprofilattico Sperimentale della Sardegna, Sassari	74	3,04
11	IRCCS Istituto Scientifico Romagnolo per lo Studio e la Cura dei Tumori, Meldola (FC)	69	2,84
12	IRCCS Ospedale Maggiore, Milano	68	2,79
13	IRCCS Istituti Ortopedici Rizzoli, Bologna	63	2,59
14	IRCCS Fondazione Santa Lucia, Roma	62	2,55
15	Istituto Zooprofilattico Sperimentale della Lombardia e dell'Emilia, Brescia	59	2,42
16	IRCCS Istituto Nazionale Neurologico "Carlo Besta", Milano	54	2,22
17	IRCCS Ospedale Generale Regionale "Casa Sollievo della Sofferenza", S. Giovanni Rotondo (FG)	52	2,14
18	Istituto Zooprofilattico Sperimentale della Sicilia, Palermo	45	1,85
19	IRCCS Istituto Regina Elena per lo Studio e la Cura dei Tumori, Roma	44	1,81
20	IRCCS Fondazione SDN, Napoli	42	1,73
21	IRCCS Istituto Nazionale per le Malattie Infettive Lazzaro Spallanzani, Roma	38	1,56
22	IRCCS Associazione Oasi Maria SS., Troina (EN)	36	1,48
23	IRCCS Istituti Clinici Scientifici Maugeri Spa SB, Pavia	36	1,48
24	IRCCS Istituto Ortopedico Galeazzi, Milano	33	1,36
25	IRCCS materno infantile "Burlo Garofolo", Trieste	32	1,32
26	IRCCS Centro di Riferimento Oncologico di Basilicata di Rionero in Vulture, Potenza	31	1,27
27	IRCCS Fondazione Istituto Auxologico Italiano, Milano	31	1,27
28	IRCCS Policlinico San Donato, Milano	31	1,27
29	IRCCS Istituto Nazionale di Riposo e Cura per Anziani Vittorio Emanuele II, Ancona	30	1,23
30	Istituto Zooprofilattico Sperimentale del Mezzogiorno, Portici (NA)	30	1,23
31	IRCCS Fondazione G.B. Bietti, Roma	26	1,07
32	IRCCS Istituto Dermatologico dell'Immacolata (IDI), Roma	24	0,99
33	Istituto Zooprofilattico Sperimentale dell'Abruzzo e del Molise, Campo Boario (TE)	23	0,95
34	IRCCS Istituto Tumori "Giovanni Paolo II", Bari	22	0,90
35	IRCCS Istituto Nazionale per lo Studio e la Cura dei Tumori "Fondazione G. Pascale", Napoli	21	0,86
36	IRCCS San Raffaele Pisana, Roma	19	0,78
37	Istituto Zooprofilattico Sperimentale del Piemonte, Liguria e Valle d'Aosta, Torino	18	0,74
38	IRCCS Centro di Riferimento Oncologico Istituto Nazionale Tumori, Aviano (PN)	17	0,70
39	IRCCS Centro Neurolesi "Bonino Pulejo", Messina	17	0,70
40	IRCCS Ospedale San Camillo, Venezia	17	0,70
41	IRCCS Arcispedale Santa Maria Nuova, Reggio Emilia	15	0,62
42	IRCCS Istituto Clinico Humanitas, Rozzano (MI)	14	0,58
43	IRCCS Centro Cardiologico Monzino, Milano	13	0,53
44	IRCCS Centro San Giovanni di Dio, Fatebenefratelli, Brescia	13	0,53
45	IRCCS Ente Ospedaliero Specializzato in Gastroenterologia "Saverio de Bellis", Castellana Grotte	13	0,53
46	IRCCS Istituto "Eugenio Medea", Associazione "La Nostra Famiglia", Bosisio Parini (LC)	13	0,53
47	IRCCS Istituto Nazionale per lo Studio e la Cura dei Tumori, Milano	13	0,53
48	IRCCS Istituto Neurologico Mediterraneo Neuromed, Pozzilli (IS)	13	0,53
49	IRCCS Istituto S. Maria e S. Gallicano, Roma	12	0,49
50	IRCCS Istituto delle Scienze Neurologiche, Bologna	11	0,45
51	Istituto Zooprofilattico Sperimentale dell'Umbria e delle Marche, Perugia	11	0,45
52	IRCCS Fondazione "Istituto Neurologico Casimiro Mondino", Pavia	10	0,41
53	IRCCS Fondazione Stella Maris, Calambrone (Pisa)	10	0,41
54	Istituto Zooprofilattico Sperimentale del Lazio e della Toscana, Roma	9	0,37
55	Istituto Zooprofilattico Sperimentale della Puglia e della Basilicata, Foggia	8	0,33
56	IRCCS Fondazione Don Carlo Gnocchi, Milano	7	0,29
57	AGENAS, Agenzia Servizi Sanitari Regionali, Roma	4	0,16
58	IRCCS Multimeditica Holding Spa, Milano	4	0,16
	<b>totale</b>	<b>2.433</b>	<b>100</b>

## RIFLESSIONI SUGLI ESITI DELLE INIZIATIVE BISA E PROGETTO E-INFRASTRUCTURES AUSTRIA

Paolo Budroni

*Library and Archive Services, Department e-Infrastructures, University of Vienna, Austria*

### ***e-Infrastructures Austria***

Il progetto *e-Infrastructures Austria* ([www.e-infrastructures.at/en/home](http://www.e-infrastructures.at/en/home)) nasce nel 2014 e vede nella sua prima fase triennale la partecipazione di 26 partner. *e-Infrastructures Austria* riunisce in questo sforzo collettivo a livello nazionale tutti gli atenei austriaci (tranne uno), nonché altre 5 istituzioni fra le quali la Biblioteca Nazionale austriaca. In qualità di Direttore dell'intero programma triennale 2014-2016 e attuale Presidente dell'assemblea generale di *e-Infrastructures Austria*, l'autore del presente contributo era affiancato da uno staff manageriale di altre 3 persone, tutte afferenti all'Università di Vienna: Barbara Sánchez Solís (Biblioteca), Raman Ganguly e José Luis Preza (Centro di calcolo). A livello operativo, questo gruppo dirigenziale era sostenuto da 12 responsabili di sottoprogetto (<http://e-infrastructures.at/en/the-project/work-package-clusters/>). Al Progetto era inoltre associato altro personale che, su richiesta, collaborava alla finalizzazione di progetti e sottoprogetti. A uno di questi gruppi di lavoro, il *Work Package Cluster-B*, era stato affidato il compito di condurre in Austria, a livello nazionale, un'indagine i cui risultati sono stati pubblicati (1).

I risultati dell'indagine *e-Infrastructures Austria* rispecchiano quelli ottenuti in un sondaggio svolto in precedenza a livello europeo, nell'ambito del progetto *Permanent Access to the Records of Science in Europe* (PARSE) (2) e non soltanto confermano alcune previsioni formulate dai colleghi di *e-Infrastructures Austria*, ma si rivelano conformi agli sforzi messi attualmente in atto dalla Commissione Europea nell'ambito dello *European Open Science Cloud* (EOSC).

Dall'indagine austriaca si evince una serie di metodologie per la gestione dei dati della ricerca in Austria che permette di:

1. comprendere come iniziare il riorientamento delle risorse da dedicare;
2. pianificare misure per migliorare le infrastrutture di supporto per la gestione dei dati;
3. identificare le migliorie da apportare ai servizi attualmente in uso;
4. individuare i servizi da offrire in futuro.

I rispondenti all'indagine, svolta nel 2015, sono stati 3026, corrispondente al 9% dei destinatari del sondaggio, cui è stato inviato il questionario. Nel confronto fra l'*e-Infrastructures Austria* e il Questionario elaborato dal Gruppo di lavoro BISA sono emersi risultati di particolare importanza per lo sviluppo di una visione strategica comune in ambito EOSC, utili per delineare direttrici di sviluppo per la gestione dei dati di ricerca e delle relative infrastrutture. Le aree esplorate sono le seguenti:

- aspettative dei ricercatori austriaci nei confronti della propria istituzione;
- tipologia di dati e di formati;
- archiviazione dei dati, *backup* e perdita di dati;
- aspetti legali ed etici;
- accessibilità e riuso;
- infrastrutture e servizi.

