

Politiche e linee guida per la gestione dei dati della ricerca: l'esperienza di IOSSG

Paola Gargiulo

Cineca

Abstract. La relazione descrive brevemente il contesto in cui il gruppo di lavoro informale IOSSG – Italian Open Science Support Group si è costituito, gli obiettivi del gruppo e le attività che svolge. Il gruppo IOSSG è nato in ambito accademico, nell'ambito della promozione, dello sviluppo e la diffusione della scienza aperta in Italia, per fornire supporto e strumenti che rispondono alle sfide che la scienza aperta pone al mondo della ricerca, con particolare attenzione, ai processi della ricerca, al ciclo di vita e gestione dei dati, ai servizi e alle infrastrutture di supporto. Le attività del gruppo consistono nell'elaborare linee guida, toolkit, raccogliere buone pratiche, condurre analisi, studi, e inoltre produrre materiali e corsi di formazione sulle tematiche della scienza aperta. Il programma di lavoro suddiviso in 4 sottogruppi contempla, per l'anno 2018, quattro linee di attività: la prima si concentra sulla produzione di linee-guida, raccomandazioni per l'implementazione di politiche istituzionali sulla gestione dei dati; la seconda sullo studio di modelli per la creazione di unico punto di accesso alle diverse competenze e risorse presenti nell'istituzione, al fine di costruire un efficiente servizio di sostegno trasversale alla ricerca; la terza sulla formazione, al fine di diffondere la conoscenza e promuovere l'acquisizione di competenze su tutti gli aspetti relativi alla gestione dei dati della ricerca, una quarta sullo studio e analisi di soluzioni per la creazione di infrastrutture a livello locale per la gestione dei dati della ricerca in un'ottica aperta e federata, basata su tecnologia open source.

Keywords. Open Science, research data management, politiche e raccomandazioni, dati aperti della ricerca

Introduzione

Quando si parla di scienza aperta non si fa riferimento solo ai dati della ricerca (open research data), 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, dei suoi risultati, che viene reso accessibile e fruibile nelle modalità più aperte, della disseminazione dei risultati, della loro conservazione, ma anche della loro divulgazione con mezzi e strumenti appropriati. Implementare la scienza aperta significa dotarsi di infrastrutture, servizi ma anche di politiche, di raccomandazioni, di studi e di analisi di percorsi formativi, di best practice. Il gruppo di lavoro IOSSG intende, in questo ambito, portare il suo contributo alla diffusione e adozione della cultura della scienza aperta in Italia.

1. Il contesto

Nel promuovere e sostenere la scienza aperta, cioè un 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, la Commissione Europea, ha definito le politiche di accesso ai risultati della ricerca (pubblicazioni e dati della ricerca) prodotti in progetti di ricerca da essa finanziati. Ha stabilito, alla fine del 2013, che tutti i beneficiari dei progetti di ricerca afferenti al programma di finanziamento Horizon 2020 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). (1) Dal 2017 anche i dati della ricerca che validano i risultati presentati nelle pubblicazioni (underlying data) vanno dotati di un piano di gestione e depositati, in un data repository, resi accessibili e fruibili, secondo i principi FAIR (Findable, Accessibile, Interoperable, Reusable), principi di rintracciabilità, accessibilità, interoperabilità e riusabilità/riproducibilità dei dati della ricerca. (2)

Nell'aprile del 2016, la Commissione ha presentato una bozza di progetto su EOSC – European Open Science Cloud Initiative, un'iniziativa o per meglio dire un processo per l'implementazione in Europa della cultura dei dati e della loro condivisione e dei relativi servizi basata su un'architettura di data infrastructure commons, e su una leggera e agile governance al fine di rispondere ai bisogni dei ricercatori e all'avanzamento della scienza e della conoscenza nel contesto di una ricerca sempre più data driven. La EOSC Declaration resa pubblica nell'ottobre scorso riassume efficacemente le finalità di questo processo. (3) In alcuni paesi europei, già da un po' di anni, sono stati messi in piede servizi nazionali volti a sostenere la cultura dei dati e i suoi sviluppi, a fornire un servizio di data stewardship, cioè supporto e strumenti per l'interoperabilità, la gestione dei dati della ricerca, per la elaborazione del piano di gestione. Tali servizi sono rivolti alle istituzioni di ricerca, ai ricercatori, al personale tecnico che vi operano e coprono vari aspetti riguardanti l'intero ciclo di vita dei dati, la raccolta, l'analisi, la gestione, la cura, il trattamento e la conservazione dei dati medesimi, le licenze d'uso, così come la creazione di materiali e corsi di formazione.