## Aspettative dei ricercatori austriaci nei confronti della propria istituzione

In generale, nell'indagine *e-Infrastructures Austria* i colleghi austriaci hanno focalizzato le proprie aspettative per lo più su tre aree. Le loro richieste, infatti, puntano soprattutto sull'offerta di personale qualificato di supporto e sullo sviluppo di linee guida e di politiche per la gestione dei dati di ricerca. Essi inoltre dichiarano che gli oneri amministrativi nella gestione dei dati sono numerosi e che si riscontra una carenza di corsi di formazione. Hanno anche indicato che sarebbe opportuno offrire maggiore assistenza tecnica e consulenza per la soluzione di questioni di carattere legale ed etico.

### Tipologie di dati e di formati

La maggior parte dei ricercatori (97%) dichiara di produrre dati di ricerca sotto forma di *file* di testo non strutturati, mentre poco meno (81%) risponde di produrre grafici e tabelle. In terza posizione, con il 67%, si pone l'utilizzo dei fogli di lavoro (*spreadsheet*) (Figura 1). Se da un lato non sorprende l'alta presenza di rispondenti afferenti a discipline tecniche – che dichiara di generare *source code* e dati configurati – si riscontra un'alta frequenza di rispondenti in merito anche da parte di chi opera in ambiti quali la medicina e l'ampio spettro delle scienze umanistiche. Oltre a ciò, la maggior parte dei ricercatori dichiara di produrre tre quarti del proprio volume di ricerca in formato digitale. I dati in formato analogico vengono utilizzati solo da un ventesimo dei partecipanti all'indagine (soprattutto in ambito umanistico).

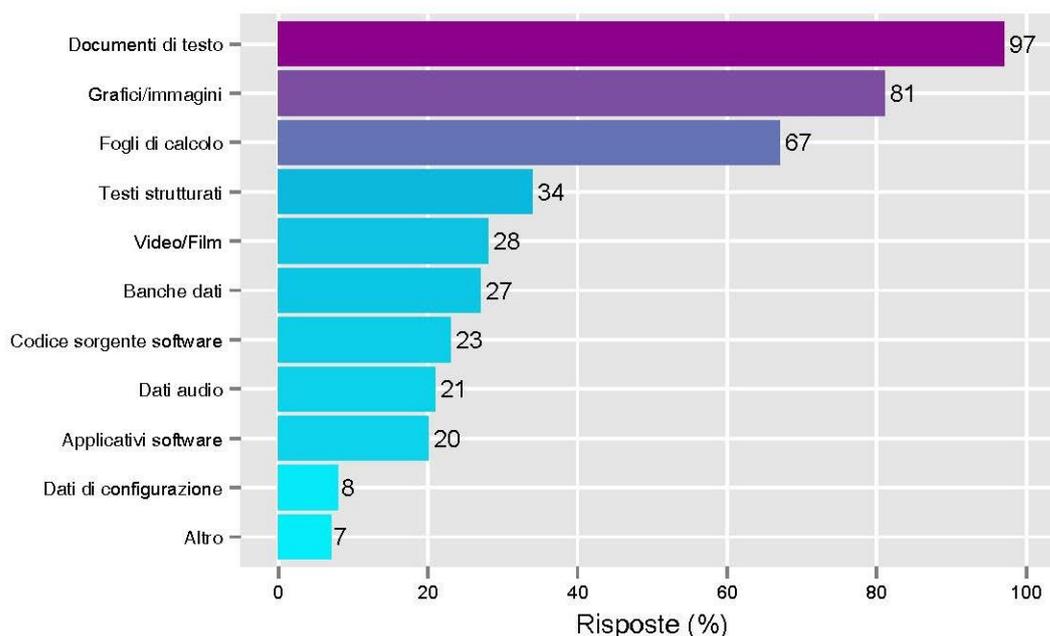


Figura 1. Tipologie di dati e formati, e-Infrastructures Austria 2015

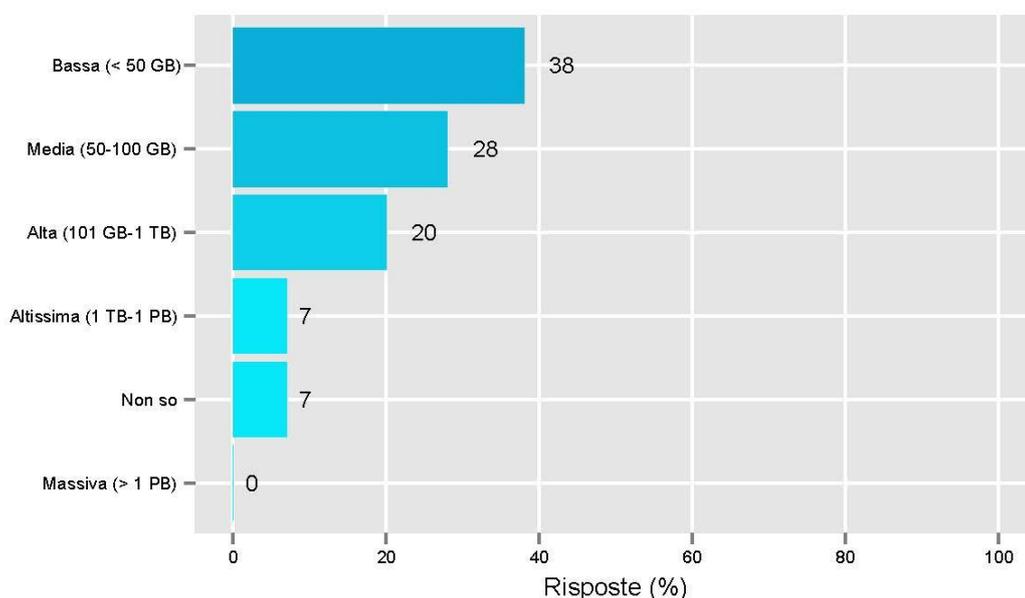
## Archiviazione dei dati, *backup* e perdita di dati

Per quel che riguarda l'archiviazione dei dati, la maggior parte dei partecipanti (54%) dichiara di optare per soluzioni di tipo *multiple storage*, il che significa un impiego simultaneo di postazioni di lavoro, del computer di casa, come anche di chiavi USB, *flash drive* e *hard drive* esterni.

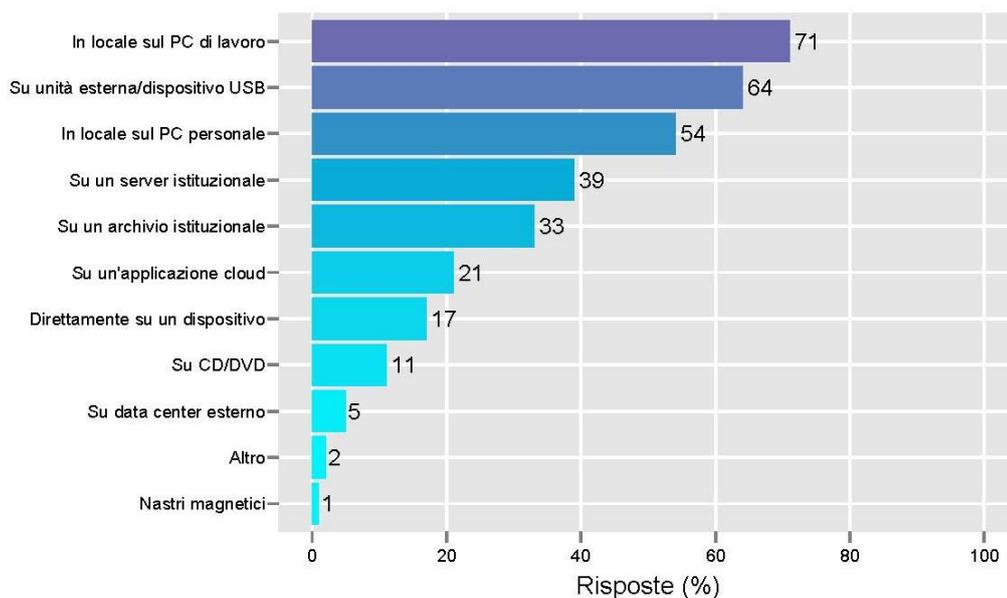
A margine di questi risultati, è interessante notare che:

1. circa la capacità di *storage* richiesta, due terzi dei partecipanti (66%) dichiarano di necessitare al massimo di 100 GB/anno o anche meno (Figura 2). Questo dato varia nel caso dei colleghi che lavorano presso una delle cinque università austriache ad indirizzo artistico e di quanti operano in ambito medico (tre università) che hanno espresso una necessità di spazio maggiore;
2. *big data is not an issue*: questo enunciato riassume sinteticamente uno dei risultati più rilevanti dell'indagine *e-Infrastructures* svolta in Austria. Soltanto il 7% dei colleghi austriaci lavora veramente con *big data*, ponendosi il problema di come gestirli – e la maggior parte di essi dichiara di sapere come e dove gestirli (*vedi* il paragrafo *Infrastrutture e servizi*). La sfida è rappresentata invece da come rispondere alle esigenze di quel folto gruppo di ricercatori (86%) che deve gestire una serie molto ampia di *dataset* sparsi e suddivisi fra migliaia di server, postazioni, USB *flash drive et similia* (Figura 3).

Nella maggior parte dei casi (90%), il produttore dei dati dichiara di essere anche la persona responsabile della loro archiviazione; due terzi di questo gruppo di persone dichiara di descrivere i dati in maniera non coerente e approssimativa; un fatto, questo, che incide notevolmente e in maniera negativa sulle future strategie di ricerca e riutilizzo dei dati. Va considerato, a tale proposito, che le attuali attività avviate a livello europeo, (EOSC, GO-FAIR, i *Data Management Plan* previsti dai programmi H2020) prevedono strategie di ricerca finalizzate al preciso recupero dei dati e al loro accesso in via permanente (il che non significa necessariamente “libero”), come anche modalità di riproducibilità e di riutilizzo. Le finalità di tali iniziative, pertanto, non sembrano allinearsi alle attuali prassi di gestione dei dati, come fin qui riportate. Quanto al fenomeno della perdita di dati, un terzo dei partecipanti all'indagine dichiara di averlo subito.



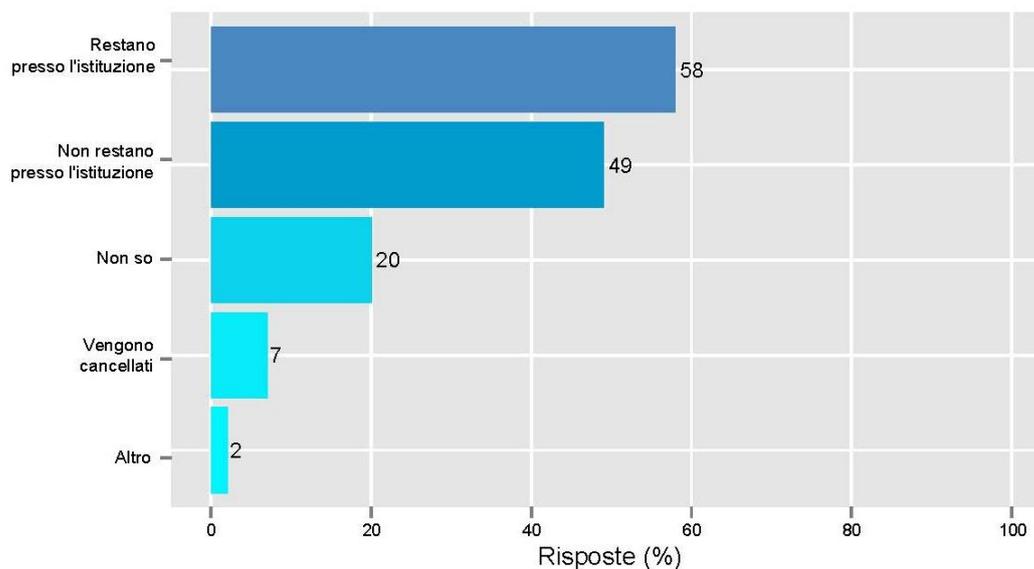
**Figura 2. Stima della media annuale del volume totale dei dati di ricerca sulla base dello spazio di archiviazione occorrente, e-Infrastructures Austria 2015**



**Figura 3. Localizzazione dei dati, e-Infrastructures Austria 2015**

### Aspetti legali ed etici

Se, da un lato, un terzo dei ricercatori dichiara di non aver mai dovuto affrontare problematiche di tipo legale (e di non aver mai vissuto situazioni legali ambigue) durante l'uso di dati non propri, dall'altro, un quinto dei dichiaranti precisa di aver sperimentato almeno una volta situazioni confuse dal punto di vista legale. Nel momento in cui si cambia posto di lavoro e ci si trasferisce in un'altra istituzione, si riscontra che in media il 7% dei dati viene distrutto e che nel 58% dei casi i dati rimangono presso l'istituzione di origine. Il resto dei dati (49%) vengono trasferiti presso la nuova istituzione (Figura 4).



**Figura 4. Posizionamento dei dati al momento di lasciare la propria istituzione, e-Infrastructures Austria 2015**

Circa un settimo dei ricercatori usa dati considerati sensibili e questa proporzione varia verso l'alto soprattutto in ambito medico. In questo caso, il 40% dei ricercatori dichiara di utilizzare dati ritenuti sensibili.

## Accessibilità e riuso

La maggior parte dei ricercatori dichiara che l'uso di dati esterni (generati da terzi) è di importanza cruciale nel proprio ambito di lavoro, mentre un quarto dei ricercatori dichiara di non utilizzare dati provenienti da terzi. L'utilizzo dei propri dati da parte di altri è previsto solo in maniera limitata. Oltre la metà (57%) dei partecipanti all'indagine consente l'accesso ai propri dati solo su richiesta e soltanto un decimo circa dei ricercatori dichiara di mettere i propri dati a disposizione di tutti. Infine, circa un decimo dei partecipanti non permette alcun tipo di accesso (Figura 5).