In Italia la situazione si presenta molto frammentaria, sebbene esistano infrastrutture di servizi operanti a livello nazionale e comunità di ricerca che si sono organizzate nella loro disciplina, non c'è alcun coordinamento ed attualmente manca una politica a livello nazionale relativamente ai dati della ricerca. Di fatto, non viene erogato nessun servizio a livello nazionale e raramente a livello istituzionale, volto a fornire conoscenze, modelli organizzativi, strumenti necessari per affrontare queste problematiche, ad eccezione di alcune comunità scientifiche che hanno una lunga esperienza di raccolta, conservazione e gestione dei dati.

2. Nascita del gruppo IOSSG

Affrontare le sfide di una scienza sempre più data intensive è fortemente sentito nelle università italiane che si confrontano con richieste di supporto interno quando la Commissione Europea così come altri finanziatori internazionali della ricerca richiedono ai

beneficiari dei progetti di ricerca di mettere a disposizione i dati della ricerca risultanti da ricerche finanziate e permetterne l'accesso e il riuso. Diventa, allora, necessario e indispensabile che le istituzioni di ricerca dei singoli paesi, particolarmente se non sono supportati da una politica nazionale, provvedano a fornire l'assistenza e le infrastrutture necessarie ai propri ricercatori nel contesto di una scienza globale.

Per rispondere a queste esigenze e per dare delle risposte all'impatto che l'iniziativa EOSC avrà localmente, è nato nella primavera del 2016 un gruppo di lavoro in ambito universitario, un gruppo trasversale di esperti, responsabili dell'ufficio ricerca, informatici, bibliotecari e responsabili degli archivi aperti della ricerca, appartenenti al Politecnico di Milano, alle Università degli Studi di Bologna, di Milano, di Padova (prima in qualità di uditore e successivamente di membro), di Torino, di Trento, di Ca 'Foscari Venezia. Il gruppo ha deciso di collaborare informalmente, di mettere a fattore comune le conoscenze e le competenze che stavano acquisendo per dare risposte concrete e supporto ai ricercatori nella gestione dei dati e fornire raccomandazioni e linee guida alle rispettive amministrazioni.

A promuovere questa collaborazione è stato il Cineca, in qualità di NOAD - National Open Access Desk per l'Italia, del progetto europeo OpenAIRE (4) finalizzato all'implementazione delle politiche a sostegno della scienza aperta della Commissione, e che in questo ruolo, svolge anche attività di disseminazione, training, promozione e per questa ragione, il gruppo ha chiesto al NOAD di coordinarne i lavori per il primo biennio.

3. Le attività

L'obiettivo principale del gruppo è sostenere le parti interessate (amministrazione universitaria, ricercatori, uffici di ricerca, servizi IT e bibliotecari) nell'affrontare le recenti sfide che la data driven science pone alle istituzioni accademiche e di ricerca, con particolare attenzione, ai processi della ricerca, al ciclo di vita e gestione dei dati, ai servizi e alle infrastrutture di supporto.

L'elaborazione di una politica istituzionale sui dati e, in particolare, di un piano di gestione dei dati e la sua implementazione richiede affrontare in modo dettagliato:

- la descrizione dei data set,
- gli standard e i metadati che verranno utilizzati,
- gli aspetti prettamente gestionali (data management, documentazione e cura dei dati),
- la sicurezza e eventuale confidenzialità dei dati,
- la condivisione, l'accesso, il riuso,
- gli aspetti legali e di proprietà intellettuale,
- la definizione di responsabilità nelle varie fasi del ciclo dei dati.

È indispensabile affrontare in primis, il contesto politico e di governance che i dati, la fornitura di servizi e di infrastrutture a livello locale pongono, e rispondere fornendo strumenti e ponendo le basi per l'acquisizione di diverse e articolate competenze per la gestione dei dati.

Le attività del gruppo consistono nell'elaborare linee guida, toolkit, raccogliere buone pratiche, condurre analisi, studi, indagini e produrre materiali per la formazione e condi-

vedere i risultati e i materiali del lavoro di gruppo con tutte le comunità scientifiche interessate e il pubblico in generale.