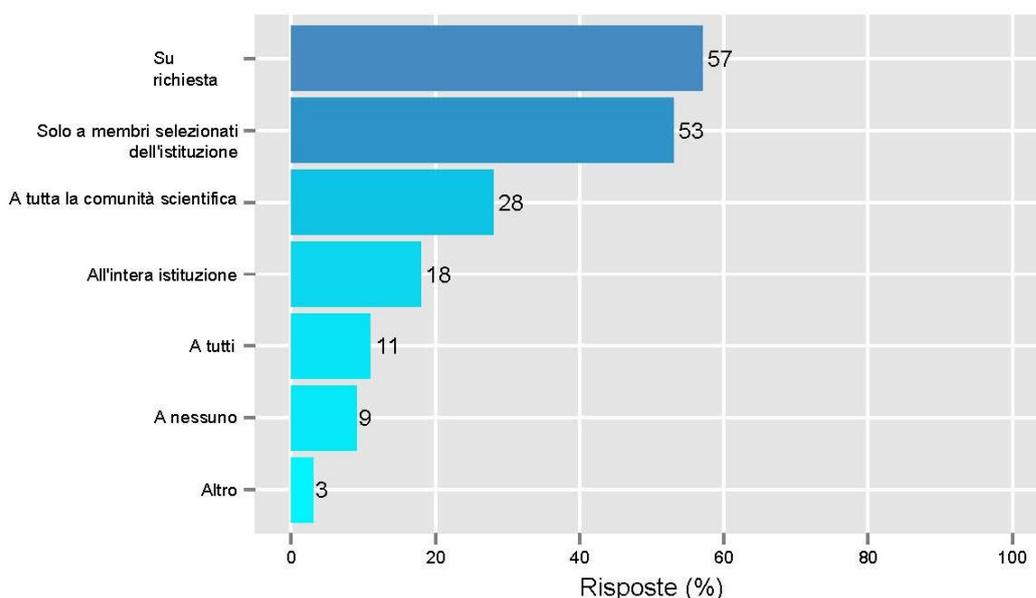
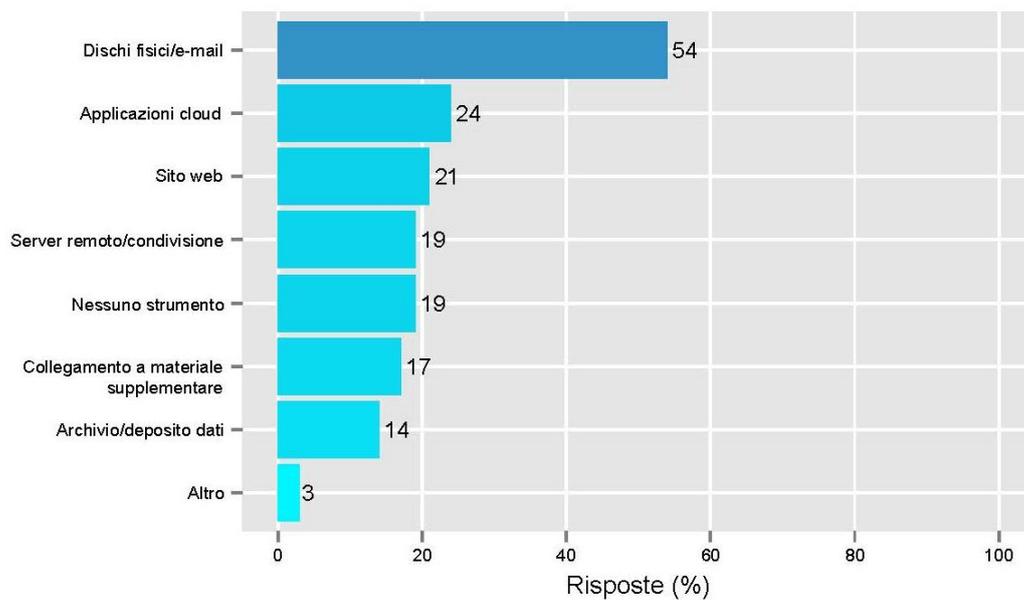


Figura 5. Accesso ai dati, e-Infrastructures Austria 2015

Dalle modalità di accesso concesse traspare anche la volatilità dei dati. Gran parte dei ricercatori (54%) dichiara di poterli rendere accessibili tramite l'utilizzo di *disk drive* o attraverso messaggi di posta elettronica. Le altre soluzioni scelte – *cloud* (24%), *remote server/ share server* (19%), sito web (21%) – non migliorano la situazione di volatilità.

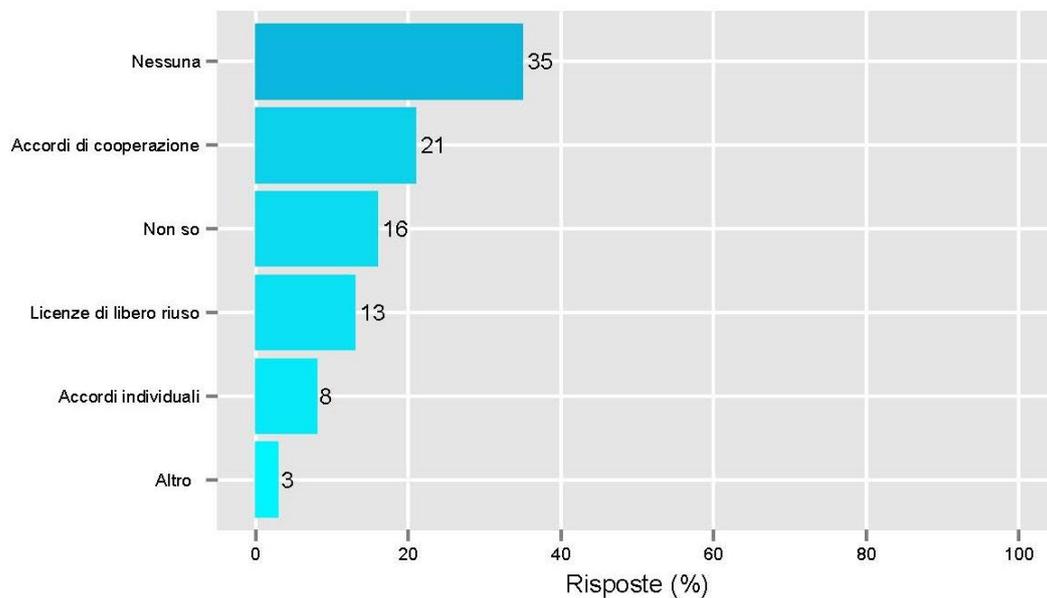
Le soluzioni stabili, come per esempio i *data archives/repositories*, sono utilizzate solo da un settimo (14%) dei ricercatori (Figura 6).

Questi risultati invitano a riflettere circa l'efficacia delle attuali campagne svolte in ambito di *accesso aperto*, iniziative che accentrano notevoli risorse e che si attuano ormai da tre lustri. Considerato che circa un terzo dei rispondenti consente il riutilizzo dei propri dati, l'atteggiamento varia in relazione alle diverse discipline: il riuso dei dati, infatti, è previsto maggiormente nel caso di aree tematiche quali geografia, biologia e chimica, mentre è accordato con minore frequenza nei casi di medicina, scienze sociali e in ambito umanistico.



**Figura 6. Modalità di accesso ai dati, e-Infrastructures Austria 2015**

Bisogna comunque evidenziare che in tre discipline (medicina, scienze sociali e ingegneria) risulta evidente che gli ostacoli di natura legale (anche presunti tali) vengono considerati il maggiore impedimento per il riutilizzo dei dati. Si rileva inoltre che solo un terzo dei partecipanti (29%) dichiara di condividere i dati sulla base di accordi strutturati e concordati previa consulenza legale. Il 66% non condivide i dati sulla base di regole riconducibili a leggi o a documenti di *policy* (Figura 7).



**Figura 7. Condizioni legali di riutilizzo, e-Infrastructures Austria 2015**

Questo risultato, a prima vista insignificante, presenta risvolti non indifferenti, perché dimostra che la richiesta dei ricercatori circa un maggior supporto in ambito legale si associa alla mancanza di un approccio organico da parte delle istituzioni di appartenenza. Ne potrebbe derivare, e probabilmente è così, che le istituzioni perdano il controllo effettivo su quello che è uno dei loro maggiori capitali, e cioè la propria produzione scientifica. Si perde quindi la sovranità sui contenuti, sui relativi metadati (che contengono la parte maggiore dell'intelligenza prodotta) e si perde il controllo sui *link* (collegamenti e relazioni) fra i differenti *set* di dati. Se a questo quadro aggiungiamo lo sconcertante scenario della perdita di sovranità nel settore delle pubblicazioni a tutto vantaggio economico dei editori commerciali, si nota come la seconda grande richiesta di linee guida e documenti di *policy*, espressa dai ricercatori, faccia rilevare carenza di *leadership* ai massimi vertici degli atenei.

In sede di Commissione Europea, si è consapevoli delle carenze in ambito di politiche sulla gestione dei dati. Per questa ragione, è stato anche approvato il progetto *Leaders Activating Research Networks* (LEARN <http://learn-rdm.eu/en/about/>), conclusosi nel mese di giugno 2017, che ha coinvolto non solo la vasta area scientifica europea, ma anche l'intera area iberoamericana e i Caraibi (tramite l'appoggio di UN-ECLA – *United Nations Economic Commission for Latin America and the Caribbean*). Il progetto ha prodotto materiali mirati alla gestione dei dati di ricerca: un *toolkit*, 20 raccomandazioni, e un'indagine sulle politiche in Europa. Il nucleo di quanto prodotto dal progetto sono un modello di *Research Data Management Policy* e linee guida. Prima della sua pubblicazione, il modello di *policy* è transitato attraverso diversi processi di valutazione, svolti fra l'altro da colleghi del *Digital Curation Centre* di Edimburgo e da colleghi delle seguenti istituzioni: Università di Cambridge, di Helsinki, dal Ministero della Scienza finlandese e dalla Commissione Europea. L'intero *toolkit* ha poi superato due revisioni in sede di Commissione Europea. La versione italiana della *model policy* è stata in seguito rimodulata dall'Università statale di Milano e dall'Università Ca' Foscari di Venezia per essere presentata in occasione dell'*Open Science Day*, organizzato a Milano dall'Università di Torino nel mese di maggio 2017.

Sul fronte degli incentivi alla pratica della condivisione dei dati, è importante evidenziare che questi non devono essere necessariamente di natura economica; reputazione e visibilità sono ritenuti cruciali, soprattutto per favorire avanzamenti di carriera, come anche per l'acquisizione di ulteriori progetti e disponibilità di nuove risorse, ai fini di ampliare i *network* e le collaborazioni e per fungere da punto di riferimento in ambito professionale.

## Infrastrutture e servizi

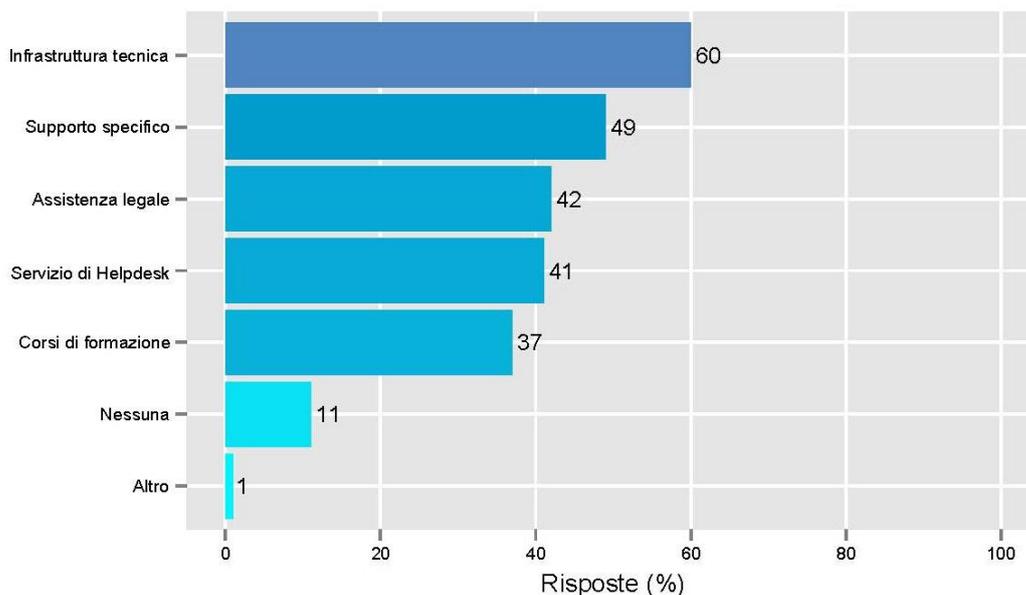
Alla domanda circa quale possa essere il contenitore privilegiato per l'archiviazione dei dati, i ricercatori rispondono, in genere, che non hanno chiare preferenze. Volendo ciononostante creare delle classifiche per poter generare una sorta di *hit-parade* fra le preferenze a scelta multipla, si giungerebbe a questa conclusione (Tabella 1).

**Tabella 1. Tipologia di archivi per il deposito dei dati, e-Infrastructures Austria 2015**

Archivio di dati preferito	Risposte (%)
Archivio disciplinare internazionale	47
Archivio istituzionale centralizzato	37
Archivio multidisciplinare internazionale	28
Archivio disciplinare nazionale	21

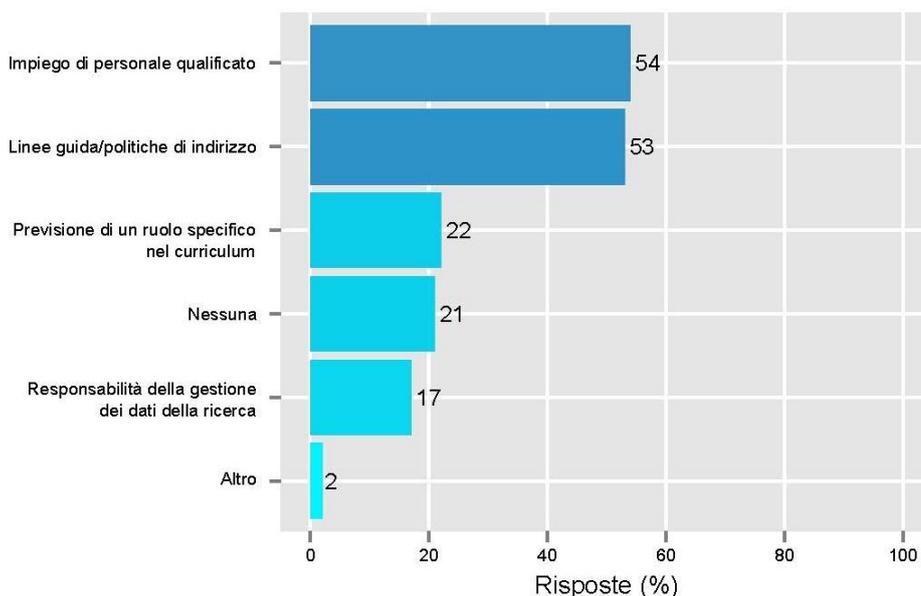
I ricercatori esprimono però con insistenza che si aspettano maggiore supporto da parte della propria istituzione per quel che riguarda la disponibilità di infrastrutture e di supporto specifico nello svolgimento di progetti, soprattutto nella gestione dei dati di ricerca. Più di un terzo (42%) dei rispondenti dichiara di desiderare maggiore assistenza legale (anche per questioni di carattere

etico), un servizio di *helpdesk* centrale e programmi di formazione da offrire non solo alla comunità scientifica, ma anche al personale che collabora con i ricercatori (Figura 8).



**Figura 8. Tipologia di supporto attesa circa la gestione dei dati, e-Infrastructures Austria 2015**

Oltre la metà (54%) degli intervistati desidera maggiore presenza di personale qualificato, nonché l'adozione di politiche e di linee guida che indirizzino verso la gestione dei dati della ricerca (53%). In un quinto (22%) delle risposte, si dichiara che le competenze di *research data management* dovrebbero diventare parte integrante del curriculum di formazione, mentre, allo stesso tempo, esperti in tal senso dovrebbero operare nei servizi di supporto offerti dall'ateneo o dall'istituzione di appartenenza (Figura 9).



**Figura 9. Ulteriori misure richieste per la gestione dei dati, e-Infrastructures Austria 2015**

## Valutazione e raccomandazioni

I risultati dell'indagine svolta in Austria con il Progetto *e-Infrastructures* vanno analizzati con un approccio analogo a quello adottato per valutare gli esiti dell'indagine BISA. Occorre infatti tenere presente il contesto di fondo europeo riferito all' EOSC, a indagini sullo sviluppo di infrastrutture, e ad analisi di impatto circa l'erogazione di ingenti finanziamenti. Va inoltre considerato che i risultati dell'iniziativa condotta per l'Austria vanno anche inseriti nel contesto dei 12 sottoprogetti attivi per tre anni nel quadro offerto e diretto dal *management* di *e-Infrastructures Austria (Work Package Clusters: <http://e-infrastructures.at/en/the-project/work-package-clusters/>)*. L'indagine austriaca è stata infatti affiancata da sondaggi di altro tipo, quali quelli che hanno preso in esame anche le infrastrutture presenti in tutte le università riguardo gli *institutional repositories* (3), nonché da ulteriori tipi di indagini, analisi, corsi di formazione, laboratori, convegni, connessi a livello nazionale e in modo capillare al progetto *e-Infrastructures Austria* (accessibili sotto *Deliverables, Activities, Activity reports* sul sito web del Progetto).