Il gruppo ha iniziato le sue attività a metà 2016 ed ha già prodotto un modello di politica istituzionale di gestione dei dati di ricerca e una checklist dettagliata per supportare la stesura di un piano di gestione dei dati di ricerca, liberamente accessibili sul sito del gruppo. (5) I lavori si basano su documenti, raccomandazioni elaborate da altri atenei all'estero, in particolare, sullo scambio con l'università di Vienna e soprattutto sui materiali e il toolkit prodotto dal progetto europeo LEARN di cui la stessa università di Vienna è partner. (6) Entrambi i documenti sono in italiano e contestualizzati al quadro giuridico e di ricerca italiano.

All'inizio di settembre 2017 il gruppo ha deciso di adottare l'acronimo IOSSG, Italian Open Science Support Group, di continuare le attività e di aprire la partecipazione ad altri esperti in qualità di osservatori quali Anna Maria Tammaro del DILL- International Master in Digital Library Learning e Simone Sacchi di LIBER, ad una piccola azienda italiana attualmente impegnata in soluzioni di gestione dei dati di ricerca e di Paolo Budroni in qualità di esperto dell' E-infrastructure Austria con cui il gruppo aveva già lavorato per la stesura di un modello di policy. IOSSG è, inoltre, aperto alla collaborazione con altri gruppi simili operanti in Italia e all'estero nel mondo della ricerca e dell'università. Ha già stabilito un contatto con il corrispondente gruppo austriaco AOSSG, Austrian Open Science Support Group.

A settembre il gruppo ha definito il programma di lavoro dettagliato per il 2018 che ha come focus l'impatto a livello locale delle politiche di EOSC ed è incentrato su queste tematiche:

- elaborazione di linee guida per l'attuazione di una politica istituzionale nella gestione dei dati di ricerca ponendo attenzione alla parte giuridico-legale, in particolare alla protezione dei dati sensibili e al riuso dei dati,
- creazione di un unico punto di accesso, che aggrega competenze e risorse diverse al fine di costruire un efficiente servizio di sostegno trasversale alla ricerca,
- definizione dei requisiti dei data repository per la conformità agli attuali standard, ai principi di certificazione e ai principi FAIR, studio e proposte di soluzioni di infrastrutture a livello locale per la gestione e la conservazione dei dati della ricerca in un'ottica aperta e federata, basata su tecnologia open source,
- campagna di formazione e di promozione della scienza aperta tramite la creazione di materiali di didattica, organizzazione di workshop per la diffusione della conoscenza di tutti gli aspetti relativi alla gestione della ricerca e ai principi FAIR.

I risultati dei lavori, una volta conclusi, vengono messi a disposizione online, liberamente accessibili e riutilizzabili sul sito del gruppo. (7)

4. Conclusioni

IOSSG auspica che le attività portate avanti quest'anno e anche negli anni a venire possano dare un contributo effettivo alla creazione e alla diffusione della cultura della scienza aperta in Italia e all'adozione di un modello EOSC anche localmente, tramite la fornitura

di strumenti, modelli di politiche, best practice, studi e soprattutto favoriscano un maggior coordinamento tra i diversi attori: infrastrutture di servizi, infrastrutture di ricerche, comunità di ricerca, istituzioni universitarie e centri/ istituti di ricerca e pongano le basi per l'adozione di una visione e una politica nazionale sulla scienza aperta.

Riferimenti bibliografici

- (1) EC - 2020 Programme – Guidelines to rules on Open Access to Scientific Publications and Open Access to Research Data in Horizon 2020, vers.3.2 (March 2017) http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf
- (2) Wilkinson MD, Dumontier M, et al . (2017) The Fair Guiding Principles for scientific data management and stewardship, Nature. <https://www.nature.com/articles/sdata201618#bx2>
- (3) EOSC - <https://ec.europa.eu/research/openscience/index.cfm?pg=open-science-cloud>
- (4) OpenAIRE – <http://www.openaire.eu>
- (5) IOSSG – Modello di Policy per la gestione dei dati della ricerca e Data Management Checklist (2017) <https://sites.google.com/view/iossg/materiali-prodotti?authuser=0>
- (6) LERU Project – <http://learn-rdm.eu/>
- (7) IOSSG – <http://sites.google.com/view/iossg>

Autori



Paola Gargiulo p.gargiulo@cineca.it

Responsabile delle attività di formazione, disseminazione sui temi relativi alla scienza aperta con particolare riferimento alle politiche dell'EC in qualità di National Open Access Desk del progetto europeo OpenAIRE ,di cui Cineca è partner.