Da una valutazione complessiva dei risultati dell'indagine condotta in Austria, si possono ricavare due categorie di raccomandazioni. La prima si riferisce ai dati dei 21 atenei presi in esame dal progetto *e-Infrastructures Austria*, la seconda si lega alle prospettive di *management* dell'intero progetto. In definitiva, le misure necessarie per la gestione dei dati di ricerca possono essere riassunte come segue:

1. adozione di politiche istituzionali per la gestione dei dati della ricerca;
2. messa in opera di infrastrutture che soddisfino le necessità di tutti i comparti disciplinari;
3. pianificazione di programmi di formazione delle professionalità che rispecchino le esigenze espresse dai punti 1 e 2;
4. implementazione di servizi per la ricerca erogati da entità di supporto qualificate e riconosciute come tali;
5. applicazione di incentivi appropriati;
6. intensificazione di cooperazioni nazionali e internazionali.

Ulteriori elementi di riflessione riguardano i seguenti punti:

### 1. Sviluppo di politiche allineate per la gestione dei dati della ricerca

L'adozione di politiche istituzionali per la gestione dei dati della ricerca va adattata ai recenti sviluppi messi in atto dalla Commissione Europea (Progetto LEARN) (4). Tale intervento di adeguamento si identifica nel *policy alignment*, azione che consiste nel conformarsi a un indirizzo di *policy* già enunciato (es. politiche *Open Access*) e a posizioni già espresse in materia di gestione dei dati, di allestimento di infrastrutture, e di servizi offerti, il tutto accompagnato da misure collaterali (5). Si tratta di orientamenti già previsti nel quadro dell'EOSC. Inoltre, le politiche di gestione dei dati dovranno tenere conto delle *educational resources* e di quei dati che gravitano in area di *cultural heritage*, in quanto parte di quell'immenso patrimonio culturale di cui ogni ateneo o istituzione dispone.

### 2. Sviluppo di un quadro comune di infrastrutture con introduzione di nuovi modelli di *Good Governance*

Si evidenzia la necessità di una riprogettazione delle infrastrutture già attualmente operanti e di un coordinamento fra le loro diverse componenti. La messa in opera di infrastrutture che soddisfino le necessità di tutte le varie aree tematiche di ricerca dovrebbe sfociare in forme di *Good Governance* (come previsto nei piani EOSC) che rispecchino le politiche introdotte (6). Questo comporta anche l'adozione di nuove terminologie; non si parlerà più semplicemente di *repositories*, ma di ecosistemi digitali nei quali si inseriranno i

*repositories*. Un quarto dei ricercatori rispondenti all'indagine *e-Infrastructures Austria* lavora con *source codes*, sviluppa applicazioni, gestisce basi di dati; tutta questa galassia di strumenti e materiali, infatti, va adeguatamente gestita, ma non nei *repositories*. Il collante che manterrà unita questo insieme di elementi non sarà di solo tipo infrastrutturale, ma anche e soprattutto di tipo giuridico; gli ecosistemi digitali devono essere concepiti, dunque, come uno spazio legalizzato unico (al pari di un'area Schengen per i dati).

### 3. Rimodellamento dei workflow accademici

Occorrono termini di uso comuni (e interoperabili) per tutti quei dati che accedono ad uno spazio condiviso. L'interoperabilità prevista dall'EOSC va comunque affiancata da misure che rendano i dati conformi alle specifiche FAIR (*Findable, Accessible, Interoperable, Reusable*), tali da rendersi liberamente disponibili all'interno delle strutture/ecosistemi digitali, sempre nel rispetto delle regole di tutela della *privacy*, delle politiche di *copyright*, delle scelte da parte di chi produce i dati, circa il loro riutilizzo, e delle volontà dichiarate nei contratti di finanziamento.

### 4. Nuovi servizi per gli ecosistemi digitali

Fra questi rientrano i programmi di formazione; la parola chiave è “convergenza delle conoscenze”, in quanto si manifesta ormai una chiara esigenza di intervenire su ambiti multidisciplinari per coinvolgere tutti i soggetti attivi nella gestione dei dati. Quanto all'attività didattica, *e-Infrastructures Austria* ha organizzato nel mese di giugno 2016 una settimana formativa a livello nazionale alla quale hanno partecipato rappresentanti di biblioteche, centri di calcolo, servizi di supporto alla ricerca, esperti di materie legali afferenti a ciascuna istituzione coinvolta (<http://e-seminar.univie.ac.at/en/program/>). Sulla base del lavoro svolto nell'ambito del progetto *e-Infrastructures Austria*, si è generato anche un *Data Management Plan* (DMP) a supporto dello sviluppo di DMPs e articolato in vari schemi (7). Un'ulteriore fase in questo processo creativo di servizi per la comunità scientifica potrebbe essere quello di istituire punti di aggregazione, da cui coordinare il lavoro di supporto e formazione. Nel 2016, sulla base dell'esperienza acquisita, è stato avviato il progetto *Open Education Austria* (<http://www.openeducation.at/en/home/>) che prevede lo sviluppo e l'offerta di un gran numero di risorse per la formazione.

### 5. Ampliamento di competenze

La problematica della gestione dei dati deve collocarsi in un quadro coordinato di iniziative fondate su azioni già delineate. La riprogettazione di infrastrutture, risorse e conoscenze deve porsi, infatti, nell'alveo di direttrici già avviate: l'EOSC, i principi FAIR, l'iniziativa GO-FAIR (<https://www.dtls.nl/fair-data/go-fair/>), l'*Amsterdam Call for Action on Open Science* ([https://en.wikipedia.org/wiki/Amsterdam\\_Call\\_for\\_Action\\_on\\_Open\\_Science](https://en.wikipedia.org/wiki/Amsterdam_Call_for_Action_on_Open_Science)), gli esiti dell'*e-Infrastructure Reflection Group* (<http://e-irg.eu/>), le raccomandazioni scaturite dal progetto LEARN, l'esperienza maturata nell'ambito dell'*International Council for Science Committee on Data for Science and Technology* (CODATA <http://www.codata.org/>). È inoltre raccomandata la partecipazione critica a convegni in materia di gestione dei dati, cui far seguire una diffusione capillare delle informazioni.

### 6. Ridefinizione delle opportunità di riutilizzo dei dati

La gestione dei dati di ricerca richiede la comprensione, in tutta la sua complessità, di quattro componenti, che, seguendo le logiche proposte dall'EOSC, sono strettamente interconnesse: infrastrutture digitali, azioni di *governance* efficienti, servizi, dati.

## Conclusioni

Si premette che i risultati dell'indagine BISA e di *e-Infrastructures Austria* sono simili alle esperienze effettuate in altri contesti internazionali. In molti campi di applicazione si è raggiunto quel grado di maturità necessario, a livello di consapevolezza e di conoscenza, per raccogliere le sfide proposte dalla *Open Science Vision* della Commissione Europea (*Amsterdam Call for Action on Open Science*, aprile 2016). Gli ambiti di riferimento sono: gestione dei dati di ricerca; nesso fra e-infrastrutture, dati e servizi; supporto alla ricerca; supporto alla didattica; supporto all'internazionalizzazione. Le raccomandazioni per una gestione efficace e condivisa dei dati possono riassumersi come segue:

- partecipare a iniziative che riguardino la creazione di politiche del *management* dei dati della ricerca;
- incentivare ogni tipo di attività che promuova l'inserimento in iniziative transnazionali per lo sviluppo di infrastrutture digitali;
- colmare le carenze individuate in esito all'indagine BISA e incrementare il grado di conoscenze tramite misure di *know-how transfer of e-infrastructures essentials*;
- generare competenze per la compilazione dei *Research Data Management Plans* (RDMPs) e in merito alla gestione dei dati;
- ispirarsi a progetti di gestione e riuso dei dati conformi alle caratteristiche FAIR;
- creare i presupposti per offrire supporto ai diversi nodi di riferimento locali (es. Ufficio Area Ricerca);
- offrire servizi trasversali;
- continuare a potenziare i servizi esistenti.

## Bibliografia

1. Budroni P, Ernst K, Frasnelli E, Ganguly R, Meyer T, Seissl M (Ed.). *Researchers and their Data. Results of an Austrian Survey. Report 2015*. Vienna: Università di Vienna; 2015. Disponibile all'indirizzo: [https://phaidra.univie.ac.at/detail\\_object/o:409473](https://phaidra.univie.ac.at/detail_object/o:409473); ultima consultazione 1/8/2017.
2. Kuipers T, van der Hoeven J. *PARSE.Insight: INSIGHT into issues of Permanent Access to the Records of Science in Europe. Survey report*. PARSE.Insight, 2009. Disponibile all'indirizzo: <http://libereurope.eu/wp-content/uploads/2010/01/PARSE.Insight.-Deliverable-D3.4-Survey-Report.-of-research-output-Europe-Title-of-Deliverable-Survey-Report.pdf>; ultima consultazione 01/08/2017.
3. Bedenig C, Blumesberger S, Danowski P, Ganguly R, Gründhammer V, Haas I, Novotny G, Paril C, Preza J L, Torggler E. *Cluster A: Dokumentenserver in den e-Infrastructures Austria Partnereinrichtungen. Erhebung 2016*. Vienna: University of Vienna; 2016. Disponibile all'indirizzo: <http://phaidra.univie.ac.at/o:459231>; ultima consultazione 01/08/2017.
4. Budroni P, Sánchez Solís B, Traub-Riehle ID, Model Policy for Research Data Management (RDM) at Research Institutions/Institutes. In: LEARN (Ed.). *LEARN toolkit of best practice for research data management. Leaders Activating Research Networks*. London: LEARN; 2017. p. 131-4.
5. Budroni P, Sánchez Solís B, Traub-Riehle ID. Guidance for Developing a Research Data Management (RDM) Policy. In: LEARN (Ed.). *LEARN toolkit of best practice for research data management. Leaders Activating Research Networks*. London: LEARN; 2017. p. 135-8.
6. Budroni P. Good Governance: Strukturen zur Sicherung der Nachhaltigkeit von Projekten im akademischen Umfeld. Das nationale Projekt e-Infrastructures Austria. *ABI Technik, Zeitschrift für Automation, Bau und Technik im Archiv-, Bibliotheks- und Informationswesen* 2016;36(1):24-36.

Disponibile all'indirizzo: <https://www.degruyter.com/view/j/abitech.2016.36.issue-1/issue-files/abitech.2016.36.issue-1.xml>; ultima consultazione 01/08/2017.

7. Blumesberger S, Budroni P, Gründhammer V, Miksa T, Pausz R, Preza JL, Sánchez Solís B. *Cluster C: Template for Data Management Plans (DMP) at Austrian Research Institutions*. Vienna: University of Vienna; 2016. Disponibile all'indirizzo: <https://phaidra.univie.ac.at/view/o:459775>; ultima consultazione 24/11/2017.

## RACCOGLIERE LA SFIDA DELLA SCIENZA APERTA

Gruppo di lavoro BISA (Bibliosan per la Scienza Aperta)

Il tema della diffusione aperta della scienza coinvolge tutti gli operatori dell'informazione scientifica, da chi la produce (i ricercatori) a chi la promuove (gli organismi finanziatori), la diffonde (strutture editoriali e biblioteche di istituzioni scientifiche) mettendo in atto modelli innovativi di comunicazione scientifica, fino a chi ne definisce ruoli e responsabilità per una corretta condivisione dei risultati (i governi e i decisori politici, in genere). In definitiva, con il concetto di *open science* ci si riferisce ad un complesso articolato di strumenti e processi volti a concretizzare il concetto di trasparenza e condivisione della conoscenza scientifica.

Molte sono le implicazioni sul piano della valutazione della ricerca e dei sistemi di incentivazione che intervengono a regolare la carriera dei ricercatori; anche i sistemi di validazione, infatti, si avvantaggiano dell'applicazione di modelli aperti, nel senso che le risorse liberamente disponibili e riutilizzabili facilitano l'individuazione del sapere e dell'innovazione, anche ai fini dell'assegnazione di fondi.

Altrettante sono le sfide che emergono nei processi di creazione di nuovi servizi e infrastrutture. Con riguardo a questo ultimo aspetto, dovrà affermarsi una competenza specifica da dedicare alla realizzazione di strategie comuni, poiché la configurazione di una rete di istituzioni come Bibliosan impone di armonizzare contesti specifici che riflettono i vari profili organizzativi interni.

L'adeguamento tecnologico, l'offerta di infrastrutture, la disponibilità economica per nuovi investimenti e la maturità culturale sul fronte delle pratiche di apertura ai risultati della ricerca variano da istituzione a istituzione e soltanto affidandosi ad una politica condivisa di gestione del trattamento dei dati si potrà pervenire a un indirizzo unitario che si riveli al contempo flessibile, in risposta ad esigenze particolari delle singole comunità di ricerca. Questo punto rappresenta sicuramente una criticità da superare con il concorso coordinato di varie professionalità nell'interesse collettivo.

Sicuramente, nella realizzazione di questo processo di integrazione, è imprescindibile volgersi alle esperienze condotte a livello internazionale e seguire le iniziative di formazione, i progetti, le indagini e le raccomandazioni messi in atto da Paesi europei ed extraeuropei, anche per avviare piattaforme di dialogo e confronto.

Un solido contesto comune di riferimento per la definizione di politiche concrete a sostegno della scienza aperta è costituito dalle priorità definite dalla Commissione Europea con la creazione dello *European Open Science Cloud* che lega il concetto di conoscenza scientifica di qualità all'immagine pubblica della scienza fondata sul coinvolgimento consapevole dei cittadini. Resta da timone della buona scienza un principio ispiratore di fondo e cioè che i risultati della ricerca finanziati da fondi pubblici debbano essere fruibili a largo raggio, senza barriere, da tutti i soggetti portatori di interesse non soltanto nel recinto della ricerca, ma anche fuori dei suoi confini. In questo sforzo di allineamento e cambiamento da declinare sul vasto fronte di una scienza trasparente e condivisa, si riconferma la spinta innovativa dei bibliotecari e operatori dell'informazione scientifica nel guidare la comunità scientifica di riferimento verso la proposta e l'adozione di buone pratiche di gestione dei dati della ricerca.

Il Gruppo di lavoro BISA, tramite il questionario sui dati della ricerca distribuito tra gli enti Bibliosan, ha voluto intraprendere un percorso iniziale di esplorazione delle pratiche in uso presso le comunità di ricerca delle varie istituzioni.

È un risultato intermedio che ha comunque sollecitato spinte e aspettative alla composizione di un quadro di iniziative volte alla definizione di un documento di *policy* condiviso. In particolare, Bibliosan intenderà farsi interlocutore attivo presso le istituzioni, preparando il terreno ad una effettiva convergenza di volontà, tra Ministero della Salute e il Ministero dell'Istruzione, dell'Università e della Ricerca.



**APPENDICE**  
**Questionario BISA**



Questionario su open data per enti Bibliosan



## Questionario su open data per enti Bibliosan

Il presente questionario si pone come un percorso conoscitivo sulle pratiche di gestione dei dati della ricerca in ambito biomedico in Italia. L'obiettivo è di raccogliere tutte le indicazioni utili ad allineare la comunità di ricerca italiana agli standard internazionali di condivisione dei dati, tramite l'elaborazione di programmi e linee guida per l'adozione di pratiche comuni.

Il questionario è stato redatto dal Gruppo di lavoro Bibliosan per la Scienza Aperta (BISA)\*, istituito, in seno al sistema Bibliosan <http://www.bibliosan.it/> il 23 giugno 2016, per sondare gli aspetti relativi alla produzione e al trattamento dei dati della ricerca, in relazione a: accessibilità, deposito, riuso, aspetti legali, formati e modalità di conservazione dei dati.

I risultati dell'indagine saranno diffusi attraverso i canali del circuito Bibliosan.

\* Componenti del Gruppo di lavoro:

- Gaetana Cognetti (IRCCS, Istituto Regina Elena per lo Studio e la Cura dei Tumori)
- Moreno Curti (IRCCS Fondazione Policlinico San Matteo, Servizio di documentazione e biblioteca scientifica)
- Paola De Castro (ISS, Settore Attività Editoriali)
- Rosalia Ferrara (ISS, Biblioteca)
- Pietro La Placa (IZS Sicilia, Biblioteca)
- Cristina Mancini (ISS, Biblioteca)
- Elisabetta Poltronieri (ISS, Settore Attività Editoriali; coordinatore)
- Franco Toni (ISS, Biblioteca; osservatore/relatore in seno al Comitato di gestione Bibliosan)
- Angela Vullo (IZS Sicilia, Dip. Sanità Territoriale)

Paola Gargiulo (Cineca, OA projects and training and support Coordinator) ha revisionato il Questionario, integrando alcuni punti e offrendo suggerimenti.

Coordinatori del Gruppo:

Elisabetta Poltronieri [elisabetta.poltronieri@iss.it](mailto:elisabetta.poltronieri@iss.it) Gaetana Cognetti [gaetana.cognetti@ifo.gov.it](mailto:gaetana.cognetti@ifo.gov.it)



Pagina 1 di 8

AVANTI

## Questionario su open data per enti Bibliosan

\*Campo obbligatorio

### Dati generali (Istituzioni e ruoli professionali dei rispondenti)

Nella compilazione del questionario le domande potranno essere a risposta singola (le scelte saranno evidenziate da un "pallino"), o a risposta multipla (le scelte saranno contrassegnate da un "quadrato").

#### 1) Selezionare l'istituzione di appartenenza \*

Scegli



#### 2) Selezionare la professione \*

- Biologo
- Chimico
- Fisico
- Medico
- Veterinario
- Psicologo
- Tecnico
- Infermiere
- Bibliotecario/Documentalista
- Editore scientifico (editor)
- Altro: \_\_\_\_\_

#### 3) Selezionare la classe di età \*

- <30
- 30-50
- >50

4) Selezionare il genere \*

Uomo

Donna

5) Che tipo di dati di ricerca produci in formato digitale?  
(possibile più di una risposta) \*

Documenti di testo

Grafici/immagini

Fogli di calcolo (es. fogli Excel)

Materiale multimediale (video, foto, audio)

Banche dati

Programmi applicativi (software)

Altro: \_\_\_\_\_

6) Abitualmente, dove archivi i tuoi dati di ricerca? (possibile più di una risposta) \*

In locale, sulla mia postazione di lavoro

Su un'unità esterna - disco rigido

Su un'unità esterna - dispositivo USB

Su un'unità esterna - CD, DVD

Su un server del mio istituto

Su un archivio istituzionale

Su un'applicazione cloud

Altro: \_\_\_\_\_

7) Calcola la media annuale del volume totale dei tuoi dati di ricerca, sulla base dello spazio di archiviazione occorrente \*

- Bassa (< 50 GB)
- Media (50 – 100 GB)
- Alta (101 GB – 1 TB)
- Altissima (> 1 TB)
- Non so

8) Normalmente, ti attieni a criteri standard di descrizione dei tuoi dati di ricerca? \*

- Sì
- No
- Non so

9) Chi è il soggetto responsabile della conservazione dei dati archiviati? (possibile più di una risposta) \*

- Io stesso
- Il responsabile del progetto o del gruppo di lavoro
- Il settore informatico dell'Ente
- La Direzione scientifica dell'Ente
- Un provider esterno
- Altro: \_\_\_\_\_

10) Hai mai subito una perdita dei dati di ricerca? \*

- Sì
- No

11) Inserire un commento sulla perdita dei dati (specificare cause perdita dati, es. furto, cancellazione accidentale)

La tua risposta \_\_\_\_\_

12) Nella tua ricerca, utilizzi dati esterni ? \*

(es. non prodotti nel tuo ambito di lavoro)

Sì

No

13) In che modo utilizzi dati esterni ? \*

(es. non prodotti nel tuo ambito di lavoro)

Sì, con rielaborazione

Sì, senza rielaborazione

Non utilizzo dati esterni

14) Sei consapevole dei risvolti legali associati all'uso di dati esterni? \*

Sì

No

15) Hai avuto problemi legali utilizzando dati esterni? \*

Sì

No

16) Cosa avviene dei dati quando lasci la tua Istituzione? \*

- Restano alla mia Istituzione
- Non restano alla mia Istituzione
- Vengono cancellati
- Non so
- Altro: \_\_\_\_\_

17) Produci o utilizzi dati di ricerca sensibili o riservati? \*

- Sempre
- Talvolta
- Quasi mai
- Mai
- Non so

18) A chi è consentito l'accesso ai dati della tua ricerca? \*

- Su richiesta
- Solo all'interno dell'Istituzione
- A tutta la comunità scientifica
- A tutti
- A nessuno
- Altro: \_\_\_\_\_

19) Attraverso quale strumento gli altri possono accedere ai tuoi dati di ricerca? (possibile più di una risposta) \*

- Nessuno strumento (Accesso ai dati non consentito)
- Dispositivo di supporto dati (es. pen drive)
- E-mail
- Applicazione cloud
- Sito web
- Server remoto/in condivisione
- Archivio dati/archivio istituzionale
- Altro: \_\_\_\_\_

20) Rendi disponibili i dati per il riuso da parte di altri? \*

- Sì
- No

21) Quale tipo di licenze d'uso sono state adottate? (possibile più di una risposta) \*

- Nessun tipo di licenza (dati non riutilizzabili da altri)
- Licenze per comunità allargate di utenti
- Licenze individuali
- Licenze di libero riuso (es. CC-by)
- Non so

22) Quali incentivi potrebbero motivarti a condividere i tuoi dati di ricerca e a lasciarli liberamente accessibili? (possibile più di una risposta) \*

- Maggiore visibilità
- Rilevanza ai fini della valutazione
- Nuovi contatti/possibilità di cooperazione
- Progresso della ricerca scientifica
- Riduzione di sprechi/duplicazioni
- Incentivi economici
- Nessuno
- Altro: \_\_\_\_\_

23) Che cosa ti trattiene dal condividere i tuoi dati di ricerca con altri? (possibile più di una risposta) \*

- Impiego di tempo e costi
- Potenziale uso scorretto dei dati (es. interpretazione impropria /falsificazione)
- Violazione delle norme sulla riservatezza dei dati
- Altre restrizioni di tipo giuridico
- Uso commerciale dei dati da parte di terzi
- Timore di essere controllati
- Mancanza di motivazione
- Altro: \_\_\_\_\_

24) Che tipo di supporto ti occorrerebbe per gestire i dati della ricerca nella tua istituzione? (possibile più di una risposta) \*

- Infrastruttura tecnica
- Assistenza legale
- Servizio di helpdesk
- Corsi di formazione
- Nessuno
- Altro: \_\_\_\_\_

25) Quale ulteriore azione ti aspetti dalla tua Istituzione? (possibile più di una risposta) \*

- Linee guida/politiche di indirizzo
- Previsione di una professionalità ad hoc per la gestione dei dati della ricerca
- Altro: \_\_\_\_\_

26) Hai ulteriori commenti o suggerimenti inerenti ai dati della ricerca?

La tua risposta \_\_\_\_\_

27) Qual è il grado di soddisfazione circa i servizi resi dalla tua Istituzione nell'ambito della gestione, conservazione, diffusione, protezione dei dati della ricerca ? \*

- per nulla soddisfatto
- poco soddisfatto
- mediamente soddisfatto
- molto soddisfatto

Autorizzo il trattamento dei dati \*

Si

No

*Serie Rapporti ISTISAN  
numero di dicembre 2017, 1° Suppl.*

*Stampato in proprio  
Settore Attività Editoriali – Istituto Superiore di Sanità*

*Roma, dicembre 2017*